Price and wage inflation inertia under time-dependent adjustments*

Giovanni Di Bartolomeo
giovanni.dibartolomeo@uniroma1.it
Sapienza University of Rome

Marco Di Pietro
marco.dipietro@uniroma1.it
Sapienza University of Rome

June, 2014

Abstract

Our paper derives and estimates a small-scale New Keynesian model with time-dependent price and wage adjustments. Specifically, we replace Calvo mechanism with a model of price-setting featuring an upward-sloping hazard function, based on the idea that the probability of resetting a price depends on time occurred since the last reset. We obtain price and wage Phillips curves including backward-looking terms, which are, therefore, endogenously derived. We micro-found wage inflation intrinsic persistence. By Bayesian estimations, we find that our model outperforms popular price-setting alternatives in terms of log-marginal likelihoods. We also test the robustness of time-dependent adjustments to policy regime shifts. Then, we perform a normative analysis exploring the nature of macro distortions induced by our pricing model and its implications for the conduct of monetary policy.

JEL classification: E31, E32, E52, C11.
Keywords: time-dependent price/wage adjustments, Calvo pricing, intrinsic inflation inertia, hybrid Phillips curves, model comparison.

1 Introduction

Our model generalizes Erceg *et al.* (2000; EHL from now on) to time-dependent price and wage adjustments *á la* Sheedy (2007). Time-dependent models imply that a price or wage change will be more likely to be observed when last price reset happened many periods ago, i.e. the probability to reset a price is time-dependent. This mechanism can be formalized by using a hazard function, which shows the relation between the probability to post a new price and the time elapsed since the last reset: if the hazard function has a positive slope the likelihood to adjust a price is an increasing function of the time (Sheedy, 2007).[1] At our knowledge, our paper is the first attempt to micro-found wage intrinsic persistence by assuming positive hazard function and to estimate the resulting wage Phillips curve by using macrodata.

Specifcally, we derive and take to the data a small-scale New Keynesian model with time-dependent price and wage adjustments. The point of departure is given by Sheedy (2007), who replaces the Calvo price-setting assumption, in which the hazard function of price changes is

*The authors are grateful to Klaus Adam, Barbara Annichiarico, Guido Ascari, Efrem Castelnuovo, Bob Chirinko, Francesco Lippi, Peter McAdam, Ricardo Reis, Lorenza Rossi, Massimiliano Tancioni, Patrizio Tirelli, seminar participants at Macro Banking and Finance (University of Milano Bicocca) and CMC Workshop (Sapienza University, Rome) for useful comments on earlier drafts. The authors also acknowledge financial support by Sapienza University of Rome.

[1] Calvo pricing model is a particular case where the hazard function is flat, i.e. the probability to reset a price is exogenously randomly assigned to all the firms independently of the last time they have reset their prices.

flat, with a model of price-setting featuring an upward-sloping hazard function. As a result of this modification, the resulting New Keynesian Phillips curve includes backward-looking terms, which are, therefore, endogenously derived without recurring to common features as automatic indexation to past inflation in price setting. Our paper borrows Sheedy's mechanism and extends it to wage setting as well, by deriving a wage inflation equation under time-dependent wage adjustment. As a result we micro-found wage inflation intrinsic persistence. By using Bayesian techniques, we compare the empirical performance of our model to several popular alternatives based on different price and wage adjustment mechanisms, including Calvo pricing. By comparing log-marginal likelihoods of different estimations, we find that our model clearly outperforms these alternatives. Finally, following Benati (2008, 2009), we also test the robustness of time-dependent adjustments to policy regime shifts.

A price adjustment with a non-constant hazard function is considered by many papers, including Taylor (1980), Goodfriend and King (1997), Dotsey *et al.* (1999), Wolman (1999), Guerrieri (2001, 2002), Mash (2004). These models are based on state or time dependent assumptions and focus on price dynamics. We follow Sheedy (2007), based on time-dependent pricing and positive hazard functions, because his approach seems to be more able to fit the macroeconomic figures, in particular to explain inflation persistence. Differently from him, we also consider wage setting. Specifically, the attractiveness of time-dependent models with positive hazard functions is that they can provide micro-foundations for a Phillips curve exhibiting "intrinsic persistence",[2] which is a stylized economic fact hard to formalize in New Keynesian DSGE models (Fuhrer, 2011). Thus, time-dependent models are somehow alternative to the assumption of price indexation to the previous inflation rate. In fact, indexation implies the so-called "hybrid" New Keynesian Phillips curve where current inflation depends on both lagged and expected future inflation. The presence of a lagged term permits to model inflation as an auto-regressive process, where past inflation is source of structural intrinsic persistence. Anyway, inflation indexation is a solution based on *ad hoc* hypothesis, which is not always supported by the microeconometric evidence (see e.g. Dhyne *et al.*, 2005; Fabiani *et al.*, 2005).

As mentioned above, positive hazard functions can provide micro-foundations for the intrinsic persistence for inflation, which is empirically observed in macro data. However, it is worth mentioning that micro evidence on price setting about the slope of the hazard function is mixed, also if the majority of the studies now support upward hazard functions. Results depend on sample, countries, periods considered and methodologies used. For instance, Nakamura and Steinsson (2008) find that the hazard function for individual prices is initially downward sloping and then flat. By contrast, Cecchetti (1986), Goette *et al.* (2005) and Ikeda and Nishioka (2007) find strong support for increasing hazard functions. Álvarez *et al.* (2005) argue that downward hazard functions may derive from a bias due to the aggregation of heterogeneous price setters when micro data are used. Nonetheless, when are taken in account samples ranging several decades, including periods of high and low inflation, micro-evidence seems to agree that hazard functions are increasing.

Regarding wages, which are the focus of our paper, evidence for hazard function with positive slope in the U.S. is supported by the micro study of Barattieri *et al.* (2010). However, a discussion about the empirical evidence of positive hazard functions in micro-economic data is outside the scope of the present paper, which focuses on the micro-foundations for the intrinsic inertia of inflation observed in macro data. For macroeconomic issues it is important to analyse aggregate hazard since its shape affects the impulse response of aggregate variables; moreover, the relationship among micro hazard and macro dynamics is not necessary a one-to-one mapping (Yao, 2011). We refer to Sheedy (2007) or Yao (2011) for a more detailed discussion about this point.

Our model is estimated for U.S. economy with Bayesian estimation techniques. After writing

[2] Following Fuhrer (2011) by "intrinsic persistence" we refer to the inertia that does not depend on the real activity, but it is proper of the inflation process, whereas we refer to "inherited persistence" as the inertia inherited by the driving process, i.e. output gap or real marginal cost.

the model in state-space form we evaluate the likelihood function using the Kalman filter. The posterior distribution of the structural parameters is obtained combining priors with the likelihood function. The estimation of the model is performed using informative priors and, as robustness check, non-informative priors for the parameters affecting the slope of the hazard function.

In a similar paper Benati (2009) analyses several models to build inflation persistence including Sheedy (2007).[3] He finds evidence of positive-sloping hazard functions, but, by considering the Great Moderation sub-sample, he also finds that the parameters encoding the hazard slope have dropped to zero in last thirty years. He concludes that these parameters depend on the monetary regime referring to the switch in the way to conduct monetary policy discussed in Clarida *et al.* (2000). However he only focuses on price inflation: we generalize his setup by considering staggered wages with possible time-dependent adjustment process in the labor markets. Stickiness and persistence in wages may have important implications for both inflation persistence and monetary policy effects. Then, by considering a sub-sample, we check if the hazard function remains strictly positive during the Great Moderation in our framework.

Moreover, following Rabanal and Rubio-Ramirez (2005), we also compare the performance of our model to others based on alternative specifications for price and wage adjustments. Our goal is to test the improvement in explaining the data, in terms of marginal likelihood, due to our mechanism to micro-found inflation persistence. Specifically, as alternatives we consider flat hazard functions (price and wage Phillips curves *á la* Calvo) with indexation, which is a popular assumption to take account for inflation persistence (see Galí and Gertler, 1999; Christiano *et al.*, 2005).

Then, we move to a normative analysis by using the linear-quadratic (LQ) approach proposed by Rotemberg and Woodford (1997) and further extended by, among others, Woodford (2002, 2003), Benigno and Woodford (2003, 2005, 2012) to study the welfare effects associated with intrinsic inflation persistence generated by a time-dependent pricing model. Specifically, we derive a LQ model that is consistent with the assumption that pricing decisions depend on the time spent from the last price reset, which in turn generates intrinsic inflation persistence, and study welfare effects and policy implications of this micro-foundation of inflation inertia. We assume the presence of an output or employment subsidy that offsets the distortion due to the market power of monopolistically-competitive price-setters, so that the steady state under a zero-inflation policy involves an efficient level of output. We consider an approximation around an efficient steady state to compare our results to those of Steinsson (2003), who investigates an alternative model of intrinsic inflation persistence under this assumption.[4]

We use our welfare approximation to explore the nature of macro distortions induced by our time-dependent pricing model. Our approximation in fact makes explicit how welfare costs are generated by time-dependent pricing adjustments. These costs are distinct from those associated with relative price dispersion and consumption fluctuations that appear in the standard New Keynesian model.

In the Calvo price setting, where the probability of resetting a price is time-independent, all firms have the same probability of changing their prices independently of their history, i.e. the time elapsed since their last update. Distortions are only related to one feature of the price setting mechanism, i.e. the (average) probability of resetting a price, which is equal to the probability of individual firms. Instead, our environment is richer as distortions can be associated to different average probabilities, as in Calvo, but also to different distributions of the probability of resetting price, since this may be not the same among firms. In other words, our setup generalizes Calvo (1983) and here we can disentangle the distortions associated to different average from those stemming from different distributions of the probability of resetting prices.

Moreover, we investigate optimal policy implications of time-dependent pricing settings. We

[3] Specifically, Benati (2009) analysed Fuhrer and Moore (1995), Galí and Gertler (1999), Blanchard and Galí (2007), Sheedy (2007), Ascari and Ropele (2009).

[4] The approach can be however generalized to the case of a distorted steady state (Benigno and Woodford, 2005, 2012).

look at how different hazard rates affects gains from commitment with respect to discretion and compare our implications for optimal policy to those derived from an alternative model for intrinsic inflation persistence based on a bounded-rational behavior—obtained by Steinsson (2003) within the same LQ approach.

Steinsson (2003) considers a model under the assumption that a fraction of the producers in the economy set their prices according to a rule of thumb, i.e. using indexation to past prices, which generates intrinsic inflation persistence. As a result, he obtains a Phillips curve that is a convex combination of a forward-looking term and a backward-looking term. As long as the relative importance of the backward term relatively to the forward one increases, he finds that recessions (expansions) are more persistent and gradual policies prevail on immediate overshooting that characterizes the purely forward-looking case under commitment, and, moreover, gains of commitment over discretion fall. Similar results can be derived by assuming that the fraction of firms not assigned to re-optimize will index their prices to the aggregate inflation of the previous period; indexation, in fact, leads to a similar reduced form Phillips curve. Both specifications however require some *ad hoc* assumptions about the price-setting process to generate the dependence of inflation on past values of inflation (Sbordone, 2007).

Clearly, as in our case, Steinsson's (2003) results are specific on the way intrinsic inflation persistence is modeled; therefore, it is important to check them compared to alternative formalizations. In contrast to Steinsson (2003), we find that commitment and inflation persistence are not in opposition. In a world with persistence, commitment gains over discretion increases in inflation inertia generated by a more asymmetric distribution of reset probabilities among firms conditional to the time elapsed since the last spell.

Finally, inspired by Sbordone (2007), we study the consequences of implementing "wrong" monetary policy due to a misinterpretation of sources of inflation persistence by assuming a considerable amount of uncertainty concerning the correct specification of the model representing the economy (model uncertainty). We consider both policy regimes (discretion and commitment). Afterwards, we also look at robust policies by using robust control techniques.

The main contributions of our paper are four and can be summarized as follows. First, the time-dependent adjustment proposed by Sheedy (2007) is extended to the wage-setting process; we thus derive an analytical solution for the wage Phillips curve with time-dependent adjustment providing micro-foundations for wage-inflation persistence, which in turn allow to obtain hump-shaped response in wage inflation to a cost-push shock, which does not emerge in the case of indexation to past-price inflation. Second, by estimating a model similar to Benati (2009) including time-dependent wage setting, we find that parameters encoding intrinsic persistence remain significantly different from zero also during the Great Moderation sub-sample–so they are "deep" in the sense of Lucas. Three, by comparing marginal likelihoods, we find that our model outperforms alternative specifications for price and wage adjustments, i.e. Calvo with indexation and Calvo augmented by Galí-Gertler mechanism for prices and wages. Finally, we investigate what are the normative impliations of considering a time-dependent adjustment for price setting.

Our results are quite robust. Among others, we successfully test their robustness by considering non-informative priors for the parameters affecting the intrinsic component of inflation inertia. Many other robustness checks are mentioned in the paper and available upon request. We also show that estimations are consistent with positively sloped hazard functions, both for prices and for wages.

The rest of the paper is organized as follows. In the next section, after introducing Sheedy mechanism, we consider a simple small-scale model characterized by price and wage Phillips curve able to account for intrinsic inflation persistence. Section 3 presents ours model estimations, whereas Section 4 compares them to EHL and its extension with different kind of inflation indexation. Section 5 shows the derivation of the welfare function when inflation inertia is achieved via a reset price probability that is time-dependent. We explore the nature of macro distortions induced by our pricing model and its implications for the conduct of monetary policy. We analyze how a change in the hazard slope affects the welfare gain of a commitment over discretion. Furthermore, we consider the cost for the Government of misunderstanding the sources of persistence in implementing its optimal policy. A final section concludes.

2 The model

Our model generalizes EHL (2000) by assuming that price and wage adjustments are governed by a time-dependent mechanism. In order to improve the empirical realism of our model we have also considered habit formation, which implies persistence in the IS curve. As we mainly differ from EHL (2000) for the derivation of the Phillips curves, the description of the demand side of the economy is not detailed. For a full derivation of the model, we refer to EHL (2000).[5] We report the log-linear deviations from the steady state.

2.1 Hazard function and Phillips curves

According to Sheedy (2007),[6] the probability to adjust a price is not random as in Calvo specification, but depends on the time elapsed since last price reset. This means that the probability to change a price is not equal among firms, but it is positive function of the time. Formally, price and wage adjustments are defined by using a hazard function, which expresses the relationship between the probability to reset a price and the duration of price stickiness. The hazard function is specified as follows:

$$\alpha_i = \alpha + \sum_{j=1}^{\min(i-1,n)} \varphi_j \left[\prod_{k=i-j}^{i-1} (1-\alpha_k) \right]^{-1}, \tag{1}$$

where α_i is the probability to change a price which last reset was i periods ago; α is the initial value of the hazard function, φ_j is its slope; n is the number of parameters that control the slope – for $n = 1$, the slope is governed by only one parameter, $\varphi_j = \varphi$.[7]

By using (1), as shown by Sheedy (2007), we can derive a price Phillips curve that depends on both expected[8] and past inflation. Formally:

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

where π_t^p is the price inflation rate and mc_t is the real marginal cost; β is the stochastic discount factor, ζ_t is a price mark-up shock; the coefficients ψ_p and k_p are function of the parameters characterizing the hazard function:

$$\begin{cases} \psi_p = \frac{\varphi_p}{(1-\alpha_p)-\varphi_p[1-\beta(1-\alpha_p)]} \\ k_p = \frac{(\alpha_p+\varphi_p)\left[1-\beta(1-\alpha_p)+\beta^2\varphi_p\right]}{(1-\alpha_p)-\varphi_p[1-\beta(1-\alpha_p)]}\eta_{cx} \end{cases} \tag{3}$$

Parameters φ_p and α_p characterize the hazard function: the former controls the slope and the latter the starting level (i.e. φ and α in (1)); $\eta_{cx} = \frac{1-\phi}{1-\phi+\phi\varepsilon_p}$ is the elasticity of a firm's marginal cost with respect to average real marginal cost, where $1-\phi$ is the labor share and ε_p is the elasticity of substitution between workers. The elasticity η_{cx} is derived from a simple Cobb-Douglas production function without capital:

$$y_t = a_t + (1-\phi)\, n_t, \tag{4}$$

[5] Details are also available upon request.

[6] Note that the hazard function used in by Sheedy (2010) is an equivalent reparametrization of Sheedy (2007). Both hazard functions lead to the same Phillips curve specification.

[7] For the sake of simplicity, we follow Sheedy (2007) using $n = 1$.

[8] Both 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 captures the dynamics of the adjustment process. See Sheedy (2007) for a discussion.

where y_t denotes output, a_t is the technology shock and n_t is the amount of hours worked.

The real marginal cost is given by:

$$mc_t = \omega_t + n_t - y_t, \tag{5}$$

where ω_t denotes the real wage.

By definition, the real wage dynamics is described by:

$$\omega_t - \omega_{t-1} = \pi_t^w - \pi_t^p. \tag{6}$$

The marginal rate of substitution, mrs_t, between consumption and hours worked is given by:

$$mrs_t = \frac{\sigma}{1-h}\left(y_t - hy_{t-1}\right) + \gamma n_t - g_t, \tag{7}$$

where σ denotes the relative risk aversion coefficient, h is an internal habit on consumption and g_t denotes a preference shifter shock. Since the labor market is characterized by imperfect competition the difference between the real wage and the marginal rate of substitution is equal to the wage mark-up:

$$\mu_t^w = \omega_t - mrs_t. \tag{8}$$

One novelty of our paper is to derive a New Keynesian wage Phillips curve that exhibits intrinsic inflation persistence from the hazard function. Formally:[9]

$$\pi_t^w = \psi_w \pi_{t-1}^w + \beta\left[1 + (1-\beta)\psi_w\right] E_t \pi_{t+1}^w - \beta^2 \psi_w E_t \pi_{t+2}^w - k_w \mu_t^w, \tag{9}$$

with

$$\begin{cases} \psi_w = \frac{\varphi_w}{(1-\alpha_w)-\varphi_w[1-\beta(1-\alpha_w)]} \\ k_w = \frac{(\alpha_w+\varphi_w)\left[1-\beta(1-\alpha_w)+\beta^2\varphi_w\right]}{(1-\alpha_w)-\varphi_w[1-\beta(1-\alpha_w)]}\Xi_w \end{cases}, \tag{10}$$

where π_t^w is the wage inflation, ψ_w and k_w are coefficients depending on the hazard parameters (as in the case for prices, φ_w and α_w control respectively the slope and the initial level of the hazard function); $\Xi_w = \frac{1}{1+\varepsilon_w\gamma}$, where ε_w denotes the elasticity of substitution between workers and γ is the inverse of the Frisch labor supply elasticity.

Inspecting equation (9) is clear that the mechanism studied in our paper provides micro-foundations for lagged terms in wage inflation equations. It is worth noticing that it micro-founds persistence directly related to wage inflation rather than price inflation as in model with indexation to past values of prices.

2.2 Closing the model: IS curve and Taylor rule

The model is closed by introducing the demand side of the economy and the monetary policy rule. The demand side (IS curve) is obtained by log-linearizing the Euler equation around the steady-state, formally:

$$y_t = \frac{1}{1+h}E_t y_{t+1} + \frac{h}{1+h}y_{t-1} - \frac{1-h}{\sigma(1+h)}\left(i_t - E_t\pi_{t+1}^p + E_t g_{t+1} - g_t\right), \tag{11}$$

where i_t is the nominal interest rate set by the central bank and lagged terms are due to the presence of internal-consumption habits.

Monetary policy is modelled as a simple Taylor rule:

$$i_t = \rho_r i_{t-1} + (1-\rho_r)\left(\delta_\pi \pi_t^p + \delta_x y_t\right) + \upsilon_t, \tag{12}$$

[9] Equation (9) is derived in Appendix A.

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; υ_t is a monetary policy shock.

Aside from the monetary disturbance,[10] all the shocks considered in the model follow an $AR(1)$ process:

$$\begin{cases} a_t = \rho_a a_{t-1} + \varepsilon_t^a, \\ g_t = \rho_g g_{t-1} + \varepsilon_t^g, \\ \zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_t^\zeta, \\ \upsilon_t = \varepsilon_t^\upsilon, \end{cases} \tag{13}$$

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

3 Empirical analysis

We estimate our model by Bayesian techniques. Our choice is motivated by the fact that Bayesian methods outperform GMM and maximum likelihood in small samples.[11] After writing the model in state-space form, the likelihood function is evaluated using the Kalman filter, whereas prior distributions are used to deliver additional non-sample information into the parameters estimation: once a prior distribution is elicited, posterior density for the structural parameters can be obtained reweighting the likelihood by a prior. The posterior is computed via numerical integration by making use of the Metropolis-Hastings algorithm for Monte Carlo integration; for the sake of simplicity all structural parameters are assumed to be independent from each other.

We use four observable macroeconomic variables: real GDP, price inflation, real wage, nominal interest rate. The dynamics is driven by four orthogonal shocks, including monetary policy, productivity, preference and price mark-up; since the number of observable variables is equal to the number of exogenous shocks the estimation does not present problems deriving from stochastic singularity.[12] The estimation of the model is performed by using informative priors and, as robustness check, non-informative priors for the parameters characterizing the slope of the hazard function.

We aim to test if the model exhibits positive hazard function, i.e. time-dependent price/wage adjustments holds. After estimating our model for the full sample (1960:1-2008:4), we also consider a smaller one (1982:1-2008:4), representative of the Great Moderation, in order to investigate if a positive hazard function still holds in a period characterized by small volatility of the shocks and more aggressive central bankers in fighting inflation. By considering only time-dependent price adjustment and flexible wages, Benati (2009) showed that during the Great Moderation, the parameters encoding the structural component of inflation persistence have dropped to zero.

Finally, we evaluate the empirical performance of our time-dependent Phillips curves to alternative specifications commonly used in literature. We consider the traditional forward-looking Phillips curves derived in EHL (2000) extended with price and wage indexation, which is often claimed as one main assumption to account for inflation persistence. Model comparison is based on log-marginal likelihood. In order to apply this methodology, we will show how models compared here are nested.

Next subsection presents the data used and prior distributions. Subsection 3.2 provides the estimation for the baseline model. Subsection 3.3 evaluates our time-dependent model against alternative specifications.

[10] Monetary policy persistence is already captured by the lagged term in (12). We have however successfully checked the robustness of our result with respect to alternative assumptions. Specifically, we have considered an $AR(2)$ process for the interest rate in equation (12). Results are available upon request.

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

[12] Problems deriving from misspecification are widely discussed in Lubik and Schorfheide (2006) and Fernández-Villaverde (2010).

3.1 Data and prior distributions

In our estimations, we use U.S. quarterly data. All the time series used come from FRED database maintained by the Federal Reserve Bank of St. Louis. The *real gross domestic product* is used as measure of the output; 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. 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; output and real wage are detrended by using the Baxter and King's bandpass filter.

Our choices about 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 and follow a Normal distribution. These values are quite standard in the literature. The smoothing parameter is assumed to follow a Beta distribution with mean 0.6 and standard deviation equal to 0.2. The same choice has been made for the consumption habit. The inverse of Frisch elasticity is a tricky parameter to estimate: our choice is based on a Gamma distribution with mean 2 and standard deviation 0.375.

For the hazard function coefficients we perform an "informative estimation" by using as priors coefficients estimated from single equation GMM;[13] we assign a Normal distribution to φ_p and φ_w with standard deviation equal to 0.2, whereas α_p and α_w follow a Beta distribution with standard deviation 0.1. As robustness check, following Benati (2009), we also estimate the model by using non-informative priors for the parameters affecting the slope of the hazard function, instead of those derived from the GMM estimations. Differently from him, we use a Uniform distribution with support $[-1, 1]$: the choice of such a large interval is motivated by the fact that we want to investigate if the hazard slope is positive, negative or zero.

We need to calibrate some parameters in order to avoid identification problems.[14] Since 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. Price elasticity is calibrated following Sheedy (2007), to be coherent with the hazard priors derived from his GMM estimation. Wage elasticity is derived as in Galí (2011) by using $\varepsilon_w = [1 - \exp(-\gamma u^n)]^{-1} = 8.85$, where we assume $\gamma = 2$ and a natural unemployment rate u^n equal to 6%, as the average rate of the period considered. Finally, all the autoregressive coefficients of the shocks follow a Beta distribution with mean 0.5 and standard deviation equal to 0.2. The prior for the shocks standard deviations is an Inverse Gamma with mean 0.01 and 2 degrees of freedom.

3.2 Estimation results

Our baseline model consists of six equations, describing: the production function (4); the real marginal cost (5); the real wage dynamics (6); the marginal rate of substitution (7); the dynamic IS (11); the Taylor rule (12). Two additional equations close the model: the price and wage Phillips curve. In our baseline estimation we consider the time-dependent form for both price and wage equation, i.e. equations (2) and (9). Shocks dynamics are described by (13).

Our estimations are reported in Table 1, which also summarizes the 90% probability intervals and our assumptions about the priors. The table describes the results for the full sample and the Great Moderation. We report posterior estimation of the shocks and structural parameters, obtained by Metropolis-Hastings algorithm, when informative priors for the hazard slope are used.

[13] We estimate φ_w and α_w by using GMM as Sheedy (2007) does for the price adjustment (details are provided by Appendix B). For the hazard characterizing price adjustment we directly use as priors the GMM estimates of Sheedy (2007).

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

Table 1 – Prior and posterior distributions[15]

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (Great Moderation)		
	Density	Mean	St. Dev.[16]	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_υ	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, since φ_p and φ_w are both positive. Thus, time-dependent mechanism seems to be able to account for inflation inertia for both prices and wages. The duration of a price spell is 3.7 quarters, whereas wages appear to be less sticky, since their duration is 2.05 quarters.[17] The response of monetary authority to inflation and output gap is in line with the Taylor principle; the estimated degree of interest rate smoothing is 0.82. All the shocks exhibit a high degree of autocorrelation, near to 0.8. With regard to the parameters characterizing the utility function (i.e. habit, relative risk aversion and inverse of Frisch elasticity), their estimations are coherent with the standard findings in literature.

By considering the Great Moderation sub-sample, differently from Benati (2009), we find that hazard function continues to exhibit positive slope, since both φ_p and φ_w are positive. This result gives us evidence that a pricing mechanism based on hazard function still holds also in a period characterized by a central bank more concerned in fighting inflation, as highlighted by the higher estimated coefficient for δ_π. As a result intrinsic persistence also holds for the Great Moderation period. Price duration is now 4.5 quarters: this is highlighted by the fact that the hazard function sloping is still positive, but smaller. This fact is in line with macroeconomic theory: as from the Great Moderation inflation has dropped, the cost of not adjusting a price is smaller and this translates in longer price spell. By contrast, computed wage stickiness is rather stable reflecting the fact that wage bargaining is more influenced by institutional factors.[18]

In Figure 1 we plot prior distribution, posterior distribution and posterior mode of the estimated parameters.

[15]The posterior distributions are obtained using Metropolis-Hastings algorithm; the procedure is implemented using the Matlab-based Dynare package. Mean and posterior percentiles come from two chains of 250,000 draws each from Metropolis-Hastings algorithm, where we discarded the initial 30% draws.

[16]For the Inverse Gamma distribution the degrees of freedom are indicated.

[17]The durations (D) of price and wage stickiness are computed by using the following relation: $D = \frac{1-\varphi}{\alpha+\varphi}$ (see Sheedy, 2007).

[18]Considering a Walrasian labor market, as in Benati (2008, 2009), may force the estimated-price Phillips curve to capture also wage stickiness present in the data. This leads to an overestimation of price duration, which, in the Great Moderation subsample, drops to zero the hazard coefficients implying a quite flat hazard function and no intrinsic persistence in price inflation.

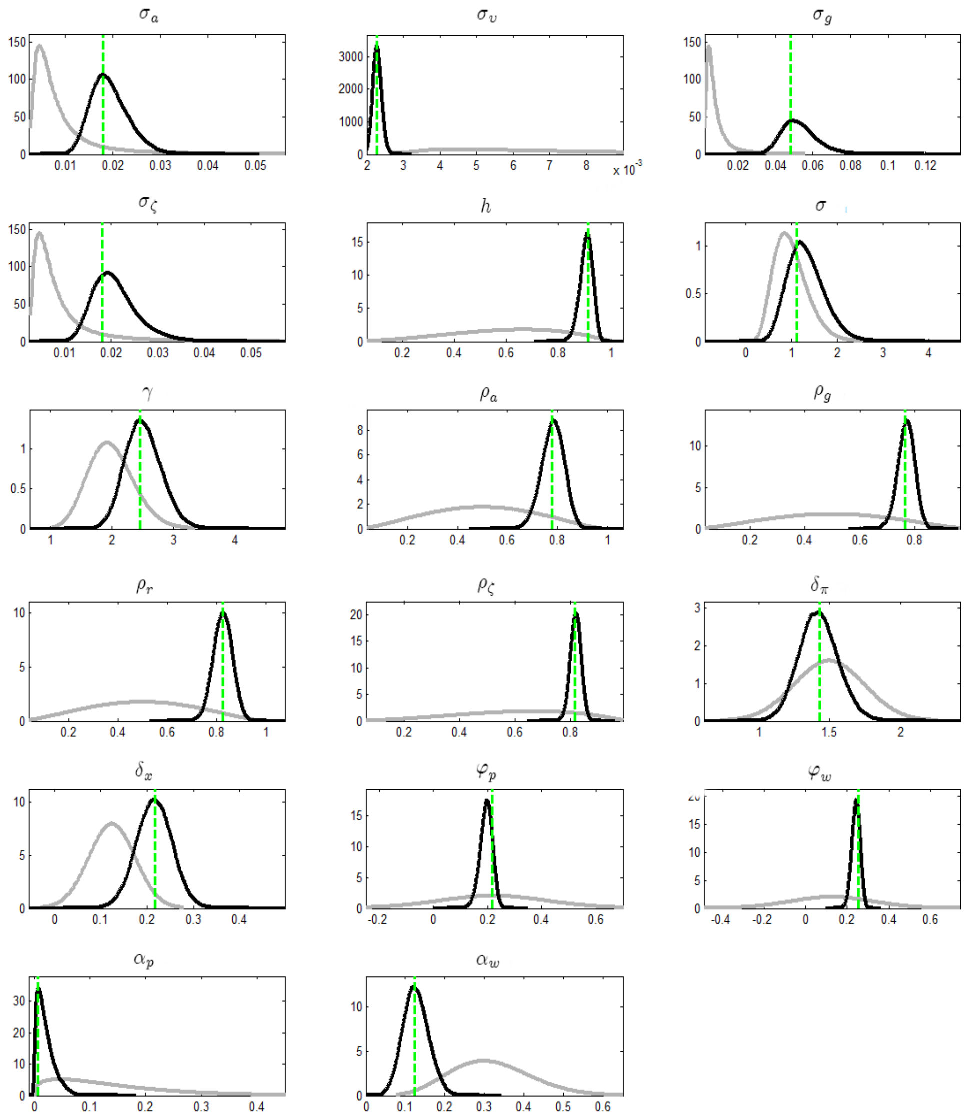

Figure 1 - Prior distribution (grey curve), Posterior distribution (green curve) and Posterior mode (dotted line) of the estimated parameters.

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 variables used, shock specifications. Thus, we have checked the robustness of our analysis by considering also uniform priors for the parameters φ_p and φ_w with support $[-1;1]$,[19] whereas the prior distributions for the remaining parameters are the same used previously. Results are provided in Table 2. The

[19] Choosing this large support we can test if the hazard slope is negative, positive or flat. The prior mean is centered on 0.

"non-informative" estimation confirms our results about the hazard function, which is still characterized by positive slope, both in full sample and during the Great Moderation; the estimated parameters for the hazard slope are very similar to the ones estimated under "informative" priors. This result shows as the hazard function mechanism is robust to a change of policy. We have also successfully checked the robustness of our results by considering different model specification (i.e., model without habit), various specifications for monetary shocks (as already mentioned) and alternative series for observable variables.[20]

Table 2 - 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_υ	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

4 Time-dependent Phillips vs. alternatives

In this section we aim to compare the empirical performance of our time-dependent Phillips curves to different specifications for price and wage adjustments. In our framework this can be easily done as the model encompasses several alternatives. Simply by setting $\varphi_p = 0$ and $\varphi_w = 0$, we obtain flat hazard functions, and therefore, price and wage Phillips curves *á la* Calvo. Moreover, different kinds of indexation can be introduced by minimal manipulations. In the following we show how to derive the EHL (2000) Phillips curves from our model and augment them by indexation and then we compare these alternatives to our baseline model in terms of log-marginal density.

4.1 Alternative price-setting mechanisms: EHL with indexation

It is easy to verify that the price Phillips curve (2) nests the EHL case. Assuming $\varphi_p = 0$, we get:

$$\pi_t^p = \beta E_t \pi_{t+1}^p + \lambda_p \left(mc_t + \zeta_t\right) \tag{14}$$

[20] In particular, we have considered a different measure for prices by using the *nonfarm business sector implicit price deflator*. With regard to wages we have considered alternatives measures given by: *average hourly earnings of production*; *business sector compensation per hour*; *hourly earnings for manufacturing sector*. Our results are available upon request.

where $\lambda_p = \frac{\alpha_p[1-\beta(1-\alpha_p)]}{1-\alpha_p}\eta_{cx}$. Equation (14) can be also augmented by indexation:

$$\pi_t^p = \frac{\iota_p}{(1+\iota_p\beta)}\pi_{t-1}^p + \frac{\beta}{1+\iota_p\beta}E_t\pi_{t+1}^p + \lambda_p^\iota (mc_t + \zeta_t) \tag{15}$$

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

Similarly, equation (9) nests the EHL case for $\varphi_w = 0$:

$$\pi_t^w = \beta E_t\pi_{t+1}^w - \lambda_w\mu_t^w \tag{16}$$

where $\lambda_w = \frac{\alpha_w[1-\beta(1-\alpha_w)]}{1-\alpha_w}\Xi_w$. It can be augmented by indexation:

$$\pi_t^w = \iota_w\pi_{t-1}^p - \iota_w\beta\pi_t^p + \beta E_t\pi_{t+1}^w - \lambda_w\mu_t^w \tag{17}$$

where ι_w denotes the degree of wage indexation to last period's inflation.

4.2 Galí-Gertler setting

A further specification to account for inflation persistence has been introduced by Galí and Gertler (1999). They proposed a modification of the Calvo mechanism by introducing partial indexation due to a backward looking rule of thumb. The Phillips curves are specified as follows:

$$\pi_t^p = \frac{\xi_p}{\Lambda_p}\pi_{t-1}^p + \frac{\beta(1-\alpha_p)}{\Lambda_p}E_t\pi_{t+1}^p + \lambda_p^\xi (mc_t + \zeta_t) \tag{18}$$

$$\pi_t^w = \frac{\xi_w}{\Lambda_w}\pi_{t-1}^p + \frac{\beta(1-\alpha_w)}{\Lambda_w}E_t\pi_{t+1}^w - \lambda_w^\xi\mu_t^w \tag{19}$$

where ξ_p measures the degree of price indexation to past inflation, ξ_p denotes the degree of wage indexation to past inflation, $\Lambda_p = 1-\alpha_p+\xi_p[\alpha_p + (1-\alpha_p)\beta]$, $\Lambda_w = 1-\alpha_w+\xi_w[\alpha_w + (1-\alpha_w)\beta]$, $\lambda_p^\xi = \frac{\alpha_p(1-\xi_p)[1-\beta(1-\alpha_p)]}{\Lambda_p}$ and $\lambda_w^\xi = \frac{\alpha_w(1-\xi_w)[1-\beta(1-\alpha_w)]}{\Lambda_w(1+\varepsilon_w\gamma)}$.

4.3 Model comparison

As shown above, our formalization nests different models of price and wage adjustment. Differences only depend on the Phillips curve parameterization. By different assumptions on φ_p, φ_w, ι_p, ι_w, ξ_p, ξ_w, we can consider positive hazard functions or flat hazard functions augmented by two different kind of indexation. We compare our baseline (BASE) to two alternative scenarios:[21]

1. EHL model with indexation (EHLind), by considering (15) and (17);
2. EHL model with indexation *á la* Galí-Gertler (GG), by considering (18) and (19).

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. The aim is to evaluate if the way in which is modeled price and wage adjustment affects the fit a model. The model with the highest log-marginal likelihood better explains the data.[22] Table 3 reports our results.

[21] We omit the comparison with a model characterized by simple forward-looking Phillips curves *á la* Calvo since this model has not intrinsic persistence. Anyway, Rabanal and Rubio-Ramirez (2005) showed that this model exhibits quite the same performance of a model with indexation.

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

Table 3 - Log-marginal data densities and Bayes factors for different models[23]

Model	Log-marginal data density	Bayes factor vs. BASE
BASE	3615.6	
BASE (non-info)	3613.6	$\exp[-2.0]$
EHLind	3569.7	$\exp[-45.9]$
GG	3564.8	$\exp[-50.8]$

The difference, in terms of marginal likelihood, between Galí-Gertler specification and EHL augmented by indexation is minimal. According to Jeffreys' scale of evidence,[24] this difference must be considered as "slight" evidence in favor of EHLind with respect to GG. However, our model clearly outperforms both the alternative considered: in particular, Bayes factor gives "very strong" evidence in favor of our specification. This means that the pricing method based on hazard functions seems to capture better inflation inertia. Under "non-informative" priors we observe a slight decrease of the marginal likelihood: this happens since under diffuse priors there is an increase of model complexity and this penalizes the marginal data density (this effect dominates the improvement in model fit).

In the comparison between our time-dependent adjustment model and the models with indexation or rule-of-thumb firms, the fit of the different models is judged by looking at marginal likelihood comparisons. Notwithstanding all inflation equations include a combination of backward-looking and forward-looking terms for the main dependent variables which may imply similar reduced forms, the differences are significant (as evidenced from the Bayes factors). Large improvements are mainly due to the fact that our mechanism provides micro-foundations for nominal wage persistence in the wage-inflation equations. This does not occur under wage indexation because this is usually associated to a price index. As a result our wage Phillips curve better captures wage dynamics. In particular, it is able to replicate hump-shaped response in wage inflation. Figure 2 shows the wage IRFs to a price mark-up shock based on estimated version of the three alternative specifications.

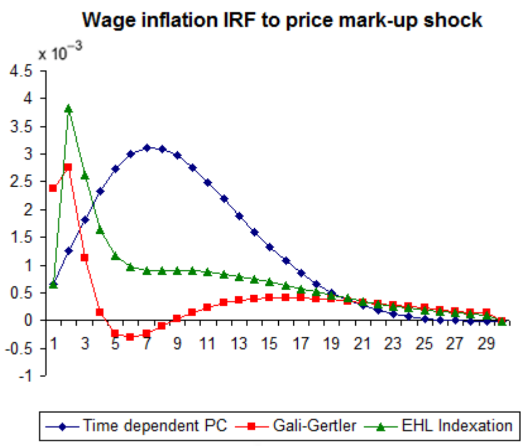

Figure 2 - IRFs of the wage inflation to 1% price mark-up shock for different model specifications.

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

[24] Jeffreys (1961) provided a scale for the evaluation of the Bayes factor indication. Odds ranging from 1:1 to 3:1 give "very slight evidence;" odds ranging from 3:1 to 10:1 constitute "slight evidence;" odds ranging from 10:1 to 100:1 constitute "strong to very strong evidence;" odds greater than 100:1 give "decisive evidence."

It is clear as a model based on time-dependent Phillips curves is able to exhibit the hump-shaped response of wage inflation to a shock, whereas the others do not. This is due to the fact that for wage inflation, under time-dependent pricing models, the persistence component is "intrinsic"; this fact explains the large difference between models highlighted by the Bayes factor.

5 Normative analysis

5.1 Welfare approximation

In what follows we consider what are the normative implications of considering a time-dependent adjustment in price setting. In order to compare our results with those of Steinsson (2003) we consider a walrasian labor market. Our price Phillips curve, as usual, is given by:

$$\pi_t = \Psi_1 \pi_{t-1} + \beta \Psi_2 E_t \pi_{t+1} - \beta^2 \Psi_1 E_{t+1} \pi_{t+2} + \kappa x_t + u_t \tag{20}$$

The welfare loss derives from a second-order Taylor approximation of the representative household's utility function and mainly depends on the form that takes the price dispersion. The details of the approach developed by Rotemberg and Woodford (1997, 1999) are widely discussed in Woodford (2003: Chapter 6), Galí (2008: Chapter 4), Benigno and Woodford (2012). In the derivation we follow Galí (2008: Chapter 4), from whom we only differ in the price dispersion evolution, which in our case relies on the use of time-dependent price setting with non-constant hazard functions.

We assume a utility function taking the following separable form:

$$U_t = \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\gamma}}{1+\gamma} \right) \tag{21}$$

where C_t is consumption; σ and γ are parameters.

A second-order approximation of (21) is:

$$U_t - U \simeq U_c Y \left(y_t + \frac{1-\sigma}{2} y_t^2 \right) + U_n N \left(n_t + \frac{1+\gamma}{2} n_t^2 \right) \tag{22}$$

where $\sigma = -\frac{U_{cc}}{U_c} Y$, $\gamma = \frac{U_{nn}}{U_n} N$. Note to obtain (22), we have used the aggregate resource constraint $C_t = Y_t$.

By integrating the production function, $Y_t = \int_0^1 Y_t(i) di = A_t \int_0^1 N_t(i)^{1-\delta} di = A_t N_t^{1-\delta}$, we can manipulate the demand function of good i to obtain:

$$\left(\frac{Y_t}{A_t} \right)^{\frac{1}{1-\delta}} \underbrace{\int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{\frac{-\varepsilon_p}{1-\delta}} di}_{D_t} = N_t \tag{23}$$

where D_t measures the degree of price dispersion.

Log-linearizing (23) around the steady state implies:

$$(1-\delta) n_t = y_t - a_t + \frac{\varepsilon_p}{2\Theta} var_i \{p_t(i)\} \tag{24}$$

since up to a second-order approximation $(1-\delta) d_t \simeq \frac{\varepsilon_p}{2\Theta} var_i \{p_t(i)\}$, where $\Theta = \frac{1-\delta}{1-\delta+\delta\varepsilon_p}$.

As we consider an efficient steady state, thus $-U_n/U_c$ equals the marginal product of labor in the steady state, $(1-\delta) Y/N$, by substituting (24) into (22), after some simple algebraic manipulations, we obtain:

$$\frac{U_t - U}{U_c Y} \simeq -\frac{1}{2} \left[\frac{\varepsilon_p}{\Theta} var_i \{p_t(i)\} + \left(\sigma + \frac{\gamma+\delta}{1-\delta} \right) x_t^2 \right] + t.i.p. \tag{25}$$

where $x_t = y_t - y_t^n$ denotes the output gap, y_t^n is the natural level of output; and $t.i.p.$ denotes the terms independent of policy. Then, as usual, we can express our welfare function as follows:

$$W = -\frac{1}{2}E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\varepsilon_p}{\Theta} var_i \{p_t(i)\} + \left(\sigma + \frac{\gamma + \delta}{1 - \delta}\right) x_t^2 \right] \tag{26}$$

In order to specify our welfare criterion, we need to find an expression that relates $var_i \{p_t(i)\}$ to π_t. In our framework, as shown by Sheedy (2007), the aggregate price level evolves as:

$$\log p_t(i) = \sum_{h=0}^{\infty} \theta_h \log P_{t-h}^* \tag{27}$$

where P_t^* stands for the reset price and θ_h denotes the share of firms using a price which last change was h periods ago. Thus, price level is an average of past reset prices weighted by the share of firms using such price at time t.

By making use of (27) and exploiting the properties of the variance, we can show that the discounted sum of price dispersion evolves in the following way:

$$\sum_{t=0}^{\infty} \beta^t var_i \{p_t(i)\} = \frac{\sum_{t=0}^{\infty} \beta^t \left(\frac{1-\alpha_p}{1+\varphi_p} \pi_t^2 + \alpha_p \varphi_p \pi_{t-1}^2 \right)}{\left[1 - \beta \frac{(1-\alpha_p)}{(1+\varphi_p)}\right] (1 + \varphi_p) (\alpha_p + \varphi_p)}, \tag{28}$$

the proof is showed in Appendix.

Once we have derived the price dispersion, we substitute (28) into (26) and obtain our welfare measure:

$$W = -\Omega \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + \Gamma_1 x_t^2 + \Gamma_2 \pi_{t-1}^2 \right] \tag{29}$$

where Ω, Γ_1, and Γ_2 are expressed as follows:

$$\Omega = \frac{\varepsilon_p}{\Theta} \frac{1}{1 - \beta \left(\frac{1-\alpha_p}{1+\varphi_p} \right)} \frac{1 - \alpha_p}{(1 + \varphi_p)^2 (\alpha_p + \varphi_p)}, \tag{30}$$

$$\Gamma_1 = \frac{1}{\Omega} \left(\sigma + \frac{\gamma + \delta}{1 - \delta} \right), \tag{31}$$

$$\Gamma_2 = \frac{1 + \varphi_p}{1 - \alpha_p} \alpha_p \varphi_p. \tag{32}$$

As for the Phillips curve, our welfare measure (29) encompasses that deriving from Calvo price setting for $\varphi_p = 0$: in such a case the weight attached to backward inflation drops to zero.

5.2 Welfare analysis and optimal policy

5.2.1 Price-setting models and distortions

In the standard Calvo pricing model, the source of distortions is the (average) probability of resetting prices, which is constant among firms in each instant of time. Distortions and price duration are in fact mapped one-by-one to this probability. In other words, firms are ex ante homogenous, facing the same probability of being extracted in the Calvo's lottery. In our generalization of Calvo (1983), instead, we can disentangle two sources of distortions stemming from the average and the distribution of the probability of resetting prices. Here reset probability is not constant, unless the hazard is flat. Firms are ex ante heterogeneous, as each face a different probability of resetting according to the time elapsed since the last reset.[25] Distortions are related to the

[25] Of course, two firms who have reset at the same time have the same probability to re-adjust.

average probability of resetting prices (as in Calvo standard model), but now they also depend on how the reset probability is distributed among firms.

A given average probability of resetting prices (ARP) can be obtained for different slopes of the hazard function (φ_p).[26] For instance, we can obtain three different scenarios consistent with $ARP = 0.3$ associated to three different distributions of the reset probability, as illustrated by the figure below.

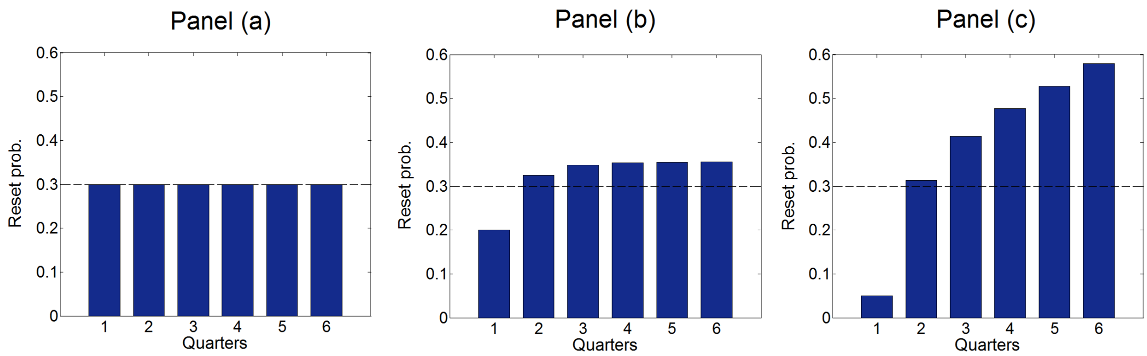

Figure 3 – Different distributions of the reset probability with the same mean.

All the panels of Figure 3 imply $ARP = 0.3$, but Panel (a) is built by assuming $\varphi_p = 0$ (Calvo price model); in Panel (b) $\varphi_p = 0.1$; in Panel (c) $\varphi_p = 0.25$. It is clear that in Panel (c) there is a greater dispersion in the probability to adjust a price, i.e. there is more heterogeneity in the reset probability distribution. Similar figures can be drawn for different average probabilities. Given the average reset probability, we can refer to the scenario (a) as Calvo; scenario (b) as the mid-dispersion case; (c) as the high-dispersion case. It is worth noticing that given the ARP level, Panels (a) shows that the probability of being able to post a new price is constant and equal among firms in the Calvo scenario. By contrast, if dispersion in the reset probability is not zero, given the ARP level, more dispersion implies lower individual probability for firms that have recently changed their prices and higher probability for those that did it in the far past to compensate and keep the APR constant.

Table 4[27] reports the welfare effects associated to several average probabilities with different dispersions. For the sake of comparison, we consider the same monetary policy rule in all cases.[28] We express the welfare effects in the more common form of welfare loss, which are further normalized to the Calvo scenario with a reset probability equal to 0.4 (i.e., our baseline). The table is built on a standard calibration of model parameters. We set the discount factor (β) equal to 0.99. The labor share is 2/3 so we set $\delta = 0.33$. We assume a net price mark-up of 20% setting $\varepsilon_p = 6$; we assume a log-linear utility function calibrating $\sigma = 1$ and $\gamma = 2$. The process governing the cost-push shock is modeled as an $AR(1)$ with an auto-regressive coefficient equal to 0.5. Our qualitative results are independent of the calibration.

Table 4 – Welfare cost of different price setting schemes

Average reset probability	Hazard function slope (φ_p)		
(ARP)	0	0.1	0.25
0.4	1.000	1.417	2.291
0.3	1.769	2.631	4.641
0.25	2.259	3.438	6.418

[26] By using the following relationship: $ARP = \alpha_p + \varphi_p$.

[27] The table is built by considering a transitory cost push shock. Thus it reports conditional welfare.

[28] We consider a policy consistent to a Taylor rule, which only responds to current inflation by a coefficient equal to 1.5. The same results hold if optimal policies are instead considered. More details are available upon request.

The table shows that a decrease in the ARP induces more distortions as in the traditional setup. Moreover, here distortions also increase in the dispersion of the reset probability. The intuition of our results is related to the idea that distortions are related to the behavior of the firms allowed to reset the prices, i.e. those extracted in the lottery. Clearly, firms that do not adjust cannot be associated to distortional behaviors. The intuition follows.

In general, independently of the hazard shape, the optimal pricing rule for the firm is to apply a mark-up over its marginal cost. However, since firms do not know when they will be able to re-optimize, they must balance the one-time cost of changing prices against the benefit of staying close to the profit maximizing mark-up over time. This balance is costly and induces distortions. The cost is increasing in the degree of price stickiness faced by firms able to re-optimize. What matters is the perceived probability of being able to update the price in the future (or the expected duration of the contract) when price has to be reset.

In the Calvo setup the probability of being able to post a new price is constant and equal among firms, independently of the fact that they are re-optimizing or not (see e.g. Figure 3: Panel (a)). A smaller average probability implies a lower individual reset probability for all the adjusting firms and thus, as said, it leads to more distortions. The same occurs if the hazard has positive slope. In this case every firm has a different probability to be selected for the lottery, depending on the time spent since last price reset (see e.g. Figure 3: Panels (b) and (c)); but a smaller ARP, in turn, will lower each individual probability—including that of the firms extracted in the lottery.

However, distortions also increase in reset probability dispersion (i.e. in the hazard slope). For any given average, in fact, more dispersion means that after the lottery the individual reset probability of selected firms dramatically falls, while that of the other increases. As a consequence, the steeper is hazard, the stronger this effect is. In other words, the expected duration of a price spell is increasing in the dispersion. As distortions depend on expected duration of adjusting firms' contracts, more dispersion entails more distortions.

The next subsection investigates how these two features of the price setting process affect the conduct of monetary policy under different assumptions about the policy regime (discretion and commitment), and how they affects the relative gains of the latter on the former.

5.2.2 Optimal policies

The Government problem consists in maximizing (29) subject to (20), conditional to the policy regime. Formally, the Government's problem at some point of time, here taken (without loss of generality) to be $t = 0$, can be expressed as minimization[29] of the following Lagrangian expression:

$$\begin{aligned}\min_{\{\pi_{t+j}\}_{j=0}^{+\infty},\,\{y_{t+j}\}_{j=0}^{+\infty}} & E_0 \sum_{j=0}^{\infty} \beta^j \Big\{ \frac{1}{2}\left[\pi_{t+j}^2 + \Gamma_1 x_{t+j}^2 + \Gamma_2 \pi_{t+j-1}^2\right] + \\ & + \lambda_{t+j}\left[\pi_{t+j} - \Psi_1 \pi_{t+j-1} - \beta\Psi_2\pi_{t+j+1} + \beta^2\Psi_1\pi_{t+j+2} - \kappa x_{t+j} - u_{t+j}\right]\Big\}\end{aligned} \tag{33}$$

where λ_{t+j} is the Lagrangian multiplier. Corresponding first order conditions are:

$$(1+\Gamma_2\beta)\,\pi_{t+j} + \lambda_{t+j} - \Psi_1\beta E_{t+j}\lambda_{t+j+1} = 0 \quad \text{for } j = 0 \tag{34}$$

$$(1+\Gamma_2\beta)\,\pi_{t+j} + \lambda_{t+j} - \Psi_1\beta E_{t+j}\lambda_{t+j+1} - \Psi_2\lambda_{t+j-1} + \Psi_1\lambda_{t+j-2} = 0 \quad \text{for } j > 0 \tag{35}$$

$$\Gamma_1 x_{t+j} - \kappa\lambda_{t+j} = 0 \quad \text{for } j \geqslant 0 \tag{36}$$

As it is well known, the system (34)-(36) leads to a dynamic inconsistency. At $t = 0$ the policymaker implements (34) and commits to act following (35) in future periods. When $t = 1$ comes, it would be optimal to pursue again (34) rather than (35). Thus, time-inconsistency arises as it is no longer optimal act as planned at $t = 0$.

[29] Without loss of generality, we are considering a welfare loss minimization instead of a welfare maximization.

Dynamic inconsistency does not affect the decisions under discretion as the discretionary regime is a process that presumes period-by-period re-optimization given the expectations. As finding a solution under discretionary regime is not trivial when intrinsic inflation persistence is considered, we adopt Söderlind (1999) method to find the equilibrium path followed by the endogenous variables.

Woodford (2003) proposes to overcome dynamic inconsistency by implementing an alternative regime known as "timeless perspective," where the Government should ignore (34). The idea is that optimal policy solves problem (33) "from some date forward as being optimal from a timeless perspective, rather than from the perspective of the particular time at which the policy is actually chosen" (Woodford, 2011; p. 744). In such a case, the solution is given by equations (35) and (36). At a generic time t, combining these equations, we eliminate λ_t and obtain:

$$\pi_t = -\frac{\Gamma_1}{\kappa(1+\Gamma_2\beta)}\left(x_t - \Psi_1\beta E_t x_{t+1} - \Psi_2 x_{t-1} + \Psi_1 x_{t-2}\right) \tag{37}$$

As the Phillips curve has an inertial component, discounted expected output gap should be considered. Therefore, the timeless perspective is characterized by the fact that the Government by reacting to the past output gap affect current inflation expectations and obtain a more favorable current trade-off between inflation and output. Thus, as the curve has two forward components, in order to manipulate future expectations the Government rule now reacts to two lags of output gap; the different signs are explained by the different effects on the expectations (see (20)). If $\varphi_p = 0$, inflation expectations at $t+2$ do not affect the trade-off and thus $\Psi_1 x_{t-2}$ drops to zero.

The impulse response functions to a cost-push shock in the two regimes are depicted in the Figure 4, where we consider three scenarios illustrated in Figure 3 (i.e., Calvo, mid-dispersion and high-dispersion under a common average reset probability equal to 0.3).

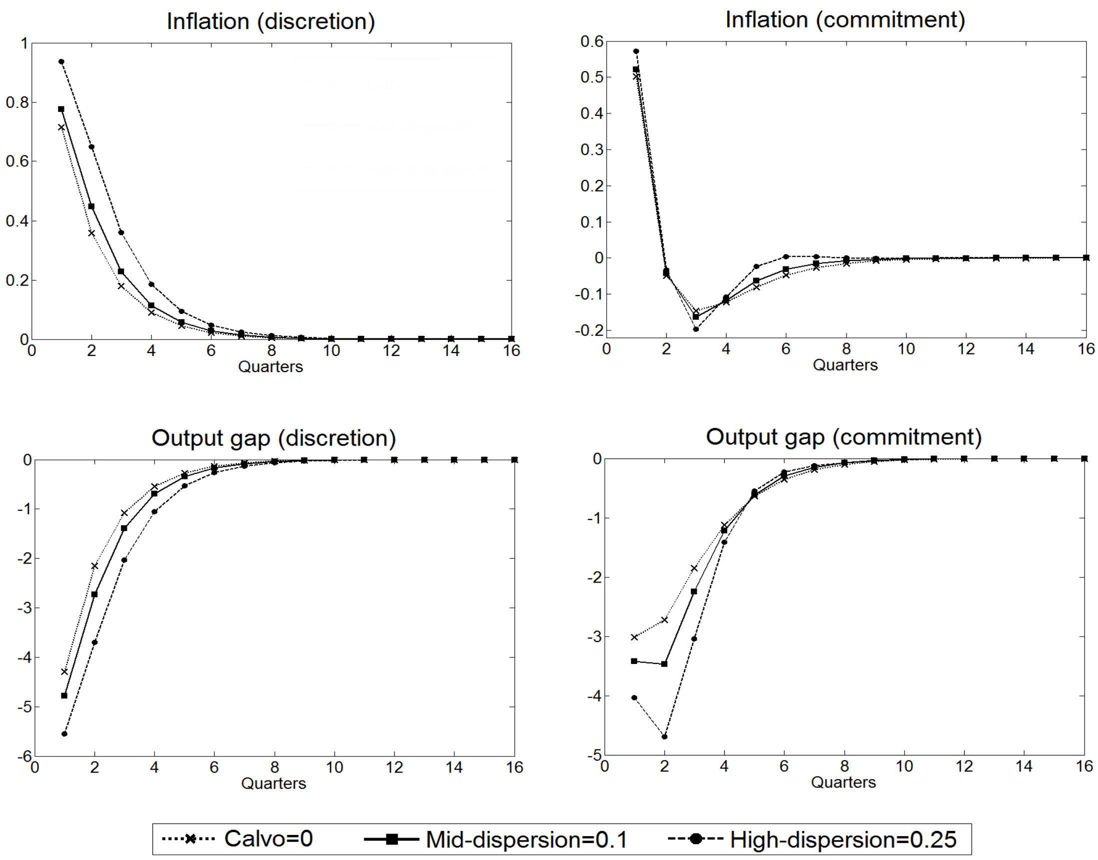

Figure 4 – Impulse response functions in the two policy regimes

Considering discretion, the comparison among the three scenarios shows that inflation and output gap follow the same shapes. They only differ on a quantitative point of view: A greater dispersion in the reset probability leads to a larger variability in both inflation and output gap. In the Calvo case we can observe lower variances, here the discretionary policy is more aggressive since it responds only to current output gap. However, our results do not depend on monetary policy efficacy, but on the different degree of distortions associated to the different scenarios—as we have shown before. The same results are observed under commitment.

Following Steinsson (2003), it is interesting to compare the relative performance of the two regimes. In our setup an increase in the dispersion of the reset price distribution leads to more inflation persistence. The following table shows the relative gains on discretion derived by implementing a timeless perspective regime. We consider different slopes for the hazard function starting from Calvo ($\varphi_p = 0$) to the high-dispersion scenario ($\varphi_p = 0.3$). In all cases we keep fixed the average probability of changing prices (APR) at 0.3. It is worth noticing that the inflation intrinsic persistence grows with φ_p. Thus, the table reports the gains associated to the commitment under different degree of inflation inertia.

Table 5 – Gains from commitment and inflation inertia

hazard function slope (φ_p)	0	0.10	0.15	0.20	0.25	0.30
% gain from commitment	54.3	58.6	62.2	66.9	72.9	78.8

The greater φ_p is, the higher the relative gain of commitment is. This result differs from Steinsson (2003), who finds that gain increases in the degree of inflation persistence. In Steinsson (2003) inflation inertia is introduced by a rule-of thumb, and results come out from the fact that as the backward component of the Phillips curve increases, the forward-looking one falls reducing the gain from commitment. In our setup the interaction is much complex as intrinsic inflation persistence is micro-founded and the effects on the different components of the Phillips curve are non-linear.

In a model characterized by high-dispersion, e.g. $\varphi_p = 0.25$, the optimizing firm knows that, in the period immediately after the price reset, it will be very unlikely to change its price again (Figure 3, Panel (c)): thus, in choosing a new price, it gives much importance to the future state of the economy, when it will probably be allowed to reset again. As a consequence, commitment gains are higher in a world with high-dispersion as expectations on the future state of the economy matter more and thus implementing a policy rule able to influence private sector expectations lead to an improvement of the inflation-output gap trade-off.

The relative effects of inflation inertia on the optimal policy regimes are thus not general but they depend on the way intrinsic inflation persistence is introduced.

5.2.3 Stabilization policies and misinterpretation of persistence sources

This section assumes model uncertainty and considers the cost of misunderstanding the true sources of inflation persistence by comparing the dynamics of implementing a stabilization policy when the policymaker overestimates the degree of intrinsic persistence to those arising from ignoring actual structural persistence. We consider both discretion and commitment. At the end of the section, by using robust control techniques (specifically, the *minimax* regret criterion), we look at robust policies.

In the spirit of Sbordone (2007), we consider two specifications, in both we set $APR = 0.3$. The first specification is a purely forward-looking New Keynesian Phillips curve ($\varphi_p = 0$), in which persistence is extrinsic, i.e. it depends on the auto-correlation of the cost push shock. The second is the high-dispersion scenario ($\varphi_p = 0.25$) that endogenously generates (intrinsic) inflation persistence by way of a lagged inflation term in the New Keynesian Phillips curve. In each specification we compute the two welfare losses associated to a conduct for monetary policy where the government attributes the observed persistence to an extrinsic or intrinsic nature, implementing the discretionary rule by assuming $\varphi_p = 0$ or $\varphi_p = 0.25$, respectively. Of course, for each specification one policy rule is "correct," while the other is based on a misinterpretation of the source of inflation persistence. The same quantitative exercise is performed for the case of commitment—by using the rule given by (37).

Our results are presented in Table 6 that reports losses in the two specifications under different policy rule. We compute the percentage loss deriving from the implementation of an optimal stabilization policy under wrong policymaker beliefs about the nature of inertia. To grasp the intuition of our result we also report the impulse response functions in Figure 5 and 6, which refer to discretion and commitment, respectively. In figures we depict the equilibrium paths of inflation and output gap to a cost-push shock when the policymaker conducts a stabilization policy that is optimal conditional to his beliefs.

Table 6 illustrates our results. It reports the welfare associated to different policy rule based on $\varphi_p = 0$ or $\varphi_p = 0.25$ under the discretion (Panel (a)) and timeless-commitment regime (Panel (b)).

Table 6 – Welfare loss for different policy rules

	(a) Discretion		(b) Commitment	
	True source of inflation persistence			
Policy rule assuming:	$\varphi_p = 0$	$\varphi_p = 0.25$	$\varphi_p = 0$	$\varphi_p = 0.25$
$\varphi_p = 0$	0.343	0.427	0.247	0.311
$\varphi_p = 0.25$	0.351	0.439	0.252	0.305
Loss difference (%)[30]	2.33	−2.73	2.02	1.96

In the case of discretion, the cost of overestimating inflation inertia in setting monetary policy when the economy is governed by a purely forward-looking Phillips curve is 2.33%. By contrast, the cost of ignoring inflation persistence when the true model is characterized by intrinsic inertia ($\varphi_p = 0.25$) is negative, i.e. -2.73%. For the discretionary regime, we therefore find that ignoring persistence always involves welfare gains. This result suggests that the costs of implementing a stabilization policy when the government overestimates the degree of intrinsic persistence are potentially higher than the costs of ignoring actual structural persistence.

The rationale of the above result can be found by inspecting the impulse response functions. The case of discretion is illustrated in Figure 5. The left panel plots the impulse response functions of inflation and output gap to a cost-push shock when the policymaker implements a policy rule assuming $\varphi_p = 0$; in contrast, right panel depicts the equilibrium path if the stabilization policy is conducted by setting $\varphi_p = 0.25$.

After the cost-push shock hits the economy, inflation immediately jumps up and then progressively returns to its steady state after about five quarters; the output gap declines at the impact and then gradually converges to the steady state. Comparing the impulse response functions of the left panel with those of the right panel we can see that ignoring intrinsic persistence always leads to more stabilization of inflation, lowering the welfare loss.

In other words, the implementation of a wrong rule in the latter scenario somehow improves the inflation-output gap trade-off. Clarida *et al.* (1999) show that as long as the extent that price setting today depends on beliefs about future economic conditions, a government may face an improved short-run output-inflation trade-off by engineering a greater contraction in output in response to inflationary pressures. Thus, it would optimal for the society to delegate monetary policy to a conservative central banker—even when monetary authorities operate under discretion. In our context this commitment is unfeasible; however, model uncertainty leads the central bank to credibly use a more aggressive (wrong) policy improving the short-run trade-off since the policy rule associated to $\varphi_p = 0$ is always more aggressive than that when $\varphi_p > 0$.

[30] The difference is obtained by considering the welfare loss of government who misinterprets the source of persistence and the loss of a government who does not.

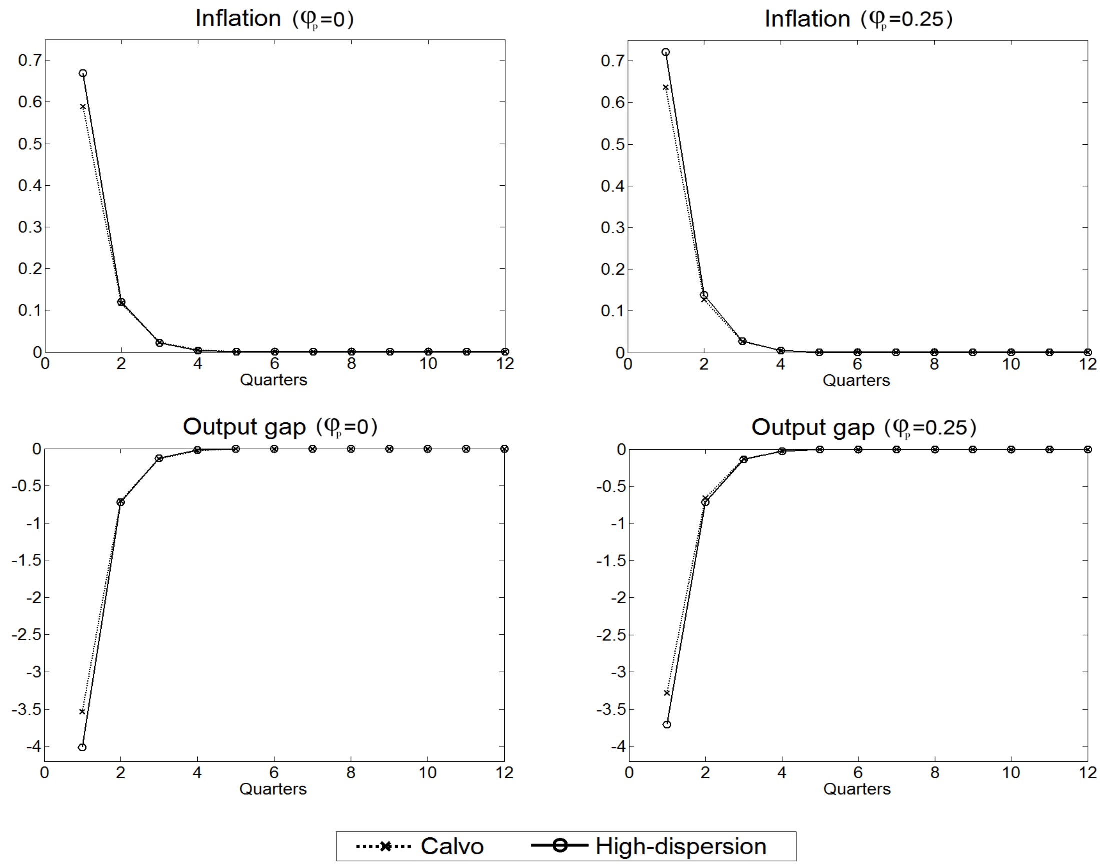

Figure 5 - Optimal stabilization policy (discretionary regime).

Regarding commitment, Table 6 results imply that misunderstanding the true source of persistence implies an additional loss in both specifications. In per cent terms, these losses are 2.02% (when the true model is a purely forward-looking Phillips, i.e. extrinsic inertia) and 1.97% (when inflation persistence is intrinsic). Now misunderstanding are always costly. In this regime there are not problem of credibility, then implementing a different rule is unlikely to lead to a welfare improvement.[31]

Figure 6 provides the impulse response functions in the commitment regime. As expected, inflation shows an initial increase, offset by a mild deflation in the following quarters; the output gap follows a path similar to that under discretion. The shape of the responses are quite similar under both the true models considered, but they differ for the initial impact of the shock. Implementing an optimal policy rule that accounts for intrinsic persistence involves more stabilization for inflation; on the other side, under this policy, output gap is more volatile. The net effect is that conducting a stabilization policy under wrong assumptions about the true source of inertia always entail welfare worsening. Thus, differently from the discretionary case, for the policymaker is optimal to correctly choose the “right” rule in order to minimize the welfare loss.

[31] Although not impossible (for a discussion, see Blake, 2001; or Jensen and McCallum, 2002).

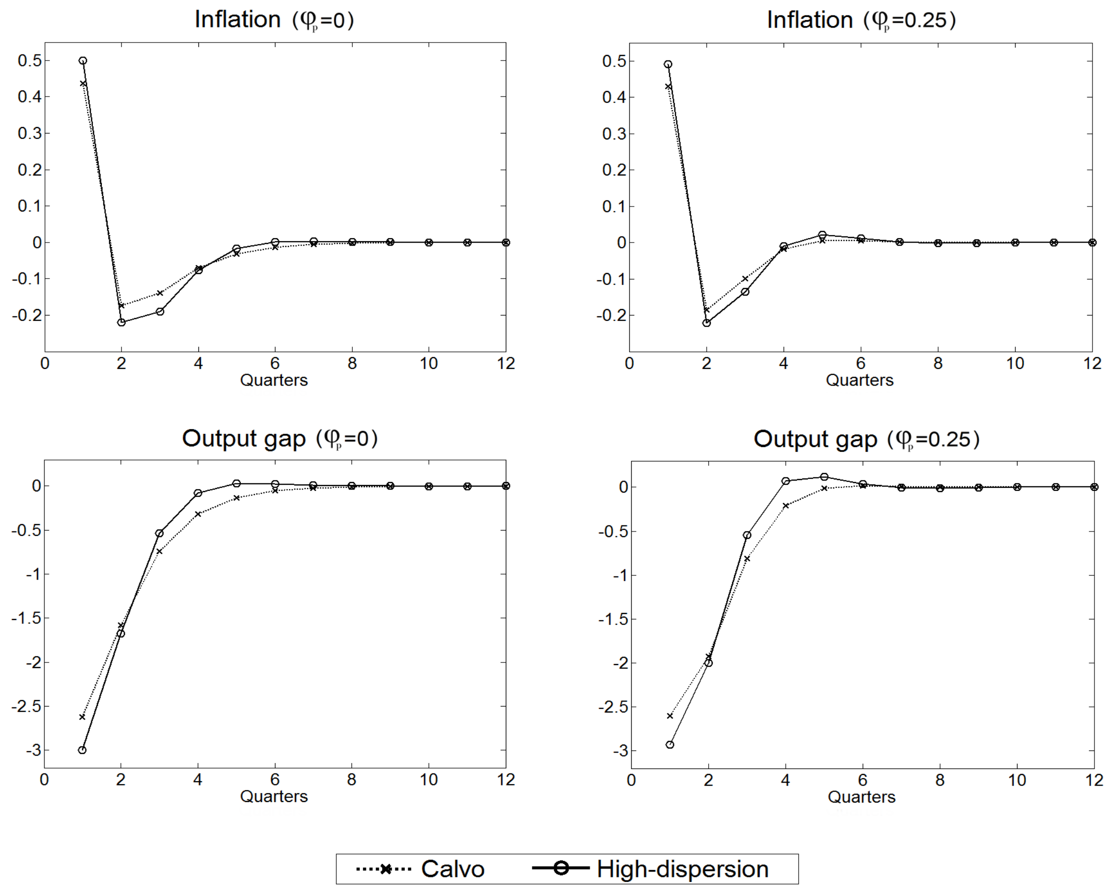

Figure 6 - Optimal stabilization policy (commitment case).

The results above (Table 6, Figure 5 and 6) show that misunderstanding the true sources of inflation persistence have several effects. In particular, it seems to be reasonable to ignore the intrinsic persistence in the case of discretion under uncertainty about the true model of the economy. In contrast, the case of commitment is less clear as here in both specifications misunderstandings are costly. In such a contexts, as emphasized by Brock *et al.* (2007) and Sbordone (2007), a good robustness criterion to select among different policy rule is the *minimax* regret criterion (Savage, 1951).[32]

The Savage's criterion can be explained as follows. Consider two possible scenarios A and B, to each of them there is a corresponding model of the economy between those of the set $M = \{m_A, m_B\}$. The policymaker can then perform two kinds of policies according to its beliefs about the true model of the economy, these polices are summarized in the set $P = \{p_A, p_B\}$. The regret is then defined as loss incurred by a certain policy given a certain model relative to what would have incurred under the policy leading to the lower loss for that model. Formally, for $p_i \in P$ and $m_j \in M$, the regret is $R(p_i, m_j) = L(p_i, m_j) - \min_{p \in P} L(p, m_j)$, where $L(\ )$ is the expected loss. The robust policy p^* is then obtained by the *minimax* regret criterion:

[32] See Stoye (2011) for its axiomatization.

$p^* = \min_{p \in P} \max_{m \in M} R(p, m)$, i.e. is the *minimax* applied to the relative loss associated with choosing a policy in absence of knowledge of the correct model.

In our case the two possible scenarios are related to the origins of inflation persistence. The regret for both discretion and commitment are reported in the following table. The robust policy is then chosen by applying the *minimax* to panel (a) for discretion and (b) for commitment.

Table 7 – Regret for different policy rules and models

	(a) Discretion		(b) Commitment	
	True source of inflation persistence			
Policy rule assuming:	$\varphi_p = 0$	$\varphi_p = 0.25$	$\varphi_p = 0$	$\varphi_p = 0.25$
$\varphi_p = 0$	0	0	0	0.006
$\varphi_p = 0.25$	0.008	0.012	0.005	0

In discretionary regime is optimal to act considering $\varphi_p = 0$, as ignoring intrinsic inertia minimize the regret. Under commitment, making a mistake about the origin of inflation inertia always involve incremental costs: A regret of 0.005 arises when persistence is enterely generated by a serially correlated shock, whereas the regret is 0.006 whether the persistence is due to a backward component in the Phillips curve. Notwithstanding the incremental costs are very similar, the regret is minimized for a policy rule following $\varphi_p = 0.25$, thus the policymaker should conduct monetary policy by following this rule.

6 Conclusions

We have built and estimated a model that considers both price and wage adjustment governed by the time-dependent mechanism described by Sheedy (2007). By making use of a hazard function, we have derived price and wage Phillips curves that are able in micro-founding price and wage inflation intrinsic persistence–as they are characterized by both forward and backward terms for inflation. We have estimated our model with Bayesian techniques. The estimation of our model has confirmed that a hazard function upward-sloping emerges. Differently from Benati (2009), who only considers price inflation, we find that the hazard function slope does not change with the policy regime, i.e. during the Great Moderation era. Finally, we have compared the empirical performance in fitting the data of our model to those of others based on popular alternative mechanisms for price and wage adjustment. By comparing log-marginal likelihoods of different estimations, we have found that our model clearly outperforms alternatives. It would be also interesting to look at other implications of the different modelling choices for price setting; e.g., implications for welfare and optimal monetary policy. However, this is beyond the scope of the current paper, and we let it for future researches.

By our welfare approximation, we explored the nature of macro distortions induced by our pricing model and its implications for the optimal monetary policy under discretion and (timeless perspective) commitment.

We disentangled two sources of distortions by considering both the average and the distribution of the probability of resetting prices, showing how welfare falls in the former, as in Calvo price setup, but also in the latter. The greater distortions also imply more variability when optimal policy is introduced, but do not affect qualitatively monetary responses under commitment or discretion. Regarding the central banker's conduct, in both regimes, monetary authority should take account of expected future output gaps because of its persistence. Furthermore, in the commitment regime, monetary policy should respond to an additional lagged term of the output gap to optimally affect expectations and thus to improve the current trade-off between inflation and output—since now expectations affect the current state of the economy for two periods.

By comparing the welfare under different policy regimes, we found that the relative effects of inflation inertia on the optimal policy regimes are different from those stemming from alternative models of inflation persistence based on rule-of-thumb assumptions or indexation mechanisms.

In our setup the relative gain of commitment over discretion increases in the degree of inflation persistence as—once intrinsic inflation persistence is micro-founded—the interaction between the different components of the Phillips curve is more complex. Therefore, relative effects of inflation inertia on the optimal policy regimes are at least not general, but they depend on the way intrinsic inflation persistence is introduced.

Assuming model uncertainty, we also analyzed the effects on optimal monetary policy of misinterpreting the sources of inflation inertia. Under a discretionary regime, we show how considering a policy behavior wrongly based on the idea that the true source of persistence is extrinsic leads to lower welfare costs compared to the case where the policymaker correctly understand the right source of inflation inertia. The rationale of this result is due to the fact that monetary authority wrongly react in a too aggressive way to inflation, but then it incidentally improves the inflation-output gap trade-off by stabilizing the inflation expectations—even in absence of credibility. This result, however, is conditional to the policy regime assumed, as under commitment implementing the wrong rule is always costly whatever the right model is. Finally, using robust control techniques, we found that robust policy implies to ignore intrinsic persistence under model uncertainty, when a discretionary regime is considered. Under commitment, despite having wrong beliefs about the correct source of inflation inertia always lead to similar increment of the welfare loss, the regret is minimized implementing an optimal policy rule based on $\varphi_p = 0.25$.

Appendix A – Wage Phillips curve derivation

Following Sheedy (2007), we assume that wages are set according to a time-dependent mechanism: the probability to change a wage depends positively on the time elapsed since last wage reset. This adjustment process can be formalized by using a hazard function.[33]

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

$$D_t(j) \equiv \min \{l \geq 0 \mid j \in \Gamma_{t-l}\} \tag{38}$$

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

We now introduce the hazard function, which expresses the relationship between the probability to post a new wage and the wage duration. The hazard function is defined by a sequence of probabilities: $\{\alpha_l\}_{l=1}^{\infty}$, where α_l represents the probability to reset a wage which remained unchanged for l periods. This probability is defined as: $\alpha_l \equiv \Pr(\Gamma_t \mid D_{t-1} = l-1)$.

Each hazard function is related to a survival one, 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: $\{\varsigma_l\}_{l=0}^{\infty}$, where ς_l denotes the probability that a wage fixed at time t will be still in use at time $t+l$. Formally, the survival function is defined by:

$$\varsigma_l = \prod_{h=1}^{l} (1 - \alpha_h) \tag{39}$$

with $\varsigma_0 = 1$. Following Sheedy (2007), we assume that the hazard function satisfies two restrictions:

$$\begin{cases} \alpha_1 < 1, \text{ meaning that is allowed a degree of wage stickiness;} \\ \alpha_\infty > 0, \text{ with } \alpha_\infty = \lim_{l\to\infty} \alpha_l. \end{cases} \tag{40}$$

The hazard function can be reparameterized by making use of a set of $n+1$ parameters and rewritten as (1), where $\{\varphi_l\}_{l=1}^{n}$ is a set of n parameters that control the hazard slope, whereas parameter α controls its initial level.

[33] We refer to Sheedy (2007) for the proofs relative to the hazard function mentioned here. See in particular his Appendix A.2 and A.5.

By making use of (39), we can rewrite the non-linear recursion (1) for the wage adjustment probabilities as a linear recursion for the corresponding survival function:

$$\varsigma_l = (1-\alpha)\varsigma_{l-1} - \sum_{h=1}^{\min(l-1,n)} \varphi_h \varsigma_{l-1-h} \tag{41}$$

The parameters $\{\varphi_l\}_{l=1}^n$ control the slope of the hazard function in the following way:

$$\begin{cases} \varphi_l = 0, \text{ for all } l = 1, ..., n \text{ the hazard is flat (Calvo case);} \\ \varphi_l \geq 0, \text{ for all } l = 1, ..., n \text{ the hazard is upward-sloping;} \\ \varphi_l \leq 0, \text{ for all } l = 1, ..., n \text{ the hazard is downward-sloping.} \end{cases} \tag{42}$$

Let $\theta_{lt} \equiv \Pr(D_t = l)$ denote the proportion of households earning at time t a wage posted at period $t-l$. The sequence $\{\theta_{lt}\}_{l=0}^{\infty}$ denotes the distribution of the duration of wage stickiness at time t. This distribution evolves over the time according to:

$$\begin{cases} \theta_{0t} = \sum\limits_{l=1}^{\infty} \alpha_l \theta_{l-1,t-1} \\ \theta_{lt} = (1-\alpha_l)\,\theta_{l-1,t-1} \end{cases} \tag{43}$$

If the hazard function satisfies the restrictions (40) and the evolution over the time of the distribution of wage duration evolves as in (43), then *a)* from whatever starting point, the economy always converges to a unique stationary distribution $\{\theta_l\}_{l=0}^{\infty}$. Hence $\theta_{lt} = \theta_l = \Pr(D_t = l)$, $\forall t$; *b)* let consider (1) and assume that the economy has converged to $\{\theta_l\}_{l=0}^{\infty}$, the following three relations are obtained:

$$\begin{cases} \theta_l = \left(\alpha + \sum\limits_{h=1}^{n} \varphi_h\right) \varsigma_l \\ \alpha^e = \alpha + \sum\limits_{l=1}^{n} \varphi_l \\ D^e = \frac{1-\sum_{l=1}^{n} l\varphi_l}{\alpha+\sum_{l=1}^{n}\varphi_l} \end{cases} \tag{44}$$

where α^e denotes the unconditional probability of wage reset and D^e represents the expected duration of wage stickiness.

Our supply side of the economy is fairly standard (see, e.g., Galí, 2008: Chapter 6). It is composed by 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: $Y_t(i) = A_t N_t(i)^{1-\phi}$, 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 $1-\phi$ is the labor share. The quantity of labor used by firm i is defined by:

$$N_t(i) = \left[\int_{\Omega} N_t(i,j)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} dj\right]^{\frac{\varepsilon_w}{\varepsilon_w - 1}} \tag{45}$$

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 employed yields to labor demand schedule:

$$N_t(i,j) = \left(\frac{W_t(j)}{W_t}\right)^{-\varepsilon_w} N_t(i) \tag{46}$$

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

$$W_t = \left[\int_0^1 W_t(j)^{1-\varepsilon_w} dj\right]^{\frac{1}{1-\varepsilon_w}} \tag{47}$$

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 $N_t(j) = \int\limits_{\Omega} N_t(i,j)di$ 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:

$$U(C_t(j), N_t(j)) = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[g_t \frac{(C_t(j) - hC_{t-1}(j))^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\gamma}(j)}{1+\gamma} \right] \right\} \tag{48}$$

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

$$P_t(j) C_t(j) + E_t [Q_{t+1,t} B_t(j)] \leq B_{t-1}(j) + W_t(j) N_t(j) + T_t(j) \tag{49}$$

where $P_t(j)$ is the price of good j, $B_t(j)$ denotes holdings of one-period bonds, 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 it is described by a CES aggregator: $C_t(j) = \left(\int\limits_{\Theta} C_t(i,j)^{\frac{\varepsilon_p - 1}{\varepsilon_p}} di \right)^{\frac{\varepsilon_p}{\varepsilon_p - 1}}$, 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 (48) internalizing the effects of labor demand (46) and taking account of (49). Households are subject to a random probability to reset price, but, according to our time-dependent mechanism, a wage change will be more likely to be observed when last price reset happened many periods ago. Formally, suppose that at time t a household sets a new wage, denoted by W_t^*,[34] if the household still earns this wage at time $\tau \geq t$ then its relative wage will be W_t^*/W_τ and the household utility can be written as $U\left[W_t^*/W_\tau; C_{\tau|t}; N_{\tau|t}\right]$;[35] by considering the survival function, the household will then choose its optimal reset wage by solving:

$$\max_{W_t^*} \sum_{\tau=t}^{\infty} \varsigma_{\tau-t} E_t \left\{ \left(\prod_{s=t+1}^{\tau} \frac{\pi_s^p}{I_s} \right) U \left[\frac{W_t^*}{W_\tau}; C_{\tau|t}; N_{\tau|t} \right] \right\} \tag{50}$$

where $\pi_s^p = P_s/P_{s-1}$ is the gross price inflation rate and $I_s = i_s/i_{s-1}$ is the gross nominal interest rate. This maximization is subject to the budget constraint (49) and the labor demand schedule (46). Equation (50) yields the following first-order condition:

$$\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} E_t \left(\frac{W_t^*}{W_\tau} \right)^{-\varepsilon_w} \left(\prod_{s=t+1}^{\tau} \frac{\pi_s}{I_s} \right) \left[U_c(C_{\tau|t}, N_{\tau|t}) \frac{N_{\tau|t}}{P_\tau} (1-\varepsilon_w) - \varepsilon_w U_n(C_{\tau|t}, N_{\tau|t}) \frac{N_{\tau|t}}{W_t^*} \right] = 0 \tag{51}$$

where $U_c(C_{\tau|t}, N_{\tau|t})$ is the marginal utility of consumption and $-U_n(C_{\tau|t}, N_{\tau|t})$ is the marginal disutility of labor. Considering that the marginal rate of substitution between consumption and leisure is $MRS_{\tau|t} = -\frac{U_n(C_{\tau|t}, N_{\tau|t})}{U_c(C_{\tau|t}, N_{\tau|t})}$, and the steady-state wage mark-up is $\mu_w = \frac{\varepsilon_w}{\varepsilon_w - 1}$, equation (51) can be rearranged and expressed in terms of the optimal wage reset as:

$$W_t^* = \left[\frac{\mu_w \left(\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \beta^{\tau-t} MRS_{\tau|t} P_\tau \right)}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \beta^{\tau-t}} \right] \tag{52}$$

[34] Since each household solves the same optimization problem henceforth index j are omitted.

[35] $C_{\tau|t}$ and $N_{\tau|t}$ denote respectively the level of consumption and the labour supply at time τ of a household which last wage reset was in period t.

Assuming that the economy has converged to $\{\theta_l\}_{l=0}^{\infty}$, then the wage level (47) can be expressed as a weighted-average of the past reset wages:

$$W_t = \left(\sum_{l=0}^{\infty} \theta_l W_{t-l}^{*\,1-\varepsilon_w} \right)^{\frac{1}{1-\varepsilon_w}} \tag{53}$$

By log-linearizing (52) and (53) around a steady-state (characterized by zero wage inflation), we get:[36]

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\beta^{\tau-t} \varsigma_{\tau-t}}{\sum_{j=0}^{\infty} \beta^j \varsigma_j} \right) \left[w_\tau - \Xi_w \mu_\tau^w \right] \tag{54}$$

$$w_t = \sum_{l=0}^{\infty} \theta_l w_{t-l}^* \tag{55}$$

Equations (54) and (55) describe the wage adjustment mechanism. The time-dependent wage Phillips curve (9) used in the paper is derived by combining them with (41) and (44).

Specifically, by combining (41) with (54), we obtain:

$$w_t^* = \beta(1-\alpha) E_t w_{t+1}^* - \sum_{l=1}^{n} \beta^{l+1} \varphi_l E_t w_{t+l+1}^* + \left[1 - \beta(1-\alpha) + \sum_{l=1}^{n} \beta^{l+1} \varphi_l \right] (w_t - \Xi_w \mu_t^w) \tag{56}$$

By making use of (44), equation (55) can be recast as follows:

$$w_t = (1-\alpha)\, w_{t-1} - \sum_{l=1}^{n} \varphi_l w_{t-1-l} + \left(\alpha + \sum_{h=1}^{n} \varphi_h \right) w_t^* \tag{57}$$

where we have used the fact that the stationary distribution of the wage duration (44) can be rewritten in recursive way as:

$$\theta_l = (1-\alpha)\theta_{l-1} - \sum_{h=1}^{\min(l-1,n)} \varphi_h \theta_{l-h-1} \tag{58}$$

with $\theta_0 = \alpha + \sum_{h=1}^{n} \varphi_h$.

The general expression for the wage Phillips curve is obtained from (56) and (57):

$$\pi_t^w = \sum_{l=1}^{n} \psi_l \pi_{t-l}^w + \sum_{l=1}^{n+1} \delta_l E_t \pi_{t+l}^w - k_w \mu_t^w \tag{59}$$

where the coefficients ψ_l, δ_l and k_w have the following parameterization:

$$\begin{aligned}
\psi_l &= \frac{\varphi_l + \sum_{h=l+1}^{n} \varphi_h \left[1 - \beta\,(1-\alpha) + \sum_{k=1}^{h-1} \beta^{k+1} \varphi_k \right]}{\chi} \\
\delta_1 &= \frac{\beta \left[(1-\alpha) - \sum_{h=1}^{n} \beta^h \varphi_h \left(\alpha + \sum_{k=1}^{h-1} \varphi_k \right) \right]}{\chi} \\
\delta_{l+1} &= -\frac{\beta^{l+1} \left[\varphi_l + \sum_{h=l+1}^{n} \beta^{h-1} \varphi_h \left(\alpha + \sum_{k=1}^{h-1} \varphi_k \right) \right]}{\chi} \\
k_w &= \frac{\Xi_w \left[\left(\alpha + \sum_{h=1}^{n} \varphi_h \right) \left[1 - \beta\,(1-\alpha) + \sum_{h=1}^{n} \beta^{h+1} \varphi_h \right] \right]}{\chi}
\end{aligned}$$

[36] Small-caps letters denote log-deviation from the steady-state.

where $\chi = (1-\alpha) - \sum_{h=1}^{n} \varphi_h \left[1 - \beta(1-\alpha) + \sum_{k=1}^{h-1} \beta^{k+1} \varphi_k\right]$, for $l = 1, ..., n$.

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 (59) becomes that reported in the paper, i.e. (9).

Appendix B – GMM estimation of the wage Phillips curve

As in Sheedy (2007) we estimate our wage Phillips curve via generalized method of moments (GMM), in order to get priors for the parameters affecting the hazard function. Since it is not easy to find an observable proxy for the wage mark-up, the latter can be expressed as a function of unemployment, as in Galí *et al.* (2011):

$$\mu_t^w = \gamma u_t \tag{60}$$

where u_t represents the unemployment gap. Therefore (59) becomes:

$$\pi_t^w = \sum_{l=1}^{n} \psi_l \pi_{t-l}^w + \sum_{l=1}^{n+1} \delta_l E_t \pi_{t+l}^w - k_w \gamma u_t \tag{61}$$

To perform a GMM estimation of (61) we need to use a set of instruments, in order to correctly identify all the coefficients. Let z_{t-1} represents a vector of observable variables known at time $t-1$: under rational expectations the error forecast of π_t^w is uncorrelated with information contained in z_{t-1}, then the following orthogonality condition holds:

$$E_t\left[\left(\pi_t^w - \sum_{l=1}^{n} \psi_l \pi_{t-l}^w - \sum_{l=1}^{n+1} \delta_l E_t \pi_{t+l}^w + k_w \gamma u_t\right) z_{t-1}\right] = 0 \tag{62}$$

Following Galí and Gertler (1999), since (62) is non-linear in the structural parameters, we normalize the orthogonality condition in the following way:

$$E_t\left[\left(\chi \pi_t^w - \chi \sum_{l=1}^{n} \psi_l \pi_{t-l}^w - \chi \sum_{l=1}^{n+1} \delta_l E_t \pi_{t+l}^w + \chi k_w \gamma u_t\right) z_{t-1}\right] = 0 \tag{63}$$

Our estimation is made using quarterly U.S. data ranging from 1960:1 to 2011:4: all the data comes from FRED database. The wage inflation is measured by the *compensation per hour*, whereas for the unemployment rate we use the *civilian unemployment rate.* The set of instruments is composed by the lags of the following observable variables: wage inflation, unemployment, price inflation, consumer price index, output gap, labor share, spread between ten-year Treasury Bond and three-month Treasury Bill yields. In particular six lags of price inflation, wage inflation and CPI, four lags for the output gap and two lags for the remaining instruments are used.[37] For the sake of simplicity we show only the estimation of (63) when $n = 1$.[38] Under the latter assumption, (62) and (63) change as follows:

$$E_t\left\{\left[\pi_t^w - \psi_w \pi_{t-1}^w - \beta(1+(1-\beta)\psi_w) E_t \pi_{t+1}^w + \beta^2 \psi_w E_t \pi_{t+2}^w + k_w \gamma u_t\right] z_{t-1}\right\} = 0 \tag{64}$$

$$E_t\left\{\frac{1}{\chi_w}\left[\pi_t^w - \psi_w \pi_{t-1}^w - \beta(1+(1-\beta)\psi_w) E_t \pi_{t+1}^w + \beta_w^2 \psi_w E_t \pi_{t+2}^w + k_w \gamma u_t\right] z_{t-1}\right\} = 0 \tag{65}$$

[37] The sample range, the data and the NKWPC specification used for GMM estimation differ from those of Bayesian estimation. We made this choice in order to avoid 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 Table 3).

[38] For $n > 1$ we test that the extra leads and lags deriving from this specification are not statistically significant.

where $\chi_w = (1-\alpha_w) - \varphi_w [1-\beta(1-\alpha_w)]$.

The structural form of (65) is estimated by imposing $\beta = 0.99$, $\varepsilon_w = 8.85$ and $\gamma = 2$; the reduced form coefficients (see (10)) are convolution of the structural parameters estimated and they are obtained by substituting these parameters into them; the standard errors are computed using the delta method.[39]

The results for the structural form estimation are reported in Table 8. We show the estimation for the structural parameters φ_w (hazard slope) and α_w (hazard initial value); moreover, we also report D^e and α_w^e (computed as in (44)) and the $J-stat$.

Table 8 – Wage Phillips curve estimation (structural form)[40]

α_w	φ_w	D^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 coefficients estimated are statistically significant and the hazard function is estimated to be upward-sloping. The $J-stat$ is a test of over-identifying moment condition: 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 (64), obtained by substituting the estimated values of α_w and φ_w into (10).

$$\pi_t^w = \underset{(0.038)}{0.197\pi_{t-1}^w} + \underset{(0.000)}{0.991E_t\pi_{t+1}^w} - \underset{(0.037)}{0.193E_t\pi_{t+2}^w} - \underset{(0.006)}{0.03u_t} \tag{66}$$

Also under this specification all the coefficients are statistically significant at 5% level (standard errors computed using delta method are reported in parentheses). Our wage Phillips curve, in line with the underlying theory, is able to capture the well-known negative relation between the unemployment gap and the wage inflation, as highlighted by the negative coefficient measuring the slope of the NKWPC. In Figure 7 we report a graphical representation for the hazard and survival functions deriving from our estimation and computed respectively by using (1) and (41). The hazard clearly shows a positive slope, meaning that a time-dependent mechanism for wage adjustment emerges.

[39] See Papke and Wooldridge (2005).

[40] The estimation has been performed by using Cliff's (2003) GMM package for MATLAB available from https://sites.google.com/site/mcliffweb/programs.

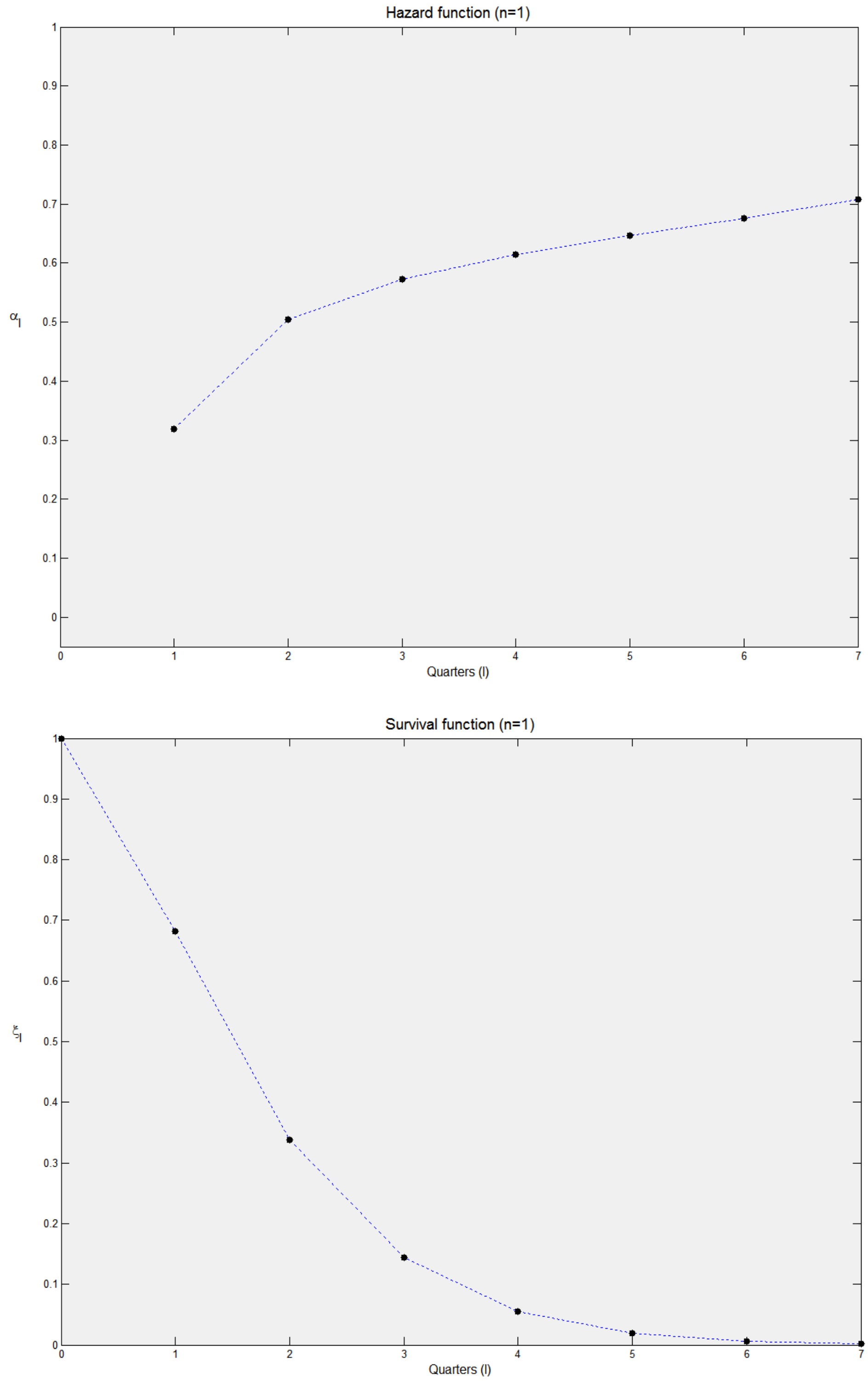

Figure 7 - Hazard and survival function deriving from GMM estimation.

Appendix C – Reduced form from the Bayesian estimation for the price and wage Phillips curves

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

$$\pi_t^p = 0.138\pi_{t-1}^p + 0.853E_t\pi_{t+1}^p + 0.01mc_t \tag{67}$$
$$\pi_t^w = 0.153\pi_{t-1}^p - 0.151\pi_t^p + 0.99E_t\pi_{t+1}^w - 0.01\mu_t^w \tag{68}$$

The EHL model with indexation *á la* Galí-Gertler (GG) implies:

$$\pi_t^p = 0.144\pi_{t-1}^p + 0.847E_t\pi_{t+1}^p + 0.009mc_t \tag{69}$$
$$\pi_t^w = 0.106\pi_{t-1}^p + 0.884E_t\pi_{t+1}^w - 0.01\mu_t^w \tag{70}$$

Finally, our time-dependent specification is associated with:

$$\pi_t^p = 0.2\pi_{t-1}^p + 0.99E_t\pi_{t+1}^p - 0.196E_t\pi_{t+2}^p + 0.012mc_t \tag{71}$$
$$\pi_t^w = 0.288\pi_{t-1}^w + 0.99E_t\pi_{t+1}^w - 0.283E_t\pi_{t+2}^w - 0.007\mu_t^w \tag{72}$$

As shown in the paper, the above reduced forms imply that models without time-dependent adjustment capture persistence in the wage equation by price inflation, whereas (71)-(72) includes a backward term for wage inflation. Both for price and wage inflation equation our Phillips curves are able to capture a higher degree of persistence, 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.

Appendix D – Price dispersion derivation

This appendix derives the relationship between price dispersion and inflation in our framework. We define $\overline{P}_t \equiv E_i \log p_t(i)$ and $\Delta_t^p = var_i\left[\log p_t(i) - \overline{P}_{t-1}\right]$, and $X_t = \log p_t(i) - \overline{P}_{t-1}$. As shown by Sheedy (2007), the price level, $\log p_t(i) = \sum_{h=0}^{\infty} \theta_h \log P_{t-h}^*$, can be written as:

$$\log p_t(i) - \overline{P}_{t-1} = (1-\alpha_p)\log p_{t-1}(i) - \varphi_p \log p_{t-2}(i) + \left(\alpha_p + \varphi_p\right)\log P_t^* - \overline{P}_{t-1}$$

adding and subtracting $\alpha_p \overline{P}_{t-1}$ to the r.h.s., it becomes:

$$\log p_t(i) - \overline{P}_{t-1} = \pi_t$$

where $\pi_t = (1-\alpha_p)\left[\log p_{t-1}(i) - \overline{P}_{t-1}\right] + \alpha_p\left[\log P_t^* - \overline{P}_{t-1}\right] + \varphi_p\left[\log P_t^* - \log p_{t-2}(i)\right]$.

Moreover, $\mu_x = E(X) = \frac{\pi_t}{1+\varphi_p}$ and

$$var(X) = \frac{(1-\alpha_p)\left[\log p_{t-1}(i) - \overline{P}_{t-1} - \mu_x\right]^2 + \alpha_p\left[\log P_t^* - \overline{P}_{t-1} - \mu_x\right]^2}{1+\varphi_p} + \\ + \frac{\varphi_p\left[\log P_t^* - \log p_{t-2}(i) - \mu_x\right]^2}{1+\varphi_p}$$

After some algebra

$$var(X) = \frac{(1-\alpha_p)\,\mu_x^2 + \alpha_p\mu_x^2 + \varphi_p\mu_x^2 - 2\mu_x\,(1-\alpha_p)\left[\log p_{t-1}(i) - \overline{P}_{t-1}\right]}{1+\varphi_p} +$$

$$+ \frac{(1-\alpha_p)\left[\log p_{t-1}(i) - \overline{P}_{t-1}\right]^2 - 2\mu_x\left\{\alpha_p\left[\log P_t^* - \overline{P}_{t-1}\right] + \varphi_p\left[\log P_t^* - \log p_{t-2}(i)\right]\right\}}{1+\varphi_p} +$$

$$+ \frac{\alpha_p\left[\log P_t^* - \overline{P}_{t-1}\right]^2 + \varphi_p\left[\log P_t^* - \log p_{t-2}(i)\right]^2}{1+\varphi_p}$$

$$= \frac{(1-\alpha_p)\,\mu_x^2 + \alpha_p\mu_x^2 + \varphi_p\mu_x^2 - 2\mu_x\,(1-\alpha_p)\left[\log p_{t-1}(i) - \overline{P}_{t-1}\right]}{1+\varphi_p} +$$

$$+ \frac{-2\mu_x\left\{\alpha_p\left[\log P_t^* - \overline{P}_{t-1}\right] + \varphi_p\left[\log P_t^* - \log p_{t-2}(i)\right]\right\}}{1+\varphi_p} +$$

$$+ \frac{\alpha_p\left[\log P_t^* - \overline{P}_{t-1}\right]^2 + \varphi_p\left[\log P_t^* - \log p_{t-2}(i)\right]^2}{1+\varphi_p}$$

Since $\log P_t^* = \frac{\pi_t + \alpha_p \log p_{t-1}(i) + \varphi_p \log p_{t-2}(i)}{\alpha_p + \varphi_p}$, we can write

$$\Delta_t^p = \frac{1-\alpha_p}{1+\varphi_p}\Delta_{t-1}^p + \frac{\alpha_p}{1+\varphi_p}\left[\frac{\pi_t + \alpha_p \log p_{t-1}(i) + \varphi_p \log p_{t-2}(i)}{\alpha_p + \varphi_p} - \overline{P}_{t-1}\right]^2 +$$

$$+ \frac{\varphi_p}{1+\varphi_p}\left[\frac{\pi_t + \alpha_p \log p_{t-1}(i) + \varphi_p \log p_{t-2}(i)}{\alpha_p + \varphi_p} - \log p_{t-2}(i)\right]^2 - \mu_x^2$$

i.e.

$$\Delta_t^p = \frac{1-\alpha_p}{1+\varphi_p}\Delta_{t-1}^p + \frac{\left(\alpha_p+\varphi_p\right)}{\left(1+\varphi_p\right)\left(\alpha_p+\varphi_p\right)^2}\pi_t^2 + \frac{\alpha_p\varphi_p^2 + \alpha_p^2\varphi_p}{\left(1+\varphi_p\right)\left(\alpha_p+\varphi_p\right)^2}\pi_{t-1}^2 - \frac{\pi_t^2}{\left(1+\varphi_p\right)^2}$$

Finally, we obtain:

$$\Delta_t^p = \left(\frac{1-\alpha_p}{1+\varphi_p}\right)\Delta_{t-1}^p + \frac{(1-\alpha_p)\,\pi_t^2 + \alpha_p\varphi_p\left(1+\varphi_p\right)\pi_{t-1}^2}{\left(1+\varphi_p\right)^2\left(\alpha_p+\varphi_p\right)} \tag{73}$$

In the case of $\varphi_p = 0$ it encompasses Calvo price dispersion. Iterating (73) forward, the degree of price dispersion in any period $t \geq 0$ under the new policy is given by:

$$\Delta_t^p = \left(\frac{1-\alpha_p}{1+\varphi_p}\right)^{t+1}\Delta_{-1}^p + \sum_{s=0}^{t}\left(\frac{1-\alpha_p}{1+\varphi_p}\right)^{t-s}\frac{(1-\alpha_p)\,\pi_s^2 + \alpha_p\varphi_p\left(1+\varphi_p\right)\pi_{s-1}^2}{\left(1+\varphi_p\right)^2\left(\alpha_p+\varphi_p\right)}$$

Then, we discount over all periods $t \geq 0$, getting (28).

References

Álvarez, L.J., P. Burriel and I. Hernando (2005), "Do decreasing hazard functions for price changes make any sense?," *ECB Working Paper Series*, No. 461.

An, S. and F. Schorfheide (2007), "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4): 113-172.

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 (2011), "Some Evidence on the Importance of Sticky Wages," *Federal Reserve Bank of Boston Working Paper*, 10-11.

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.

Benigno, P. and M. Woodford (2004), "Optimal monetary and fiscal policy: A linear-quadratic approach," *NBER Macroeconomics Annual 2003*: 271-364.

Benigno, P. and M. Woodford (2005), "Inflation stabilization and welfare: The case of a distorted steady state," *Journal of the European Economic Association*, 3(6): 1185-1236.

Benigno, P. and M. Woodford (2012), "Linear-quadratic approximation of optimal policy problems," *Journal of Economic Theory*, 147(1): 1-42.

Bils, M. and P.J. Klenow (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5): 947-985.

Blake, A.P. (2001), "A 'Timeless Perspective' on optimality in forward-looking rational expectations models," *NIESR Discussion Papers 188*, National Institute of Economic and Social Research.

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.

Brock, W.A., S.N. Durlauf, J.M. Nason and G. Rondina (2007), "Simple versus optimal rules as guides to policy," *Journal of Monetary Economics*, 54(5): 1372-1396.

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.

Dhyne, E., L.J. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünnemann, 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. and M. Gertler (1999), "Inflation Dynamics: a Structural Econometric Analysis," *Journal of Monetary Economics*, 44(2): 195-222.

Galí, J. (2008), "Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework," *Princeton University Press*.

Galí, J. (2011), "The Return of Wage Phillips Curve," *Journal of the European Economic Association*, 9(3): 436-461.

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, Development 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.

Goette, L., R. Minsch, and J.R. Tyran (2005), "Micro evidence on the adjustment of sticky-price goods: It's how often, not how much," *Discussion Papers 05-20*, University of Copenhagen, Department of Economics.

Guerrieri, L. (2001), "Inflation Dynamics," *International Finance Discussion Paper*, No. 715.

Guerrieri, L. (2002), "The Inflation Persistence of Staggered Contracts," *International Finance Discussion Paper*, No. 734.

Ikeda, D. and S. Nishioka (2007), "Price setting behavior and hazard functions: evidence from Japanese CPI microdata," *Bank of Japan Working Paper Series*, No. 07-E-19.

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.

Jensen, C. and B. McCallum, (2002), "The non-optimality of proposed monetary policy rules under timeless perspective commitment," *Economics Letters*, 77(2): 163-168.

Justiniano, A. and G. Primicieri (2008), "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 94(1): 190-217.

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.

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.

Mash, R. (2004), "Optimizing Microfoundations for Inflation Persistence," *Economics Series Working Papers,* University of Oxford, Department of Economics, No. 183.

Nakamura E. and J. Steinsson (2008), "Five Facts about Prices: A reevaluation of menu cost models," *Quarterly Journal of Economics*, 123(4): 1415-1464.

Nakamura E. and J. Steinsson (2013), "Price rigidity: Microeconomic evidence and macroeconomic implications," *Annual Review of Economics*, 5(1): 133-163.

Papke, L.E. and J.M. Wooldridge (2005), "A Computational Trick for Delta-method Standard Errors," *Economics Letters,* 86: 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.

Rotemberg, J. J. (1982), "Sticky prices in the United States," *Journal of Political Economy*, 90(6): 1187-1211.

Rotemberg, J.J. and M. Woodford (1997), "An optimization-based econometric framework for the evaluation of monetary policy," *NBER Macroeconomic Annual 1997*: 297-361.

Rotemberg, J. and M. Woodford (1999), "Interest rate rules in an estimated sticky price model," in J.B. Taylor (ed.), *Monetary Policy Rules*, The University of Chicago Press, Chicago: 57-119.

Savage, L.J. (1951), "The theory of statistical decision," *Journal of the American Statistical Association*, 46(253): 55-67.

Sbordone, A. (2007), "Inflation persistence: Alternative interpretations and policy implications," *Journal of Monetary Economics*, 54(5): 1311-1339.

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.

Söderlind (1999), "Solution and estimation of RE macromodels with optimal policy," *European Economic Review*, 43(4-6): 813-823

Steinsson, J. (2003), "Optimal monetary policy in an economy with inflation persistence," *Journal of Monetary Economics*, 50: 1425-1456.

Stoye, J. (2011), "Statistical decisions under ambiguity," *Theory and Decision*, 70(2): 129-148.

Walsh, C.E. (2010), "Monetary theory and policy," *Cambridge MIT Press*, Cambridge.

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(0): 195-214.

Woodford, M. (1999), "Commentary: How should monetary policy be conducted in an era of price stability?," *New Challenges for Monetary Policy*, Federal Reserve Bank of Kansas City, 277–316.

Woodford, M. (2002), "Inflation stabilization and welfare," *The B.E. Journal of Macroeconomics*, 2(1): 1-53.

Woodford, M. (2003), "Interest and prices: Foundations of a theory of monetary policy," *Princeton University Press*, Princeton.

Woodford, M. (2011), "Optimal monetary stabilization policy," *Handbook of Monetary Economics Volume 3B*: 723-828.

Wolman, A.L. (1999), "Sticky Prices, Marginal Cost, and the Behaviour of Inflation," *Federal Reserve Bank of Richmond Economic Quarterly*, 85(4): 29-48.

Yao F. (2011), "Non-constant Hazard Function and Inflation Dynamics," *SFB 649 Discussion Papers, 2009-030*, Humboldt University, Berlin, Germany.