Inflation persistence and price time-dependent adjustment

Marco Di Pietro

Sapienza University of Rome
Doctoral School of Economics

Thesis Supervisor: Prof. Giovanni Di Bartolomeo

Academic Year 2013/14

Foreword

This thesis aims to micro-found inflation persistence in a New Keynesian DSGE framework by recurring to time-dependent pricing mechanism. In particular, we assume that the probability to adjust a price is not exogenous, as in Calvo specification, but it is positive function of the time spent since last price reset: according to Sheedy (2007), this method is implemented by using an increasing hazard function.

In Chapter 1 we present a small review of several methods proposed in literature to make the Phillips curve able to account for inflation persistence.

Chapter 2 and Chapter 3 are devoted to a positive analysis. In particular, in Chapter 2, titled "*Wage time-dependent adjustment*", we micro-found, via time-dependent pricing relying, in turn, on the use of positive hazard functions, a Time-dependent New Keynesian Wage Phillips curve (TDWPC henceforth) embedded of a backward component and exhibiting intrinsic inertia. We then estimate this TDWPC by GMM in order to investigate the hazard shape properties for wages and what is the number of leads and lags components statistically significant for our TDWPC specification.

In Chapter 3, "*Price and wage inflation intrinsic inertia*", we develop a standard small-scale New Keynesian DSGE model characterized by time-dependent Phillips curves both for prices and wages. We then estimate this model by Bayesian techniques and show that a positive hazard function emerges for both the TDPCs; therefore, we successfully test the robustness of our result to a change of policy regime, given by the Great Moderation. Hence, we compare the empirical performance of our model, measured by the log-marginal likelihood, with that of a model where inflation persistence is achieved by some form of indexation to past inflation: we found that our model clearly outperforms the alternatives considered as, in our specification, we are able to capture

the hump-shaped response of wage inflation to a cost-push shock, coherently with the VAR evidence.

In Chapter 4, titled "*Welfare and optimal monetary policy*", we move to a normative analysis, evaluating how the distortions present in our model, due to different hazard rates, influence the social welfare and how they affect the gains deriving from a commitment.

Table of Contents

Foreword . . . . . **i**

1 Modeling inflation inertia: a review **1**

1.1 Introduction 1

1.2 Models with indexation 3

1.2.1 Galí-Gertler (1999) setting 4

1.2.2 Christiano et al. (2005) dynamic indexation 5

1.3 Sticky information models: Mankiw and Reis (2002) 6

1.4 Contracting models 7

1.4.1 Fuhrer and Moore (1995) 7

1.4.2 Blanchard and Galí (2007) 8

1.5 Trend in‡ation: Ascari and Ropele (2007) 9

1.6 Time-dependent vs. State-dependent models 11

2 Wage time-dependent adjustment **12**

2.1 Introduction 12

2.2 The hazard function 15

2.3 Wage Phillips Curve derivation 18

2.4 GMM Estimation 23

2.5 Conclusions 28

3 Price and wage in‡ation intrinsic inertia **29**

3.1 Introduction 29

3.2 The model 32

3.2.1 Hazard function and Phillips curves 33

3.2.2 Closing the model: IS curve and Taylor rule 36

3.3 Empirical analysis 37

3.3.1 Data and prior distributions 39

3.3.2 Estimation results 40

3.4 Time-dependent Phillips vs. alternatives 45

3.4.1 Alternative price-setting mechanisms: EHL with indexation . . . 46

3.4.2 Galí-Gertler setting 46

3.4.3 Model comparison 47

3.5 International evidence 50

3.6 Conclusions 51

4 Welfare and optimal monetary policy **52**

4.1 Introduction 52

4.2 The economy 57

4.3 Welfare approximation 60

4.4 Welfare analysis and optimal policy 63

4.4.1 Price-model features and distortions 63

4.4.2 Optimal policies 66

4.5 Conclusions 70

Appendix A – Time-dependent Wage Phillips curve derivation . . . **72**

A.1 Wage inflation dynamics 72

A.2 Wage Phillips curve 75

Appendix B – Robustness test **82**

B.1 Alternative time series 82

B.2 Alternative data filtering 84

B.3 Different model specifications 86

B.3.1 Model without persistence 86

B.3.2 Taylor rule AR(2) 88

Appendix C – Bayesian estimation for different countries **90**

C.1 Australia 90

C.2 United Kingdom 91

C.3 France 92

C.4 Germany 93

C.5 Italy 94

Appendix D – Phillips curves reduced form **96**

Appendix E – Welfare function derivation **97**

E.1 Proof of Lemma1 97

E.2 Final step 100

References **102**

Chapter 1

Modeling inflation inertia: a review

1.1 Introduction

New Keynesian dynamic stochastic general equilibrium models (NK DSGE henceforth) are now the workhorse for macroeconomic modelling. They are widely used in the field of macroeconomics in order to analyze the effects of monetary and fiscal policy, evaluating the impact of several frictions (e.g. hiring costs in the labor market, credit crunch), accounting for business cycle fluctuations, defining the policy rule followed by a Central Bank to implement the optimal monetary policy, analyzing the inflation process. Such class of models, in their easier specification, are simply defined by three equations: a demand schedule (called dynamic IS) derived from the Euler equation for the consumption, which defines the negative relation between the output level and the real interest rate; a supply curve, the New Keynesian Phillips curve (NKPC from now on), which describes the dynamics of the inflation process; finally, the model is closed by the monetary rule implemented by the Central Bank, that usually takes the form of a Taylor rule.[1]

Further developments on macroeconomic and econometric research allowed us to evaluate the empirical performance of these models, showing that they are able to replicate quite well the main macroeconomic figures; on the other hand, models relying

[1] For a wider analysis on this kind of rules see Taylor (1993, 2001).

on Real Business Cycle theory fail to account some important macroeconomic facts, especially the well-known productivity-employment puzzle.[2] This empirical difference is due to the hypotheses at the basis of these two approaches: whereas RBC models are developed in a framework characterized by price flexibility and perfect competition, NK DSGE models assume a market structure with monopolistic competition and nominal rigidity under the form of sluggish adjustment of prices.

The stickiness of prices is usually introduced through the Calvo (1983) mechanism: under this framework in each period a proportion of firms, θ, cannot change their price, whereas the remaining fraction $(1-\theta)$ can reoptimize. The firm able to adjust its price sets it in order to maximize the expected discounted flow of future profits: the maximization takes in account also future profits since the firm knows that the price set will remain fixed for more than one period. Thus, Calvo pricing model involves infrequent price change based on an exogenous probability to reset a price: this probability, equal to $(1-\theta)$, follows an exogenous Poisson process and the price duration is given by $1/(1-\theta)$. Under this pricing method it is possible to derive a NKPC that takes the following form:[3]

$$\widehat{\pi}_t^p = \beta E_t \widehat{\pi}_{t+1}^p + \kappa \widehat{mc}_t \tag{1}$$

where $\widehat{\pi}_t^p$ denotes the price inflation, $\widehat{mc}_t$ is the real marginal cost, β is the stochastic discount factor and $\kappa = \frac{(1-\theta)(1-\beta\theta)}{\theta}$ measures the elasticity of inflation to real marginal cost. As described previously, this relation tells us that inflation positively depends on the expectation on future inflation and on the real marginal cost. Alternatively, (1) can be rewritten considering the output gap instead of real marginal cost: in fact,

[2] For a discussion over this issue see Galí (1999) and Basu *et al.* (2006).

[3] The hat over a variable denotes the log-deviation from the steady state.

these two variables are proportional and this allows us to relate inflation directly to the output.

A problem deriving with this NKPC specification is that it lacks of backward terms, so it totally fails to account for the intrinsic persistence component of inflation. Looking at the macro data, the inflation series exhibits a high degree of autocorrelation, denoting the necessity to add a lagged term to (1), treating in this way inflation as an auto-regressive process.

In last years a large stream of macroeconomic research has been devoted to develop mechanisms to add inertia in the NKPC and able to fit the data better than (1). These models differ among them mainly for the assumption on which relies the pricing decision. Time by time have been proposed pricing models based on indexation, sticky information, real wage rigidity, positive steady-state inflation, time-dependent and state-dependent price adjustment

All these models try to provide micro-foundations in order to obtain a NKPC embedding at least a lagged term, being so able to account for the inertia observed in the data. Each approach will be analyzed more in depth in next sections.

1.2 Models with indexation

The earlier contribution to endow the NKPC of a backward term is made by introducing indexation to past inflation. In this way it is possible to derive a Phillips curve where current inflation exhibits inertial behavior since it depends also on past inflation. A shortcoming of this approach is that indexation to past inflation is not always verified in the micro data (Fabiani *et al.*, 2005; Dhyne *et al.*, 2005), in particular since the Great Moderation. Indexation is introduced by assuming a fraction of rule of thumb

firms in resetting their price (Galí and Gertler, 1999) or through dynamic indexation to past inflation (Christiano *et al.*, 2005).

1.2.1 Galí-Gertler (1999) setting

In their seminal work Galí and Gertler (1999) derived a hybrid Phillips curve characterized by both forward and backward terms for inflation. Their framework is similar to Calvo, but they depart from the latter with regard to the behaviour of the firms able to reoptimize their price. As in Calvo a fraction of firms, θ, cannot reoptimize their price and continue to charge the price used in the previous period. The remaining fraction, $(1-\theta)$, is composed by the firms able to change their price. These firms are divided into two groups: a forward-looking fraction, $(1-\omega)$, is able to reset the price and this is set in order to maximize the expected flow of profits; on the other side, the remaining fraction, ω, is composed by backward-looking firms that will index their price to past inflation. Thus, the aggregate price level evolves in the following way:

$$p_t = (1-\theta)\, p_t^* + \theta p_{t-1} \tag{2}$$

The evolution of p_t^* is then given by:

$$p_t^* = (1-\omega)\, p_t^f + \omega p_t^b \tag{3}$$

where p_t^f is the price set by forward-looking firms and p_t^b denotes the price set by backward-looking firms. The price p_t^f is fixed in the same way of the traditional Calvo

model, whereas p_t^b is set according to:

$$p_t^b = p_{t-1}^* + \pi_{t-1}^p \tag{4}$$

This approach allows us to derive a hybrid NKPC specified as follows:

$$\widehat{\pi}_t^p = \gamma_f \beta E_t \widehat{\pi}_{t+1}^p + \gamma_b \widehat{\pi}_{t-1}^p + \nu \widehat{mc}_t \tag{5}$$

where $\gamma_f = \frac{\theta}{\theta+\omega[1-\theta(1-\beta)]}$, $\gamma_b = \frac{\omega}{\theta+\omega[1-\theta(1-\beta)]}$ and $\nu = \frac{(1-\theta)(1-\beta\theta)(1-\omega)}{\theta+\omega[1-\theta(1-\beta)]}$.

In this specification ω is the degree of indexation to past inflation, whereas γ_b represents the degree of backwardness of inflation. Moreover, Galí and Gertler (1999) provide an estimation of (5) performed via GMM: they obtain that the degree of backwardness is quite significant, but the forward component is still dominant (about 0.75, against a value of 0.25 for the backward term).[4]

1.2.2 Christiano *et al.* (2005) dynamic indexation

Another contribution to this issue comes from Christiano *et al.* (2005). Their approach is a variant of Calvo where the only difference is that they assume automatic indexation to past inflation for the share of firms not able to change their price. The aggregate price level evolves according to:

$$p_t = (1-\theta)\, p_t^* + \theta\omega\pi_{t-1}^p p_{t-1} \tag{6}$$

where p_t^* is the reset price chosen by a firm in order to maximize his expected flow

[4] The reader is referred to Galí and Gertler (1999) for the complete results and robustness analysis about this NKPC specification.

of profits and the parameter $\omega \in (0,1)$ measures the degree of indexation to past inflation.[5] The corresponding Phillips curve is:

$$\widehat{\pi}_t^p = \frac{\beta}{1+\omega\beta} E_t \widehat{\pi}_{t+1}^p + \frac{\omega}{1+\omega\beta} \widehat{\pi}_{t-1}^p + \varkappa \widehat{mc}_t \tag{7}$$

where $\varkappa = \frac{(1-\theta)(1-\beta\theta)}{\theta(1+\omega\beta)}$. This framework implies that each period there will be a price change, due either to price reoptimization or to dynamic indexation: this is at odds with microeconomic evidence since firms can take their price fixed also for long periods.

1.3 Sticky information models: Mankiw and Reis (2002)

A further possible explanation to inflation persistence is given by Mankiw and Reis (2002) in their sticky information model: they assume that information, and not price, is sticky and borrow Calvo mechanism to micro-found the sluggish adjustment in information. Under this framework there is an exogenous probability that a fraction λ of the firms receive new information about the state of the economy and, as a consequence, it is able to reoptimize his price, fixing a plan for its future prices; this pricing model is an extension of Fischer (1977), as prices are predetermined. The remaining fraction $(1-\lambda)$ does not receive any information and leaves its price unchanged. At time t the optimal price of a representative firm is:

$$p_t^* = p_t + \alpha y_t \tag{8}$$

The firm is concerned about its optimal price that is set accordingly with the

[5] In their original specification Christiano *et al.* (2005) assume that the price is fully indexed to past inflation (i.e. $\omega = 1$).

variation in the output gap. A firm that received last information i periods ago updates its price in the following way:

$$x_t = E_{t-i} p_t^* \tag{9}$$

Therefore the aggregate price level is:

$$p_t = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i x_t \tag{10}$$

By combining equations (8), (9) and (10) emerges a Phillips curve with the following form:

$$\pi_t^p = \left[\frac{\alpha\lambda}{1-\lambda}\right] y_t + \lambda \sum_{i=0}^{\infty} (1-\lambda)^i E_{t-1-i} \left[\pi_t + \alpha (y_t - y_{t-1})\right] \tag{11}$$

Here, inflation depends on output gap and past expectations of current inflation and output gap growth. Since these variables are predetermined, they introduce an inertial component in the Phillips curve.

1.4 Contracting models

This class of models concentrates on real wage rigidity and its implication on the Phillips curve specification. We shall analyze two contributions: Fuhrer and Moore (1995) and Blanchard and Galí (2007).

1.4.1 Fuhrer and Moore (1995)

Fuhrer and Moore (1995) depart from Taylor contracting model (1980) assuming that agents care about relative real wage. The nominal wage contract proposed by Fuhrer

and Moore has the following form:

$$w_t - p_t = \frac{1}{2}\left[w_{t-1} - p_{t-1} + E_t\left(w_{t+1} - p_{t+1}\right)\right] + \delta y_t \tag{12}$$

where w_t is the wage contract negotiated at time t and y_t is the excess of demand. Under this perspective the price level is defined as:

$$p_t = \frac{1}{2}\left(w_t + w_{t-1}\right) \tag{13}$$

Then, by substituting (12) in (13) and rearranging we get:

$$\widehat{\pi}_t^p = \frac{1}{2}\left(\widehat{\pi}_{t-1}^p + E_t\widehat{\pi}_{t+1}^p\right) + \delta\left(y_t + y_{t-1}\right) \tag{14}$$

It is clear as (14) is able to account for the persistence own of the inflation process, since appears also a backward term for inflation. Differently from Galí and Gertler (1999) and Christiano *et al.* (2005) this model has a further form of inertia given due to the presence of a lagged term for the output. Moreover, the weight attached to past inflation is higher than the one estimated by Galí and Gertler (1999).

1.4.2 Blanchard and Galí (2007)

Blanchard and Galí (2007) show that introducing real wage rigidities it is possible to derive a Phillips curve able to account for inflation inertia. This approach relies on the assumption that real wage evolves as follows:

$$\omega_t = \tau\omega_{t-1} + (1-\tau)\, mrs_t \tag{15}$$

where τ measures the degree of real wage rigidity. Thus, the Phillips curve is:

$$\widehat{\pi}_t^p = \beta E_t \widehat{\pi}_{t+1}^p + \frac{\lambda}{1-\tau L} x_{2t} \tag{16}$$

where L is the lag operator, $\kappa = [(1-\theta)(1-\beta\theta)]\,\theta^{-1}$, and x_{2t} is a linear combination of the output and output growth depending on the extent of real wage rigidity; x_{2t} is given by:

$$x_{2t} = (1-\sigma)^{-1}\left[(1-\tau)(1+\gamma)\,y_t + \sigma\tau\,(y_t - y_{t-1})\right] \tag{17}$$

where $(1-\sigma)$ represents the labor share and γ is the labor supply elasticity. The key parameter for the persistence in this model is τ: the higher is τ, the more will account the inertial component in (17).

1.5 Trend inflation: Ascari and Ropele (2007)

An interesting innovation is the one introduced by Ascari (2004) with regard to the deterministic steady-state around we take the log-linearization. His point of departure is that, in post-war period, for all the advanced economies we observe positive inflation rates and, moreover, many Central Banks adopt a policy rule based on inflation targeting on slight, but positive, inflation rate (usually around 2%). Thus, the common practice to log-linearize NK DSGE models around a steady-state characterized by zero inflation is unrealistic. His suggestion is to take the log-linearization around a positive steady-state inflation; under this modification the Phillips curve is characterized by intrinsic inertia, deriving from the behaviour of the relative price dispersion measure. According to Ascari and Ropele (2007), the "Trend inflation" approach leads to a

Phillips curve that has the following form:

$$\widehat{\pi}_t^p = \beta E_t \widehat{\pi}_{t+1}^p + \kappa \left[(1+\sigma_n) \widehat{y}_t + \sigma_n \widehat{s}_t \right] + \frac{\theta\beta(\overline{\pi}-1)(1-\theta)}{\theta} E_t \left[(\varepsilon - 1) \widehat{\pi}_{t+1}^p + \widehat{\phi}_{t+1} \right] \quad (18)$$

where $\overline{\pi}$ is the steady-state of inflation, $\widehat{s}_t$ is the relative price dispersion, $\kappa = (1-\theta)(1-\beta\theta)/\theta$, ε is the elasticity of substitution between goods, σ_n is the inverse of Frisch elasticity and $\widehat{\phi}_t$ is an auxiliary variable defined as follows:

$$\widehat{\phi}_t = \theta\beta E_t \left[(\varepsilon - 1) \widehat{\pi}_{t+1}^p + \widehat{\phi}_{t+1} \right] \quad (19)$$

The relative price dispersion evolves as follows:

$$\widehat{s}_t = \xi \widehat{\pi}_t^p + \theta \overline{\pi}^{\varepsilon} \widehat{s}_{t-1} \quad (20)$$

where ξ is a convolution of deep parameters and trend inflation; the variable $\widehat{s}_t$ plays a pivotal role since it is characterized by a backward term which arises in a Phillips curve embedded by intrinsic persistence. As we can see from (18), the steady-state level of trend inflation has an impact on the determination of current inflation.

Some things are worth of mention: first, if we consider $\overline{\pi} = 1$, corresponding to zero inflation steady-state, equation (18) reduces to the standard NKPC in equation (1). In addition, taking a log-linearization around a positive inflation steady-state has non-negligible effects on the dynamics of the model: the short-run properties of the model depend on the value of trend inflation assumed for the log-linearization. Moreover, a log-linear approximation around a non-zero steady-state inflation rate has implication also for the long-run properties of the model: the steady-state value of

the output gap is affected by the level of trend inflation assumed. Finally, as showed in Ascari and Ropele (2007), the steady-state inflation level influences also the determinacy/indeterminacy regions: they find that the Taylor principle seems to be very sensible to trend inflation and the determinacy region tends to shrink when a positive level of steady-state inflation is assumed.

1.6 Time-dependent vs. State-dependent models

Two further classes of pricing models are the time-dependent (TDP) and state-dependent (SDP): they differentiate one another for the assumptions on the mechanism that give rise to price change.

In time-dependent models the price adjustment depends on the time spent since last price reset: these models usually recur to a hazard function to micro-found this mechanism. Therefore, the probability to change price could be endogenous, being positive or negative function of the time (depending on the hazard shape assumed) or it can be exogenous, i.e. it is always the same regardless the time spent from last reset (e.g. Calvo). Under TDP approach, prices can be reoptimized each period, as in the standard model, or can be predetermined for multi periods (Taylor, 1980).

On the other side we have state-dependent models: here, pricing decisions are taken in order to face a shock that alters the state of the economy. Thus, price change happens whether the benefit deriving from setting a new price offsets the menu cost supported by a firm. Therefore, the timing of the price change depends endogenously on the underlying economic conditions. A shortcoming of SDP models, for which they are rarely used in literature, is their complex analytic tractability.

Chapter 2

Wage time-dependent adjustment

2.1 Introduction

In this chapter we depart from Calvo hypothesis of random probability of wage adjustment by assuming that there is a positive relation between the time spent from last wage reset and the probability to change a wage. In particular we follow the approach designed by Sheedy (2007) who uses a non-constant hazard function to micro-found the price adjustment process. The idea is that this process is time-dependent: a hazard function with positive slope can well approximate this mechanism. 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 focus on 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",[6] which is a stylized economic fact hard

[6] 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.

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 work, 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 work, which focuses on the micro-foundations for the intrinsic inertia of inflation observed in macro data. For macroeconomic issues it is important to analyze 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.

In what follows we extend the work of Sheedy by investigating if a positive hazard function also emerges for wage adjustment. In particular we derive a Wage Phillips curve where the wage reset process has a time-dependent structure, obtaining in this way a TDWPC with both forward and backward terms, able to exhibit intrinsic persistence. Moreover, we empirically test, via GMM estimation, the statistical significance of the forward/backward terms emerging from the Phillips curve; we also test if the hazard is able to satisfy some conditions with regard to its slope and shape. 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.

The chapter is organized as follows. In the next section we strictly follow Sheedy (2007) in presenting the hazard function main characteristics; Section 2.3 presents the derivation of the Wage Phillips curve based on the time-dependent hypothesis; Section 2.4 shows the empirical test performed to investigate how this approach works; a final section concludes.

2.2 The hazard function

Following Sheedy (2007), we assume that wages are set according to a time-dependent mechanism: the probability to post a new wage positively depends on the time elapsed since last wage reset. This adjustment process can be formalized by using a hazard function.[7]

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:

$$D_t(j) \equiv \min \{l \geq 0 \mid j \in \Gamma_{t-l}\} \tag{21}$$

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: $\alpha_l \equiv \Pr(\Gamma_t \mid D_{t-1} = l-1)$.

The hazard function is specified as follows:

$$\alpha_l = \alpha + \sum_{h=1}^{\min(l-1,n)} \varphi_h \left[\prod_{k=l-h}^{l-1} (1-\alpha_k) \right]^{-1}, \tag{22}$$

where α_l is the probability to change a wage which last reset was l periods ago; α is the initial value of the hazard function, φ_j is its slope; n is the number of parameters that

[7] 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.

control the slope – for $n = 1$, the slope is governed by only one parameter, $\varphi_j = \varphi$.[8]

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:

$$\varsigma_l = \prod_{h=1}^{l} (1 - \alpha_h) \tag{23}$$

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

$$\begin{cases} \alpha_1 < 1, \text{ meaning that is allowed a degree of wage stickiness;} \\ \alpha_\infty > 0, \text{ with } \alpha_\infty = \lim_{l \to \infty} \alpha_l. \end{cases} \tag{24}$$

The hazard function can be reparameterized by making use of a set of $n + 1$ parameters and rewritten as (22), where $\{\varphi_l\}_{l=1}^{n}$ is a set of n parameters that control the hazard slope, whereas parameter α controls its initial level.

By making use of (23), we can rewrite the non-linear recursion (22) for the wage adjustment probabilities as a linear recursion for the corresponding survival function:

$$\varsigma_l = (1 - \alpha)\varsigma_{l-1} - \sum_{h=1}^{\min(l-1,n)} \varphi_h \varsigma_{l-1-h} \tag{25}$$

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

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

$$\begin{cases} \varphi_l = 0, \text{ for all } l = 1, ..., n \text{ the hazard is flat (Calvo case);} \\ \varphi_l \geq 0, \text{ for all } l = 1, ..., n \text{ the hazard is upward-sloping;} \\ \varphi_l \leq 0, \text{ for all } l = 1, ..., n \text{ the hazard is downward-sloping.} \end{cases} \tag{26}$$

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:

$$\begin{cases} \theta_{0t} = \sum_{l=1}^{\infty} \alpha_l \theta_{l-1,t-1} \\ \theta_{lt} = (1 - \alpha_l)\, \theta_{l-1,t-1} \end{cases} \tag{27}$$

If the hazard function satisfies the restrictions (24) and the evolution over the time of the distribution of wage duration evolves as in (27), 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 (22) and assume that the economy has converged to $\{\theta_l\}_{l=0}^{\infty}$, the following three relations are obtained:

$$\begin{cases} \theta_l = \left(\alpha + \sum_{h=1}^{n} \varphi_h\right) \varsigma_l \\ \alpha^e = \alpha + \sum_{l=1}^{n} \varphi_l \\ D^e = \frac{1 - \sum_{l=1}^{n} l\varphi_l}{\alpha + \sum_{l=1}^{n} \varphi_l} \end{cases} \tag{28}$$

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

2.3 Wage Phillips Curve derivation

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:

$$N_t(i) = \left[\int_{\Omega} N_t(i,j)^{\frac{\varepsilon_w - 1}{\varepsilon_w}} dj \right]^{\frac{\varepsilon_w}{\varepsilon_w - 1}} \tag{29}$$

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:

$$N_t(i,j) = \left(\frac{W_t(j)}{W_t} \right)^{-\varepsilon_w} N_t(i) \tag{30}$$

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:

$$W_t = \left[\int_0^1 W_t(j)^{1-\varepsilon_w} dj \right]^{\frac{1}{1-\varepsilon_w}} \tag{31}$$

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 $N_t(j) =$

$\int_{\Omega} N_t(i,j)di$ 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:

$$U(C_t(j), N_t(j)) = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[g_t \frac{(C_t(j) - hC_{t-1}(j))^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\gamma}(j)}{1+\gamma} \right] \right\} \tag{32}$$

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:

$$P_t(j) C_t(j) + E_t [Q_{t+1,t} B_t(j)] \leq B_{t-1}(j) + W_t(j) N_t(j) + T_t(j) \tag{33}$$

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: $C_t(j) = \left(\int_{\Theta} C_t(i,j)^{\frac{\varepsilon_p - 1}{\varepsilon_p}} di \right)^{\frac{\varepsilon_p}{\varepsilon_p - 1}}$, 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 (32) internalizing the effects of labor demand (30) and taking account of (33). 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^*,[9] if the household still earns this wage at time

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

$\tau \geq t$ then its relative wage will be W_t^*/W_τ and the household utility can be written as $U\left[W_t^*/W_\tau; C_{\tau|t}; N_{\tau|t}\right]$;[10] by considering the survival function, the household will then choose its optimal reset wage by solving:

$$\max_{W_t^*} \sum_{\tau=t}^{\infty} \varsigma_{\tau-t} E_t \left\{ \left(\prod_{s=t+1}^{\tau} \frac{\pi_s^p}{I_s} \right) U \left[\frac{W_t^*}{W_\tau}; C_{\tau|t}; N_{\tau|t} \right] \right\} \tag{34}$$

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 (33) and the labor demand schedule (30). Equation (34) yields the following first-order condition:

$$\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} E_t \left(\frac{W_t^*}{W_\tau} \right)^{-\varepsilon_w} \left(\prod_{s=t+1}^{\tau} \frac{\pi_s}{I_s} \right) \left[U_c(C_{\tau|t}, N_{\tau|t}) \frac{N_{\tau|t}}{P_\tau} (1-\varepsilon_w) + \right.$$
$$\left. -\varepsilon_w U_n(C_{\tau|t}, N_{\tau|t}) \frac{N_{\tau|t}}{W_\tau} \frac{W_\tau}{W_t^*} \right] = 0 \tag{35}$$

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 (35) can be rearranged and expressed in terms of the optimal wage reset as:

$$W_t^* = \left[\frac{\mu_w \left(\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \beta^{\tau-t} MRS_{\tau|t} P_\tau \right)}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t} \beta^{\tau-t}} \right] \tag{36}$$

Assuming that the economy has converged to $\{\theta_l\}_{l=0}^{\infty}$, then the wage level (31) can

[10] $C_{\tau|t}$ 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.

be expressed as a weighted-average of the past reset wages:

$$W_t = \left(\sum_{l=0}^{\infty} \theta_l W_{t-l}^{*\,1-\varepsilon_w} \right)^{\frac{1}{1-\varepsilon_w}} \tag{37}$$

By log-linearizing (36) and (37) around a steady-state (characterized by zero wage inflation), we get:[11]

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\beta^{\tau-t} \varsigma_{\tau-t}}{\sum_{j=0}^{\infty} \beta^j \varsigma_j} \right) [w_\tau - \Xi_w \mu_\tau^w] \tag{38}$$

$$w_t = \sum_{l=0}^{\infty} \theta_l w_{t-l}^* \tag{39}$$

where $\Xi_w = \frac{1}{1+\varepsilon_w \gamma}$. Equations (38) and (39) describe the wage adjustment mechanism. The time-dependent wage Phillips curve is derived by combining them with (25) and (28).

Specifically, by combining (25) with (38), we obtain:

$$w_t^* = \beta(1-\alpha) E_t w_{t+1}^* - \sum_{l=1}^{n} \beta^{l+1} \varphi_l E_t w_{t+l+1}^* + \left[1 - \beta(1-\alpha) + \sum_{l=1}^{n} \beta^{l+1} \varphi_l \right] (w_t - \Xi_w \mu_t^w) \tag{40}$$

By making use of (28), equation (39) can be recast as follows:

$$w_t = (1-\alpha)\, w_{t-1} - \sum_{l=1}^{n} \varphi_l w_{t-1-l} + \left(\alpha + \sum_{h=1}^{n} \varphi_h \right) w_t^* \tag{41}$$

where we have used the fact that the stationary distribution of the wage duration (28)

[11] Small-caps letters denote log-deviation from the steady-state. Details about the TDWPC derivation are provided in Appendix A.

can be rewritten in recursive way as:

$$\theta_l = (1-\alpha)\theta_{l-1} - \sum_{h=1}^{\min(l-1,n)} \varphi_h \theta_{l-h-1} \tag{42}$$

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

The general expression for the wage Phillips curve is obtained from (40) and (41):

$$\pi_t^w = \sum_{l=1}^{n} \psi_l \pi_{t-l}^w + \sum_{l=1}^{n+1} \delta_l E_t \pi_{t+l}^w - k_w \mu_t^w \tag{43}$$

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

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

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 (43) has the following specification:

$$\pi_t^w = \psi_w \pi_{t-1}^w + \beta\left[1 + (1-\beta)\psi_w\right] E_t \pi_{t+1}^w - \beta^2 \psi_w E_t \pi_{t+2}^w - k_w \mu_t^w, \tag{44}$$

with

$$\begin{cases} \psi_w = \frac{\varphi_w}{(1-\alpha_w)-\varphi_w[1-\beta(1-\alpha_w)]} \\ k_w = \frac{(\alpha_w+\varphi_w)\left[1-\beta(1-\alpha_w)+\beta^2\varphi_w\right]}{(1-\alpha_w)-\varphi_w[1-\beta(1-\alpha_w)]}\Xi_w \end{cases}, \tag{45}$$

where π_t^w is the wage inflation, ψ_w and k_w are coefficients depending on the hazard parameters (φ_w and α_w control respectively the slope and the initial level of the hazard function); $\Xi_w = \frac{1}{1+\varepsilon_w\gamma}$.

2.4 GMM Estimation

As in Sheedy (2007) we estimate our wage Phillips curve via generalized method of moments (GMM): we want to investigate if a positive hazard function for wages emerges (i.e. $\varphi_w > 0$) and what is the number of backward/forward component of (43) statistically significant, in order to correctly specify the number of parameters affecting the hazard slope. 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):

$$\mu_t^w = \gamma u_t \tag{46}$$

where u_t represents the unemployment gap. Therefore (43) becomes:

$$\pi_t^w = \sum_{l=1}^{n} \psi_l \pi_{t-l}^w + \sum_{l=1}^{n+1} \delta_l E_t \pi_{t+l}^w - k_w \gamma u_t \tag{47}$$

To perform a GMM estimation of (47) 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:

$$E_t\left[\left(\pi_t^w - \sum_{l=1}^{n}\psi_l\pi_{t-l}^w - \sum_{l=1}^{n+1}\delta_l E_t\pi_{t+l}^w + k_w\gamma u_t\right)z_{t-1}\right] = 0 \tag{48}$$

Following Galí and Gertler (1999), since (48) is non-linear in the structural parameters, we normalize the orthogonality condition in the following way:

$$E_t\left[\left(\chi\pi_t^w - \chi\sum_{l=1}^{n}\psi_l\pi_{t-l}^w - \chi\sum_{l=1}^{n+1}\delta_l E_t\pi_{t+l}^w + \chi k_w\gamma u_t\right)z_{t-1}\right] = 0 \tag{49}$$

Our estimation is made using quarterly U.S. data ranging from 1960:1 to 2011:4: all the data comes from FRED database maintained by the Federal Reserve Bank of St. Louis. 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. For the sake of brevity we only show the estimation of (49) when $n = 1$.[12] Under the latter assumption, (48) and (49) change as follows:

$$E_t\left\{\left[\pi_t^w - \psi_w\pi_{t-1}^w - \beta(1+(1-\beta)\,\psi_w)E_t\pi_{t+1}^w + \beta^2\psi_w E_t\pi_{t+2}^w + k_w\gamma u_t\right]z_{t-1}\right\} = 0 \tag{50}$$

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

$$E_t \left\{ \frac{1}{\chi_w} \left[\pi_t^w - \psi_w \pi_{t-1}^w - \beta(1 + (1-\beta)\,\psi_w) E_t \pi_{t+1}^w + \beta_w^2 \psi_w E_t \pi_{t+2}^w + k_w \gamma u_t \right] z_{t-1} \right\} = 0 \tag{51}$$

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

The structural form of (51) is estimated by imposing $\beta = 0.99$, $\gamma = 2$, and $\varepsilon_w = 8.85$; the elasticity of substitution between workers is derived as in Galí (2011) by using $\varepsilon_w = \left[1 - \exp\left(-\gamma u^n\right)\right]^{-1} = 8.85$, where we assume a natural unemployment rate u^n equal to 6%, as the average rate of the period considered herein. The reduced form coefficients (45) 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.[13]

The results for the structural form estimation are reported in Table 1. 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 (28)) and the $J - stat$.

Table 1 – Wage Phillips curve estimation (structural form)[14]

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 estimated coefficients 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

[13] See Papke and Wooldridge (2005).

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

are satisfied (the model is "valid"). We now report the reduced form of (50), obtained by substituting the estimated values of α_w and φ_w into (45).

$$\pi_t^w = \underset{(0.038)}{0.197\pi_{t-1}^w} + \underset{(0.000)}{0.991E_t\pi_{t+1}^w} - \underset{(0.037)}{0.193E_t\pi_{t+2}^w} - \underset{(0.006)}{0.03u_t} \tag{52}$$

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 TDWPC. In Figure 1 we report a graphical representation for the hazard and survival functions deriving from our estimation and computed respectively by using (22) and (25). The hazard clearly shows a positive slope, meaning that a time-dependent mechanism for wage adjustment emerges.

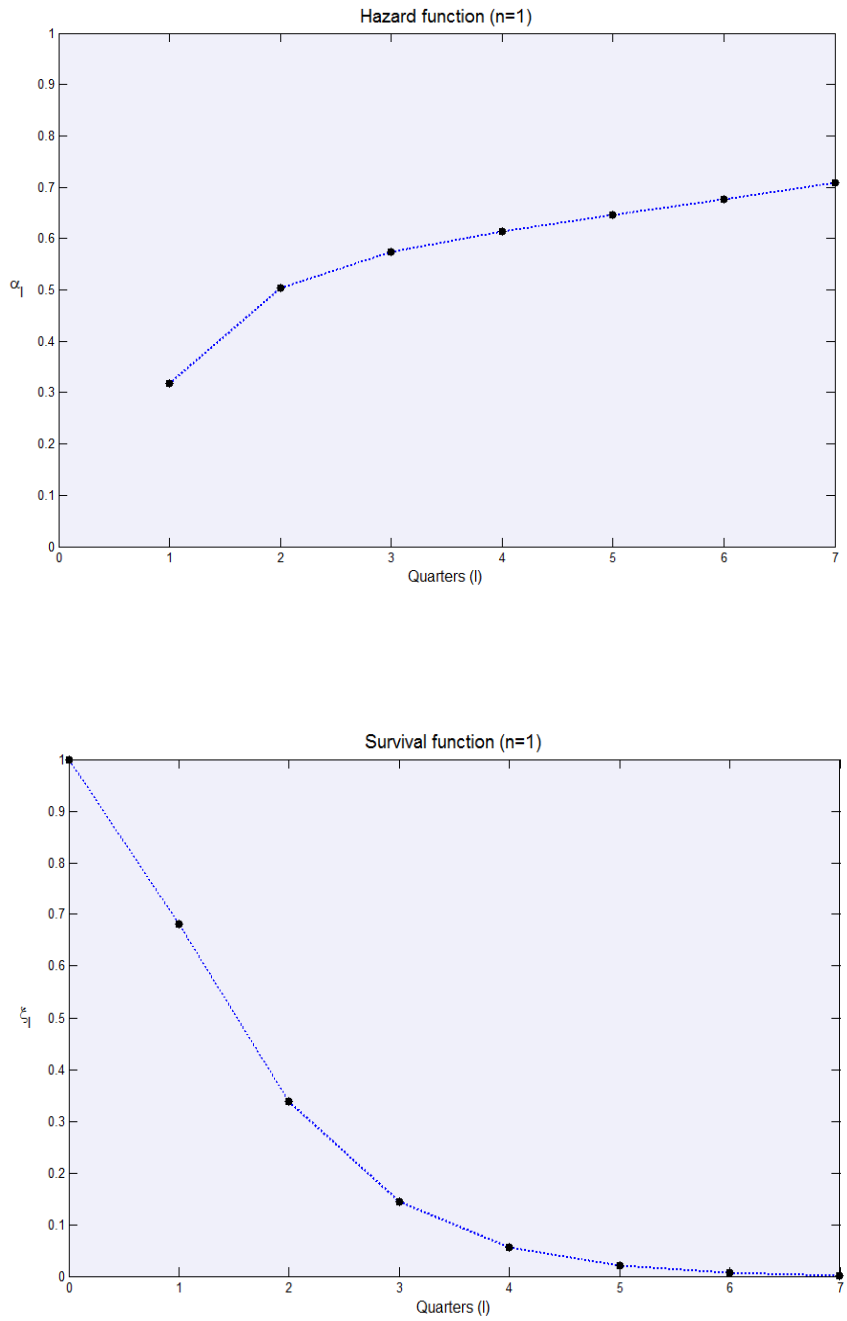

Figure 1 - Hazard and survival function deriving from GMM estimation.

2.5 Conclusions

In this chapter we have extended the results found by Sheedy (2007) to wage inflation. Assuming monopolistic competition on labor market and staggered wage setting, we derived a Wage Phillips curve able to account for inflation persistence, simply relying on the assumption of time-dependent adjustment based on the use of a positive hazard function. The empirical test performed confirms that a positive hazard function emerges for wage adjustment process. Thus, under this framework, we are able to obtain a Phillips curve exhibiting intrinsic persistence, without the necessity to impose some form of indexation.

Chapter 3

Price and wage inflation intrinsic inertia

3.1 Introduction

In this chapter we derive and estimate 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 the time occurred since the last reset; thus, we consider price and wage Phillips curves including backward-looking terms, which are, therefore, endogenously derived.

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).[15]

[15] 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.

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 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. 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.

Our model is estimated for U.S. economy with Bayesian estimation techniques. After writing 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) analyzes several models to build inflation persistence including Sheedy (2007).[16] He finds evidence of positive-sloping hazard functions,

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

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).

The main contributions of our work are three and can be summarized as follows. First, the micro-foundation of wage inflation persistence, relying on the assumption of time-dependent adjustment, allows us obtain hump-shaped response of 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.

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. Moreover, we provide estimations of our model for several countries that switched their policy regime and show that, in each case, a positive hazard functions still emerge.

The rest of the chapter is organized as follows. In the next section, after briefly 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 3.3 presents ours model estimations, whereas Section 3.4 compares them to EHL and its extension with different kind of inflation indexation. Section 3.5 shows estimations for a set of countries that recently changed their monetary regime. A final section concludes.

3.2 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). We report the log-linear deviations from the steady state.

3.2.1 Hazard function and Phillips curves

According to Sheedy (2007),[17] 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:

$$\alpha_i = \alpha + \sum_{j=1}^{\min(i-1,n)} \varphi_j \left[\prod_{k=i-j}^{i-1} (1 - \alpha_k) \right]^{-1}, \tag{53}$$

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, $\varphi_j = \varphi$.[18]

By using (53), as shown by Sheedy (2007), we can derive a price Phillips curve that depends on both expected[19] and past inflation. Formally:

$$\pi_t^p = \psi_p \pi_{t-1}^p + \beta \left[1 + (1 - \beta) \psi_p \right] E_t \pi_{t+1}^p - \beta^2 \psi_p E_t \pi_{t+2}^p + k_p \left(mc_t + \zeta_t \right), \tag{54}$$

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

[17] 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.

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

[19] 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.

the parameters characterizing the hazard function:

$$\begin{cases} \psi_p = \frac{\varphi_p}{(1-\alpha_p)-\varphi_p[1-\beta(1-\alpha_p)]} \\ k_p = \frac{\left(\alpha_p+\varphi_p\right)\left[1-\beta(1-\alpha_p)+\beta^2\varphi_p\right]}{(1-\alpha_p)-\varphi_p[1-\beta(1-\alpha_p)]}\eta_{cx} \end{cases} \tag{55}$$

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 (53)); $\eta_{cx} = \frac{1-\phi}{1-\phi+\phi\varepsilon_p}$ 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:

$$y_t = a_t + (1-\phi)\, n_t, \tag{56}$$

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:

$$mc_t = \omega_t + n_t - y_t, \tag{57}$$

where ω_t denotes the real wage.

By definition, the real wage dynamics is described by:

$$\omega_t - \omega_{t-1} = \pi_t^w - \pi_t^p. \tag{58}$$

The marginal rate of substitution, mrs_t, between consumption and hours worked

is given by:

$$mrs_t = \frac{\sigma}{1-h}\left(y_t - hy_{t-1}\right) + \gamma n_t - g_t, \tag{59}$$

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:

$$\mu_t^w = \omega_t - mrs_t. \tag{60}$$

We assume monopolistic competition on the labor market and derive a New Keynesian wage Phillips curve that exhibits intrinsic inflation persistence from the hazard function. Formally:[20]

$$\pi_t^w = \psi_w \pi_{t-1}^w + \beta\left[1 + (1-\beta)\psi_w\right] E_t \pi_{t+1}^w - \beta^2 \psi_w E_t \pi_{t+2}^w - k_w \mu_t^w, \tag{61}$$

with

$$\begin{cases} \psi_w = \frac{\varphi_w}{(1-\alpha_w)-\varphi_w[1-\beta(1-\alpha_w)]} \\ k_w = \frac{(\alpha_w+\varphi_w)\left[1-\beta(1-\alpha_w)+\beta^2\varphi_w\right]}{(1-\alpha_w)-\varphi_w[1-\beta(1-\alpha_w)]}\Xi_w \end{cases}, \tag{62}$$

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 (61) is clear that the mechanism studied in our paper provides

[20] Equation (61) is derived in Chapter 2.

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.

3.2.2 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:

$$y_t = \frac{1}{1+h} E_t y_{t+1} + \frac{h}{1+h} y_{t-1} - \frac{1-h}{\sigma(1+h)} \left(i_t - E_t \pi^p_{t+1} + E_t g_{t+1} - g_t \right), \tag{63}$$

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:

$$i_t = \rho_r i_{t-1} + (1-\rho_r)\left(\delta_\pi \pi^p_t + \delta_x y_t\right) + \upsilon_t, \tag{64}$$

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,[21] all the shocks considered in the model

[21] Monetary policy persistence is already captured by the lagged term in (64). 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 (64). Results are showed in Appendix B.

follow an $AR(1)$ process:

$$\begin{cases} a_t = \rho_a a_{t-1} + \varepsilon_t^a, \\ g_t = \rho_g g_{t-1} + \varepsilon_t^g, \\ \zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_t^\zeta, \\ \upsilon_t = \varepsilon_t^\upsilon, \end{cases} \tag{65}$$

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 $j = \{a, g, \zeta\}$.

3.3 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.[22] 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

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

present problems deriving from stochastic singularity.[23] 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. Section 3.4 provides the estimation for the baseline model. Section 3.5 evaluates our time-dependent model against alternative specifications.

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

3.3.1 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;[24] 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

[24]The GMM estimation for the Wage Phillips curve is showed in Chapter 2. For the hazard characterizing price adjustment we directly use as priors the GMM estimates of Sheedy (2007).

The sample range, the data and the TDWPC 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).

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.[25] 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.

3.3.2 Estimation results

Our baseline model consists of six equations, describing: the production function (56); the real marginal cost (57); the real wage dynamics (58); the marginal rate of substitution (59); the dynamic IS (63); the Taylor rule (64). Two additional equations close the model: the price and wage Phillips curve. In our baseline estimation we consider

[25] 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).

the time-dependent form for both price and wage equation, i.e. equations (54) and (61). Shocks dynamics are described by (65). Our estimations are reported in Table 2, 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.

Table 2 – Prior and posterior distributions[26]

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (Great Moderation)		
	Density	Mean	St. Dev.[27]	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

[26] 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. The scale factor of the covariance matrix has been set in order to achieve an acceptance rate of about 25%.

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

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.[28] 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.[29]

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

[29] 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.

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

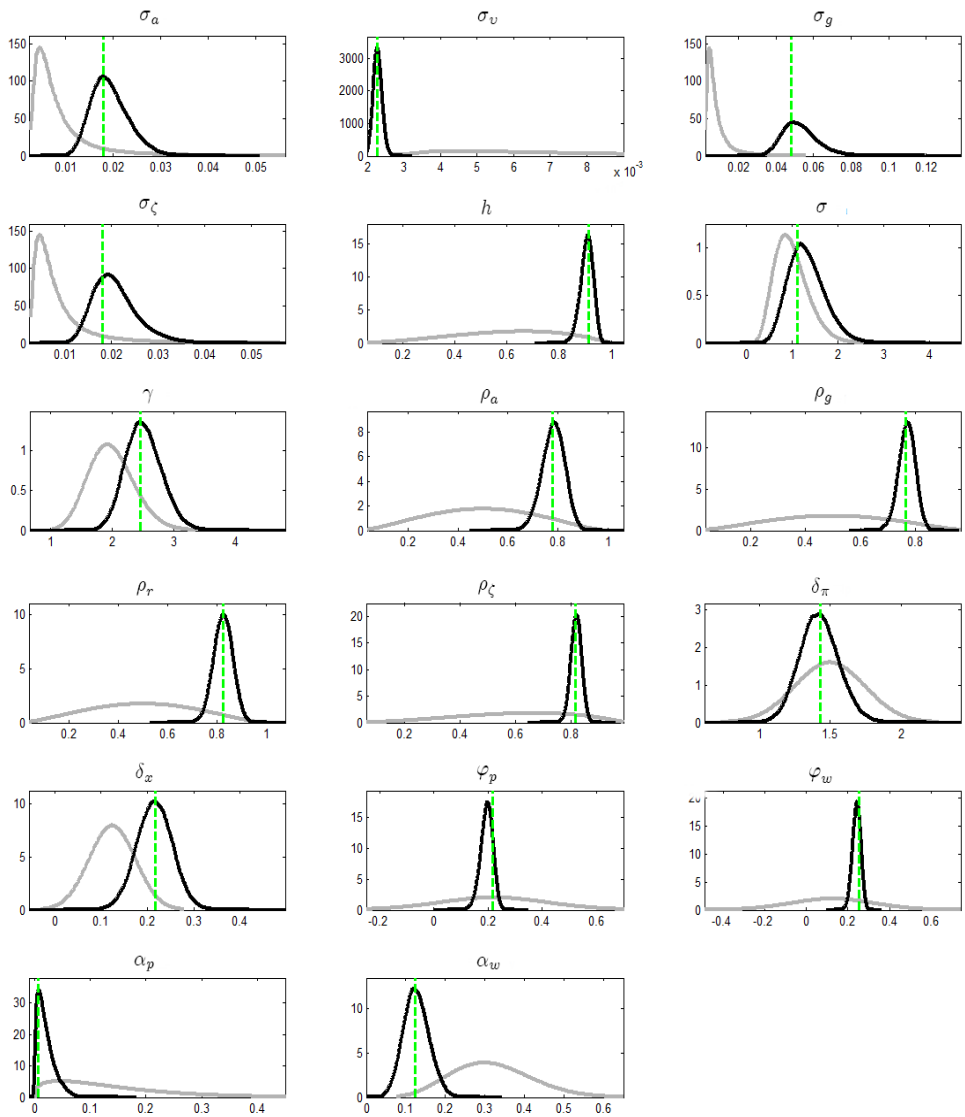

Figure 2 - Prior distribution (grey curve), Posterior distribution (green curve) and Posterior mode (dotted line) of the estimated parameters.

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 $[-1; 1]$,[30] whereas the prior distributions for the remaining parameters are the same used previously. Results are provided in Table 3. The "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.[31]

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

[31] These results are available in Appendix B.

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

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

3.4 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.

3.4.1 Alternative price-setting mechanisms: EHL with indexation

It is easy to verify that the price Phillips curve (54) nests the EHL case. Assuming $\varphi_p = 0$, we get:

$$\pi_t^p = \beta E_t \pi_{t+1}^p + \lambda_p \left(mc_t + \zeta_t \right) \tag{66}$$

where $\lambda_p = \frac{\alpha_p[1-\beta(1-\alpha_p)]}{1-\alpha_p}\eta_{cx}$. Equation (66) can be also augmented by indexation:

$$\pi_t^p = \frac{\iota_p}{(1+\iota_p\beta)}\pi_{t-1}^p + \frac{\beta}{1+\iota_p\beta}E_t\pi_{t+1}^p + \lambda_p^\iota \left(mc_t + \zeta_t \right) \tag{67}$$

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 (61) nests the EHL case for $\varphi_w = 0$:

$$\pi_t^w = \beta E_t \pi_{t+1}^w - \lambda_w \mu_t^w \tag{68}$$

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

$$\pi_t^w = \iota_w \pi_{t-1}^p - \iota_w \beta \pi_t^p + \beta E_t \pi_{t+1}^w - \lambda_w \mu_t^w \tag{69}$$

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

3.4.2 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:

$$\pi_t^p = \frac{\xi_p}{\Lambda_p}\pi_{t-1}^p + \frac{\beta\left(1-\alpha_p\right)}{\Lambda_p}E_t\pi_{t+1}^p + \lambda_p^{\xi}\left(mc_t + \zeta_t\right) \tag{70}$$

$$\pi_t^w = \frac{\xi_w}{\Lambda_w}\pi_{t-1}^p + \frac{\beta\left(1-\alpha_w\right)}{\Lambda_w}E_t\pi_{t+1}^w - \lambda_w^{\xi}\mu_t^w \tag{71}$$

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

3.4.3 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:[32]

1. EHL model with indexation (EHLind), by considering (67) and (69);

2. EHL model with indexation *á la* Galí-Gertler (GG), by considering (70) and (71).

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

[32] 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.

with the highest log-marginal likelihood better explains the data.[33] Table 4 reports our results.

Table 4 - Log-marginal data densities and Bayes factors for different models[34]

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,[35] 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

[33] 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).

[34] 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).

[35] 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."

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.[36] 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 3 shows the wage IRFs to a price mark-up shock based on estimated version of the three alternative specifications.

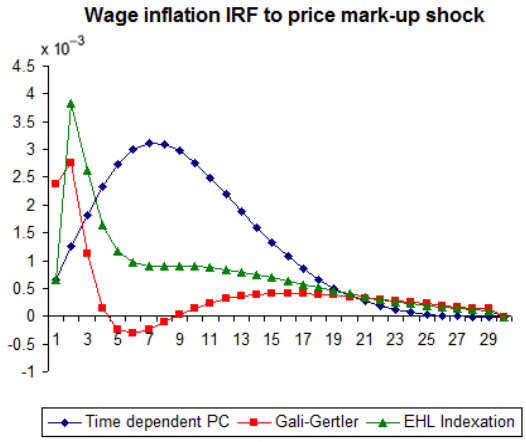

Figure 3 - IRFs of the wage inflation to 1% price mark-up shock for different model specifications.

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.

[36] The reduced form of the estimated Phillips curves considered herein are reported in Appendix D.

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.

3.5 International evidence

We have estimated our model also for a wider set of countries in order to investigate if time-dependent Phillips curves, based on upward hazard function, still emerge in the sample considered. The set we use is characterized by countries that have changed their monetary regime, concerning more on inflation stabilization, allowing us to analyze if the hazard parameters strictly remain positive after the regime switch. In particular we consider Australia and United Kingdom that, respectively since 1994:1 and 1992:4, adopt an inflation targeting regime; moreover, we have estimated the model for France, Germany and Italy that, since 1999:1, joined the European Monetary Union (EMU). In Table 5 we report our estimations for the hazard parameters, both for the full sample and the sub-sample aforementioned.[37]

Table 5 – Posterior mean estimation for different countries

	Australia		UK		France		Germany		Italy	
	full	inf. targ.	full	inf. targ.	full	EMU	full	EMU	full	EMU
α_p	0.05	0.09	0.05	0.07	0.07	0.10	0.04	0.08	0.04	0.07
φ_p	0.26	0.23	0.22	0.19	0.12	0.08	0.25	0.18	0.20	0.15
α_w	0.20	0.29	0.27	0.23	0.14	0.20	0.19	0.20	0.12	0.15
φ_w	0.24	0.16	0.13	0.17	0.16	0.07	0.23	0.18	0.27	0.20

Our estimations confirm the results found in Section 3.3: the hazard function is increasing and its slope, although slightly drops, remain strictly positive also after

[37] In Appendix C we provide details on the data, prior and sample considered for each country and show the results for complete set of estimated parameters.

a change in the way to conduct monetary policy. These results furtherly confirm our evidence that time-dependent adjustment is a valid approach to micro-found and capture inflation persistence.

3.6 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. This result is driven by the dynamics of the wage inflation that, under our TDWPC specification, is given by a hump-shaped response to a cost push shock.

Chapter 4

Welfare and optimal monetary policy

4.1 Introduction

The analysis of the optimal monetary policy implemented by the Central Bank is a central issue in macroeconomics. Since the seminal work of Kydland and Prescott (1977) many research papers has been devoted to investigate how a central banker should behaves in facing a shock, in order to maximize the social welfare. In its simpler specification, according to Clarida *et al.* (1999), we can consider a simple New Keynesian DSGE macroeconomic model described by two equations: an IS schedule, representing the demand side of the economy, and a New Keynesian Phillips Curve (NKPC henceforth) for the supply side. Formally, we have:

$$x_t = E_t x_{t+1} + \omega \left(i_t - E_t \pi_{t+1} - r_t^n \right) \tag{72}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + u_t \tag{73}$$

where $\omega < 0$, $\kappa > 0$ and u_t is a cost-push shock assumed to follow an $AR(1)$ process; all the variables are expressed in log-deviation from their steady state. Equation (72) is the IS curve, which relates the output gap, expressed as the log-deviation of the

output from its natural level,[38] with the expectations of future output gap level and real interest rate gap; the supply curve (73) describes the inflation dynamics as a process depending on the expectations of future inflation and on real activity, the latter measured by the output gap.

In order to derive the optimal policy rule followed by the monetary authority, we need to compute the Central Bank's loss function. By considering a second-order Taylor approximation of the households utility function, it is possible to derive a welfare function having a quadratic form:

$$W_t = \frac{1}{2}\left(\pi_t^2 + \phi y_t^2\right) \tag{74}$$

where ϕ is the relative weight attached to output gap stabilization. From (74), it is clear as the Central Bank has two objectives: stabilizing inflation and output gap. The policymaker will solve an optimization problem by minimizing the discounted sum of (74) subject to the constraint given by equation (73);[39] the Lagrangian for this policy problem can be written as:

$$\Im_t = \sum_{i=0}^{\infty} \beta^i E_t \left[\frac{1}{2}\left(\pi_{t+i}^2 + \phi x_{t+i}^2\right) + \lambda_{t+i}\left(\pi_{t+i} - \beta E_t \pi_{t+i+1} - \kappa x_{t+i} - u_{t+i}\right)\right] \tag{75}$$

Clarida *et al.* (1999) showed that the policy that minimize (75) is given by a commitment to a rule; nonetheless, this policy is not time-consistent. Woodford (1999, 2003) suggests to solve this problem by using a "timeless perspective", where the

[38] For natural level we mean the level of output arising when prices are flexible.

[39] As common practice in literature, we assume that the Central Bank can directly control inflation as an instrument: thus, we can neglect equation (72) in the optimization problem.

optimal policy rule, in each period, should be:

$$\pi_t = -\frac{\phi}{\kappa}\left(y_t - y_{t-1}\right) \tag{76}$$

A shortcoming of monetary models based on equations (72) and (73) is that they lack of empirical realism since inflation is a persistent process and this inertia has not negligible effects on the dynamic properties of the model. In particular, a purely forward-looking specification for the NKPC fails to account for the effects of persistence on the dynamics of the model. Considering a hybrid version of (73), has implications both for the welfare function form and, consequently, for the derivation of the optimal policy rule (see Steinsson, 2003). Moreover, introducing persistence has important effects for the evaluation of the cost of a disinflation.[40]

Several methods has been introduced in literature to embed the NKPC of a backward term: we borrow from Sheedy (2007, 2010) a log-linear-economy representation which departs from the standard Calvo mechanism assuming that a firm has a higher probability of changing its price the longer it has kept it unchanged. This price assumption leads to a hybrid time-dependent Phillips curve exhibiting intrinsic inertia *à la* Fuhrer and Moore (1995), as it embeds a backward-term for inflation. Formally, our pricing model implies a non-negative hazard function. It generalizes the Calvo price setting, which emerges as a particular case when the hazard is flat.

A growing number of papers has been recently devoted to analyze the effects of a persistent NKPC on optimal monetary policy. Steinsson (2003) considers a NKPC with rule of thumb *á la* Galí-Gertler (1999) and shows that, within this context, the

[40] Many papers study the cost of a disinflation under inflation persistence, see e.g. Fuhrer (1995) and Ascari and Ropele (2012).

Central Bank's loss function should include also lagged terms for inflation. Ascari and Ropele (2007), analyze the implications for the optimal monetary policy when the log-linearization of the system is taken around a positive steady state inflation: their findings are that the gains of a commitment are influenced by the level of trend-inflation and the degree of interest rate smoothing progressively increases as the assumed level of trend-inflation is raised.

We use 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 extend Sheedy (2007) by deriving a consistent second-order approximation of welfare. As said we follow Woodford's (2003) approach. We assume that 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 result to those of Steinsson (2003), who investigates an alternative model of intrinsic inflation persistence under this assumption.[41]

We use our welfare approximation to explore the nature of macro distortions induced

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

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 the costs associated with relative price dispersion and consumption fluctuations that appear in the standard New Keynesian model.

In the Calvo price setting, 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 spell. 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 allow us to 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 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 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. indexing 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 find that recessions (expansions) are more persistent, gradual policies prevail on immediate overshooting that characterizes the purely forward-looking case under commitment, and commitment vs. discretion gains fall.

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 word with persistence, commitment vs. discretion gains increases in inflation inertia generated by a more asymmetric distribution of reset probabilities among firms conditional to the time elapsed since the last spell.

The rest of the paper is organized as follows. Section 4.2 presents the economic setup considered and explains how to embed a New Keynesian Phillips Curve with intrinsic persistence by using positive hazard functions. Section 4.3 shows the derivation of the welfare function when inflation inertia is achieved via a reset price probability that is time-dependent. In Section 4.4 we explore the nature of macro distortions induced by our pricing model and its implications for the conduct of monetary policy. We also analyze how a change in the hazard slope affects the welfare gain of a commitment over discretion. A final section concludes.

4.2 The economy

The core of the New Keynesian approach is price stickiness; i.e., a situation in which the nominal price is resistant to change. The most common way to introduce it is the Calvo staggered contracts model,[42] where each firm faces a sort of lottery for resetting

[42] Common alternatives include Rotemberg (1982) price model, sticky information, state-dependent models. See Chapter 6 of Walsh's (2010) textbook for details.

price. In every period of time, each firm has a constant probability of being able to change its price. The alternative mechanism we consider here is instead a lottery where the probability of resetting the price, defined by a hazard function, is a function of time. Following Sheedy (2007, 2010), we assume that a firm has a higher probability of changing its price the longer it has kept it unchanged. Now, the firm faces a probability distribution over possible probabilities of resetting conditional to the time elapsed from the last adjustment.

Formally, in order to derive the Phillips curve, we first need to specify the hazard function, which is defined as the event rate at time h conditional on survival until time h or later, where the event is "be able to reset the price." The hazard function can be written as:

$$\alpha(h) = \alpha_p + \sum_{j=1}^{\min(h-1,n)} \varphi_p(j) \left\{ \prod_{k=h-j}^{h-1} [1 - \alpha(k)] \right\}^{-1}, \tag{77}$$

where $\alpha(h)$ is the probability to change a price which last reset was h periods ago; α_p is the initial value of the hazard, $\varphi_p(j)$ is its slope; n is the number of parameters that control the slope, for $n = 1$ the slope is governed by only one parameter, $\varphi_p(j) = \varphi_p$.

Knowing (77), the representative firm i chooses its price to maximize the profit, as it enjoys monopoly power on the goods market, constrained by the total demand for its product (i.e., $Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon_p} Y_t$, where $\frac{P_t(i)}{P_t}$ is the relative price, Y_t is the aggregate demand, and ε_p firm's price elasticity). We assume that representative firm has the following production function: $Y_t(i) = A_t N_t(i)^{1-\delta}$, where labor (N_t) is the only factor of production, A_t denotes the technology, and $\delta \in (0, 1)$. Formally, the representative

firm i solves

$$\max_{P_t^*(i)} \sum_{t=0}^{\infty} \varsigma_t E_0 \beta^t \left\{ P_t(i) \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon_p} Y_t - W_t \left[\frac{1}{A_t} \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon_p} Y_t \right]^{\frac{1}{1-\delta}} \right\} \tag{78}$$

where $\varsigma_l = \prod_{h=1}^{l} (1 - \alpha_h)$ with $\varsigma_0 = 1$ represents the probability that a price remains fixed for l periods (survival function) and β is the discount factor.

By assuming a competitive labor market, under very common assumptions in New Keynesian framework, the solution of the problem (78) leads to the following log-linear inflation equation:

$$\pi_t = \Psi_1 \pi_{t-1} + \beta \Psi_2 E_t \pi_{t+1} - \beta^2 \Psi_1 E_{t+1} \pi_{t+2} + \kappa x_t + u_t \tag{79}$$

where π_t is the inflation rate; is the x_t output gap; u_t is a cost push shock, modelled as an $AR(1)$ process. We always indicate deviations with lowercase letters. Coefficients in (79) are:

$$\begin{cases} \Psi_1 = \frac{\varphi_p}{(1-\alpha_p) - \varphi_p [1 - \beta(1-\alpha_p)]}, \\ \Psi_2 = 1 + (1-\beta)\, \Psi_1, \\ \kappa = \frac{(\alpha_p + \varphi_p)\left[1 - \beta(1-\alpha_p) + \beta^2 \varphi_p\right]}{(1-\alpha_p) - \varphi_p [1 - \beta(1-\alpha_p)]} \Xi_p, \end{cases}$$

with $\Xi_p = \left(\sigma + \frac{\gamma+\delta}{1-\delta}\right) \frac{1-\delta}{1-\delta+\delta\varepsilon_p}$, where σ, γ are parameters characterizing the representative consumer's utility function (see (80), below).

Equation (79) is our Phillips curve. Coefficient κ measures its slope and also depends on the hazard function parameters (α_p, φ_p). This Phillips curve captures the intrinsic component of inflation inertia, since it has a backward term. The negative coefficient attached to π_{t+2} must be interpreted as a reduction in the total degree of

its forwardness. Equation (79) nests the standard New Keynesian Phillips curve as a special case for $\varphi_p = 0$ (flat hazard).

4.3 Welfare approximation

The welfare loss derives from a second-order Taylor approximation of the households 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), Benigno and Woodford (2012). In the derivation we follow Galí (2008), 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:

$$U_t = \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\gamma}}{1+\gamma} \right) \tag{80}$$

where C_t is consumption; σ and γ are parameters.

A second-order approximation of (80) is:

$$U_t - U \simeq U_c Y \left(y_t + \frac{1-\sigma}{2} y_t^2 \right) + U_n N \left(n_t + \frac{1+\gamma}{2} n_t^2 \right) \tag{81}$$

where $\sigma = -\frac{U_{cc}}{U_c} Y$, $\gamma = \frac{U_{nn}}{U_n} N$. Note to obtain (81), 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:

$$\left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\delta}} \underbrace{\int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{\frac{-\varepsilon_p}{1-\delta}} di}_{D_t} = N_t \tag{82}$$

where D_t measures the degree of price dispersion.

Log-linearizing (82) around the steady state implies:

$$(1-\delta)\, n_t = y_t - a_t + \frac{\varepsilon_p}{2\Theta} var_i \left\{p_t(i)\right\} \tag{83}$$

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

As we consider an efficient steady state, thus $-\frac{U_n}{U_c}$ equals the marginal product of labor in the steady state, $(1-\delta)\left(\frac{Y}{N}\right)$, by substituting (83) into (81), after some simple algebraic manipulations, we obtain:

$$\frac{U_t - U}{U_c Y} \simeq -\frac{1}{2}\left[\frac{\varepsilon_p}{\Theta} var_i \left\{p_t(i)\right\} + \left(\sigma + \frac{\gamma+\delta}{1-\delta}\right) x_t^2\right] + t.i.p. \tag{84}$$

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:

$$W = -\frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\varepsilon_p}{\Theta} var_i \left\{p_t(i)\right\} + \left(\sigma + \frac{\gamma+\delta}{1-\delta}\right) x_t^2\right] \tag{85}$$

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

level evolves as:

$$\log p_t(i) = \sum_{h=0}^{\infty} \theta_h \log P_{t-h}^* \tag{86}$$

where P_t^* stands for the reset price and θ_h denotes the share of firms which last price 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 (86) and exploiting the properties of the variance, we can show that the discounted sum of price dispersion evolves in the following way:

$$\sum_{t=0}^{\infty} \beta^t var_i \{p_t(i)\} = \frac{\sum_{t=0}^{\infty} \beta^t \left(\frac{1-\alpha_p}{1+\varphi_p} \pi_t^2 + \alpha_p \varphi_p \pi_{t-1}^2 \right)}{\left[1 + \varphi_p - \beta (1 - \alpha_p)\right] \left(\alpha_p + \varphi_p\right)}, \tag{87}$$

the proof is provided in the Appendix E.

Once we have derived the price dispersion, we substitute (87) into (85) and obtain our welfare measure:

$$W = -\Omega \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + \Gamma_2 x_t^2 + \Gamma_1 \pi_{t-1}^2 \right] \tag{88}$$

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

$$\Omega = \frac{\varepsilon_p}{\Theta} \frac{1}{1 - \beta \left(\frac{1-\alpha_p}{1+\varphi_p} \right)} \frac{1 - \alpha_p}{\left(1 + \varphi_p\right)^2 \left(\alpha_p + \varphi_p\right)}, \tag{89}$$

$$\Gamma_1 = \left(\frac{1 + \varphi_p}{1 - \alpha_p} \right) \alpha_p \varphi_p, \tag{90}$$

$$\Gamma_2 = \frac{1}{\Omega} \left(\sigma + \frac{\gamma + \delta}{1 - \delta} \right). \tag{91}$$

As for the Phillips curve, our welfare measure (88) encompasses that deriving from

Calvo price setting for $\varphi_p = 0$: in such a case the weight attached to backward inflation drops to zero.

4.4 Welfare analysis and optimal policy

4.4.1 Price-model features 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, in this setup, firms are ex ante homogenous, facing the same probability of being extracted in the Calvo's lottery.

In our framework, we disentangle two sources of distortions stemming from the average and the distribution of the probability of resetting prices. Here, differently from Calvo (1983), reset probability is not constant. Firms are ex ante heterogeneous, as each faces a different probability of resetting according to the time elapsed since the last reset.[43] Thus, distortions are related to the average probability of changing prices (as in Calvo model), but now they also depend on how the reset probability is distributed among firms.

The average probability of resetting prices (ARP) can be written as:

$$ARP = \alpha_p + \varphi_p \tag{92}$$

A given average probability of changing prices can be obtained for different slopes of the hazard function (φ_p). For instance, we can obtain three different scenarios

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

consistent with $ARP = 0.3$ associated to three different distributions of the reset probability, as illustrated by the figure below.

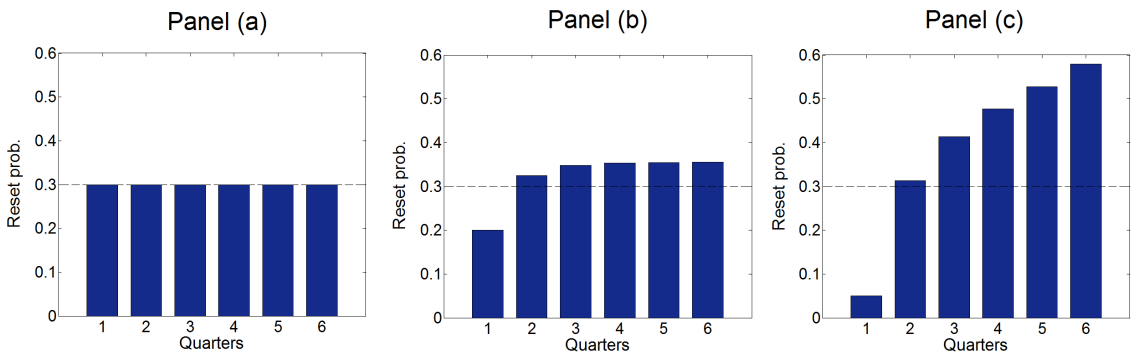

Figure 4 – Different distributions of the reset probability with the same mean.

All the panels of Figure 4 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 reset probability, i.e. there is more heterogeneity in the reset probability distribution. Similar figures can be drawn for different average probabilities. Given the average probability to adjust a price, we can refer to the scenario (a) as Calvo; scenario (b) as the mid-dispersion case; (c) as the high-dispersion case.

In order to study the welfare effects of these two features of the price setting process, we use numerical simulations based 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 by setting $\sigma = 1$ and $\gamma = 2$. The process governing the cost push shock

is modeled as an $AR(1)$ with a coefficient equal to 0.5. Our qualitative results are independent of the calibration.

Table 6 reports the welfare effects.[44] We consider different average probabilities and different dispersions. For the sake of comparison, we consider the same monetary policy rule in all cases.[45] We express the welfare effects in the more common form of welfare losses, which are further normalized to the Calvo scenario with a reset probability equal to 0.4.

Table 6 – 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

The table shows that a decrease in the average reset probability induces more distortions as in the traditional setup. Moreover, here distortions also increase in the dispersion of the reset probability, due to a higher hazard slope. As the hazard becomes steeper there is an increase of the variance of the output gap with respect to that of inflation.

The next subsection investigates how the 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.

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

[45] 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 holds if optimal policy are instead considered. More details are available upon request.

4.4.2 Optimal policies

The Government problem consists in maximizing (88) subject to (79), 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 of the following Lagrangian expression:

$$\min_{\{\pi_{t+j}\}_{j=0}^{+\infty},\ \{y_{t+j}\}_{j=0}^{+\infty}} - E_0 \sum_{j=0}^{\infty} \beta^j \left\{ \frac{1}{2} \left[\pi_{t+j}^2 + \Gamma_1 \pi_{t+j-1}^2 + \Gamma_2 x_{t+j}^2 \right] + \right. \tag{93}$$
$$\left. + \lambda_{t+j} \left[\pi_{t+j} - \Psi_1 \pi_{t+j-1} - \beta \Psi_2 \pi_{t+j+1} + \beta^2 \Psi_1 \pi_{t+j+2} - \kappa x_{t+j} - u_{t+j} \right] \right\}$$

where λ_{t+j} is the Lagrangian multiplier. Without loss of generality, we are considering a welfare loss minimization instead of a welfare maximization.

Corresponding first order conditions are:

$$(1 + \Gamma_1 \beta)\, \pi_{t+j} + \lambda_{t+j} - \Psi_1 \beta E_{t+j} \lambda_{t+j+1} = 0 \quad \text{for } j = 0 \tag{94}$$

$$(1 + \Gamma_1 \beta)\, \pi_{t+j} + \lambda_{t+j} - \Psi_1 \beta E_{t+j} \lambda_{t+j+1} - \Psi_2 \lambda_{t+j-1} + \Psi_1 \lambda_{t+j-2} = 0 \quad \text{for } j > 0 \tag{95}$$

$$\Gamma_2 x_{t+j} - \kappa \lambda_{t+j} = 0 \quad \text{for } j \geqslant 0 \tag{96}$$

As it is well known, the system (94)-(96) leads to a dynamic inconsistency. At $t = 0$ the policymaker implements (94) and commits to act following (95) in future periods. When $t = 1$ comes, it would be optimal to pursue again (94) rather than (95). Thus, time-inconsistency arises as it is no longer optimal act as planned at $t = 0$.

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. In this case the derivatives with respect to the terms in the Lagrangian expression (93) involving expectations are all equal to zero. The discretionary solution is thus obtained from (94) and (96). By combining these first-order conditions to eliminate λ_t, we obtain the following optimal policy rule:

$$\pi_t = -\frac{\Gamma_2}{\kappa(1+\Gamma_1\beta)}\left(x_t - \Psi_1\beta x_{t+1}\right) \tag{97}$$

The above expression says that inflation has to respond to the current output gap, but not to its past values, as in the textbook case; moreover, inflation also reacts to the discounted expected output gap (if $\varphi_p > 0$). So equation (97) generalizes the well known "lean against the wind" policy: If the hazard is not flat, expected output gap should be also considered since the Phillips curve has an inertial component.

Woodford (2003) proposes to overcome dynamic inconsistency by implementing an alternative regime known as "timeless perspective," where the Government should ignore (94). The idea is that optimal policy solves problem (93) "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 (95) and (96). Combining these equations, we eliminate λ_{t+j} and obtain:

$$\pi_{t+j} = -\frac{\Gamma_2}{\kappa(1+\Gamma_1\beta)}\left(x_{t+j} - \Psi_1\beta E_{t+j}x_{t+j+1} - \Psi_2 x_{t+j-1} + \Psi_1 x_{t+j-2}\right) \tag{98}$$

Again, as the Phillips curve has an inertial component, discounted expected output gap should be considered. Moreover, the timeless perspective is characterized by the fact that the Government, by reacting to the past output gap, affects 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 (79)). If $\varphi_p = 0$, inflation expectations at $t+2$ do not affect the trade off and $\Psi_1 x_{t+j-2}$ drops to zero.

The impulse response functions to a cost push shock in the two regimes are depicted in the Figure 5, where we consider three scenarios illustrated in Figure 4 (i.e., Calvo, mid-dispersion and high dispersion under a common ARP equal to 0.3).

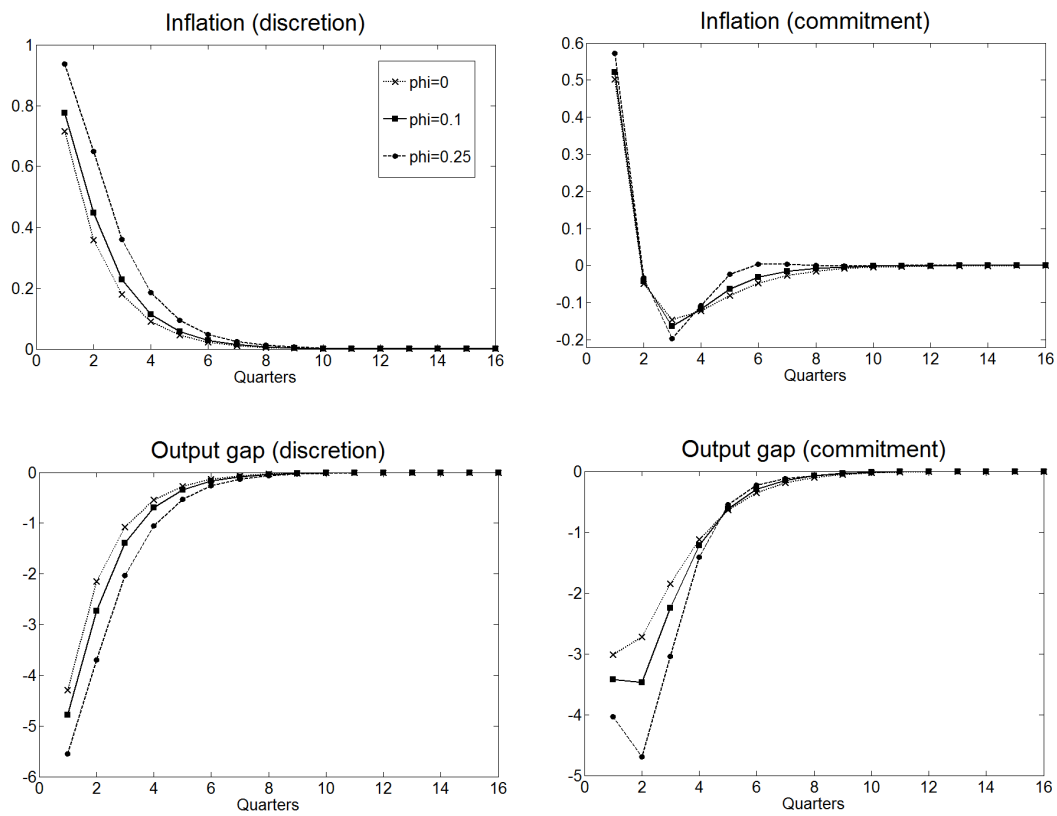

Figure 5 – Impulse response functions in the two policy regimes

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. However, this result does 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. In 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 ($\varphi_p = 0$) up to the high-dispersion scenario ($\varphi_p = 0.30$). 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 below reports the gains associated to the commitment under different degree of inflation inertia.

Table 7 – 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 results 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. 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.

4.5 Conclusions

In this chaprer, we derived a linear-quadratic model that is consistent with intrinsic inflation persistence. Inflation inertia is micro-founded assuming that the probability to update a price of each firm is function of the time, more specifically we assumed that firms' pricing decisions depend on the time spent from the last price reset. 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 authorities 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.

Finally, 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 increase 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.

Appendix A – Time-dependent Wage Phillips curve derivation

In what follows we show the complete derivation of the time-dependent Wage Phillips curve (44) for the general case $n = 1$.

A.1 Wage inflation dynamics

Let us divide equation (37) by W_t getting:

$$1 = \frac{1}{W_t}\left(\sum_{l=0}^{\infty}\theta_l W_{t-l}^{*\,1-\varepsilon_w}\right)^{\frac{1}{1-\varepsilon_w}} \tag{99}$$

Now we log-linearize[46] (36) and (99) around a steady-state with zero wage inflation $\overline{\pi}^w = 1$, $\overline{MRS} = \frac{\overline{W}}{\overline{P}\mu_w}$ and $\frac{\overline{W}^*}{\overline{W}} = \overline{\varpi} = 1$.

We begin from equation (36):

$$\overline{W}^* + \overline{W}^* w_t^* = \left[\frac{\mu_w\left(\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}\frac{\overline{W}}{\overline{P}\mu_w}\overline{P}\right)}{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}\right] + \frac{\overline{W}}{\overline{P}\mu_w}\sum_{\tau=t}^{\infty} mrs_{\tau|t}\frac{\mu_w\overline{P}\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}} +$$
$$+\frac{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}\overline{P}\mu_w\frac{\overline{W}}{\overline{P}\mu_w}}{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}p_\tau + \log\mu_w\frac{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}\frac{\overline{W}}{\overline{P}\mu_w}\mu_w\overline{P}}{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}$$

dividing by $\overline{W}^*$ we get

$$w_t^* = \frac{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}mrs_{\tau|t} + \frac{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty}\varsigma_{\tau-t}\beta^{\tau-t}}p_\tau + \log\mu_w$$

[46] Variables with the bar denote steady-state values. Small-caps letters denote the log-deviation from the steady-state.

Let $\log\mu_w = \mu$

$$w_t^* = \mu + \sum_{\tau=t}^{\infty} \left(\frac{\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t}\beta^{\tau-t}} \right) \left[mrs_{\tau|t} + p_\tau \right] \tag{100}$$

we know that $mrs_{\tau|t} = mrs_\tau + \gamma \left(n_{\tau|t} - n_\tau \right)$, hence:

$$mrs_{\tau|t} = mrs_\tau - \varepsilon_w \gamma \left(w_t^* - w_\tau \right)$$

As a consequence (100) becomes:

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t}\beta^{\tau-t}} \right) \left[mrs_\tau - \varepsilon_w \gamma w_t^* + \varepsilon_w \gamma w_\tau + p_\tau + \mu \right]$$

$$\varepsilon_w \gamma w_t^* + w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t}\beta^{\tau-t}} \right) \left[mrs_\tau + \varepsilon_w \gamma w_\tau + p_\tau + \mu \right]$$

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t}\beta^{\tau-t}} \right) \left[\frac{mrs_\tau + \varepsilon_w \gamma w_\tau + p_\tau + \mu}{1 + \varepsilon_w \gamma} \right]$$

Moreover, knowing that $mrs_\tau = w_\tau - p_\tau - \mu_\tau^w$ and assuming that μ_τ^w denotes the log-deviation of the wage mark-up from its steady-state, we get:

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t}\beta^{\tau-t}} \right) \left[\frac{w_\tau - p_\tau - \mu_\tau^w + \varepsilon_w \gamma w_\tau + p_\tau}{1 + \varepsilon_w \gamma} \right]$$

$$w_t^* = \sum_{\tau=t}^{\infty} \left(\frac{\varsigma_{\tau-t}\beta^{\tau-t}}{\sum_{\tau=t}^{\infty} \varsigma_{\tau-t}\beta^{\tau-t}} \right) \frac{1 + \varepsilon_w \gamma}{1 + \varepsilon_w \gamma} w_\tau - \frac{1}{1 + \varepsilon_w \gamma} \mu_\tau^w$$

Let $t + i = \tau$, then:

$$w_t^* = \sum_{i=0}^{\infty} \left(\frac{\beta^i \varsigma_i}{\sum_{j=0}^{\infty} \beta^j \varsigma_j} \right) \left[w_{t+i} - \mu_{t+i}^w \right] \tag{101}$$

We now switch to equation (99), it can be rewritten as follows:

$$1 = \theta_0 \varpi_t^{1-\varepsilon_w} + \theta_1 \left(\frac{W_{t-1}^*}{W_t}\right)^{1-\varepsilon_w} + \theta_2 \left(\frac{W_{t-2}^*}{W_t}\right)^{1-\varepsilon_w} + \theta_3 \left(\frac{W_{t-3}^*}{W_t}\right)^{1-\varepsilon_w} + \dots$$

$$1 = \theta_0 \varpi_t^{1-\varepsilon_w} + \theta_1 \left(\frac{W_{t-1}^*}{W_t}\right)^{1-\varepsilon_w} \frac{W_{t-1}^{1-\varepsilon_w}}{W_{t-1}^{1-\varepsilon_w}} + \theta_2 \left(\frac{W_{t-2}^*}{W_t}\right)^{1-\varepsilon_w} \frac{W_{t-2}^{1-\varepsilon_w}}{W_{t-2}^{1-\varepsilon_w}} + \theta_3 \left(\frac{W_{t-3}^*}{W_t}\right)^{1-\varepsilon_w} \frac{W_{t-3}^{1-\varepsilon_w}}{W_{t-3}^{1-\varepsilon_w}} + \dots$$

$$1 = \theta_0 \varpi_t^{1-\varepsilon_w} + \theta_1 \left(\frac{W_{t-1}^*}{W_t}\right)^{1-\varepsilon_w} \frac{W_{t-1}^{1-\varepsilon_w}}{W_{t-1}^{1-\varepsilon_w}} + \theta_2 \left(\frac{W_{t-2}^*}{W_t}\right)^{1-\varepsilon_w} \frac{W_{t-2}^{1-\varepsilon_w}}{W_{t-2}^{1-\varepsilon_w}} + \theta_3 \left(\frac{W_{t-3}^*}{W_t}\right)^{1-\varepsilon_w} \frac{W_{t-3}^{1-\varepsilon_w}}{W_{t-3}^{1-\varepsilon_w}} + \dots$$

$$1 = \theta_0 \varpi_t^{1-\varepsilon_w} + \theta_1 \varpi_{t-1}^{1-\varepsilon_w} \pi_t^{w^{\varepsilon_w - 1}} + \theta_2 \varpi_{t-2}^{1-\varepsilon_w} \pi_t^{w^{\varepsilon_w - 1}} \pi_{t-1}^{w^{\varepsilon_w - 1}} + \dots$$

$$1 = \sum_{l=0}^{\infty} \theta_l \varpi_{t-l}^{1-\varepsilon_w} \left(\prod_{h=0}^{l-1} \pi_{t-h}^{w^{\varepsilon_w - 1}}\right) \tag{102}$$

Equation (102) can be furtherly rewritten as:

$$1 = \sum_{i=0}^{\infty} \theta_l \left(\frac{W_{t-l}^*}{W_{t-l}}\right)^{1-\varepsilon_w} \left(\prod_{h=0}^{l-1} \pi_{t-h}^{w^{\varepsilon - 1}}\right)$$

$$1 = \theta_0 \left(\frac{W_t^*}{W_t}\right)^{1-\varepsilon_w} + \theta_1 \left(\frac{W_{t-1}^*}{W_{t-1}}\right)^{1-\varepsilon_w} \left(\frac{W_t}{W_{t-1}}\right)^{\varepsilon_w - 1} + \theta_2 \left(\frac{W_{t-2}^*}{W_{t-2}}\right)^{1-\varepsilon_w} \left(\frac{W_t}{W_{t-1}} \frac{W_{t-1}}{W_{t-2}}\right)^{\varepsilon_w - 1} + \dots$$

$$1 = \theta_0 \left(\frac{W_t^*}{W_t}\right)^{1-\varepsilon_w} + \theta_1 \left(\frac{W_{t-1}^*}{W_t}\right)^{1-\varepsilon_w} + \theta_2 \left(\frac{W_{t-2}^*}{W_t}\right)^{1-\varepsilon_w} + \theta_3 \left(\frac{W_{t-3}^*}{W_t}\right)^{1-\varepsilon_w} + \dots$$

Log-linearizing around the deterministic steady-state given by $\overline{\pi}^w = 1$, $\overline{\varpi} = \frac{\overline{W}^*}{\overline{W}} = 1$ we get:

$$1 = \sum_{l=0}^{\infty} \theta_l + \theta_0(1-\varepsilon_w)\frac{1}{\overline{W}}\overline{W}^* w_t^* + \theta_1(1-\varepsilon_w)\frac{1}{\overline{W}}\overline{W}^* w_{t-1}^* + \theta_2(1-\varepsilon_w)\frac{1}{\overline{W}}\overline{W}^* \hat{W}_{t-2}^* + \ldots +$$
$$-\theta_0(1-\varepsilon_w)\frac{1}{\overline{W}^2}\overline{W}^*\overline{W} w_t - \theta_1(1-\varepsilon_w)\frac{1}{\overline{W}^2}\overline{W}^*\overline{W} w_t - \theta_2(1-\varepsilon_w)\frac{1}{\overline{W}^2}\frac{1}{\overline{W}^2}\overline{W}^*\overline{W} w_t - \ldots$$

since $1 = \sum_{l=0}^{\infty} \theta_l$, then:

$$\begin{aligned}(1-\varepsilon_w) w_t \sum_{l=0}^{\infty} \theta_l &= (1-\varepsilon_w) \sum_{l=0}^{\infty} \theta_l w_{t-l}^* \\ w_t &= \sum_{l=0}^{\infty} \theta_l w_{t-l}^* \end{aligned} \tag{103}$$

A.2 Wage Phillips curve

By exploiting the recursive parametrization of the hazard and survival functions, equations (101) and (103) can be rewritten in alternative way as follows:

$$w_t^* = \sum_{i=0}^{\infty} \left(\frac{\beta^i \varsigma_i}{\sum_{j=0}^{\infty} \beta^j \varsigma_j} \right) \left[w_{t+i} - \frac{1}{1+\varepsilon_w \gamma} \mu_{t+i}^w \right]$$

we call $\frac{1}{1+\varepsilon_w \gamma} = \Xi_w$

$$w_t^* = \frac{w_t - \lambda \mu_t^w}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\beta\varsigma_1\left[w_{t+1} - \Xi_w \mu_{t+1}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\beta^2\varsigma_2\left[w_{t+2} - \Xi_w \mu_{t+2}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots$$

by using equation (25):

$$w_t^* = \frac{w_t - \Xi_w \mu_t^w}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\beta(1-\alpha)\left[w_{t+1}-\Xi_w\mu_{t+1}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\left[\beta^2(1-\alpha)^2-\beta^2\varphi_1\right]\left[w_{t+2}-\Xi_w\mu_{t+2}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^3(1-\alpha)^3-2\beta^3\varphi_1(1-\alpha)-\beta^3\varphi_2\right]\left[w_{t+3}-\Xi_w\mu_{t+3}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^4(1-\alpha)^4-2\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_2(1-\alpha)+\beta^4\varphi_1^2-\beta^4\varphi_3-\beta^4\varphi_2(1-\alpha)\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots$$

$$w_{t+1}^* = \frac{w_{t+1} - \Xi_w \mu_{t+1}^w}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\beta(1-\alpha)\left[w_{t+2}-\Xi_w\mu_{t+2}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\left[\beta^2(1-\alpha)^2-\beta^2\varphi_1\right]\left[w_{t+3}-\Xi_w\mu_{t+3}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^3(1-\alpha)^3-2\beta^3\varphi_1(1-\alpha)-\beta^3\varphi_2\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^4(1-\alpha)^4-2\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_2(1-\alpha)+\beta^4\varphi_1^2-\beta^4\varphi_3-\beta^4\varphi_2(1-\alpha)\right]\left[\hat{W}_{t+5}-\Xi_w\hat{\mu}_{t+5}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots$$

$$w_{t+2}^* = \frac{w_{t+2} - \Xi_w \mu_{t+2}^w}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\beta(1-\alpha)\left[w_{t+3}-\Xi_w\mu_{t+3}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\left[\beta^2(1-\alpha)^2-\beta^2\varphi_1\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^3(1-\alpha)^3-2\beta^3\varphi_1(1-\alpha)-\beta^3\varphi_2\right]\left[w_{t+5}-\Xi_w\mu_{t+5}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^4(1-\alpha)^4-2\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_2(1-\alpha)+\beta^4\varphi_1^2-\beta^4\varphi_3-\beta^4\varphi_2(1-\alpha)\right]\left[w_{t+6}-\Xi_w\mu_{t+6}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots$$

$$w_{t+3}^* = \frac{w_{t+3} - \Xi_w \mu_{t+3}^w}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\beta(1-\alpha)\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \frac{\left[\beta^2(1-\alpha)^2-\beta^2\varphi_1\right]\left[w_{t+5}-\Xi_w\mu_{t+5}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^3(1-\alpha)^3-2\beta^3\varphi_1(1-\alpha)-\beta^3\varphi_2\right]\left[w_{t+6}-\Xi_w\mu_{t+6}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} +$$
$$+\frac{\left[\beta^4(1-\alpha)^4-2\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_2(1-\alpha)+\beta^4\varphi_1^2-\beta^4\varphi_3-\beta^4\varphi_2(1-\alpha)\right]\left[w_{t+7}-\Xi_w\mu_{t+7}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots$$

we now group in the following way:

$$
\begin{aligned}
w_t^* = & \tfrac{w_t - \Xi_w \mu_t^w}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \\
& + \tfrac{\beta(1-\alpha)\left[w_{t+1}-\Xi_w\mu_{t+1}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \tfrac{\left[\beta^2(1-\alpha)^2\right]\left[w_{t+2}-\Xi_w\mu_{t+2}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \\
& + \tfrac{\left[\beta^3(1-\alpha)^3-\beta^3\varphi_1(1-\alpha)\right]\left[w_{t+3}-\Xi_w\mu_{t+3}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \tfrac{\left[\beta^4(1-\alpha)^4-2\beta^4\varphi_1(1-\alpha)^2-\beta^4\varphi_2(1-\alpha)\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots \\
& + \tfrac{\left[-\beta^2\varphi_1\right]\left[w_{t+2}-\Xi_w\mu_{t+2}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \tfrac{\left[-\beta^3\varphi_1(1-\alpha)\right]\left[w_{t+3}-\Xi_w\mu_{t+3}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \tfrac{\left[-\beta^4\varphi_1(1-\alpha)^2+\beta^4\varphi_1^2\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots \\
& + \tfrac{\left[-\beta^3\varphi_2\right]\left[w_{t+3}-\Xi_w\mu_{t+3}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \tfrac{\left[\beta^4\varphi_2(1-\alpha)\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots \\
& + \tfrac{\left[-\beta^4\varphi_3\right]\left[w_{t+4}-\Xi_w\mu_{t+4}^w\right]}{1+\beta\varsigma_1+\beta^2\varsigma_2+\beta^3\varsigma_3+\ldots} + \ldots
\end{aligned}
$$

As we can see, 2nd and 3rd row of the latter equation are equal to $\beta(1-\alpha)E_t w_{t+1}^*$; 4th row is equal to $(-\beta^2\varphi_1 E_t w_{t+2}^*)$; 5th row is equal to $(-\beta^3\varphi_2 E_t w_{t+3}^*)$ and so on. Thus, we can rewrite this equation as:

$$
w_t^* = \beta(1-\alpha)E_t w_{t+1}^* - \sum_{l=1}^{n} \beta^{l+1}\varphi_l E_t w_{t+l+1}^* + \left[1-\beta(1-\alpha)+\sum_{l=1}^{n}\beta^{l+1}\varphi_l\right](w_t - \Xi_w\mu_t^w) \tag{104}
$$

Recursive parametrization of the hazard function implies a recursion for the stationary distribution of the duration of wage stickiness $\{\theta_l\}_{l=0}^{\infty}$:

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

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

Therefore, (103) can be rewritten in recursive form:

$$
w_t = \underbrace{\theta_0 w_t^*}_{\left(\alpha+\sum_{h=1}^{n}\varphi_h\right)\hat{W}_t^*} + \theta_1 w_{t-1}^* + \theta_2 w_{t-2}^* + \ldots
$$

$$w_t = \left(\alpha + \sum_{h=1}^{n} \varphi_h\right) w_t^* + \left[(1-\alpha)\,\theta_0\right] w_{t-1}^* + \left[(1-\alpha)\,\theta_1 - \varphi_1\theta_0\right] w_{t-2}^* +$$
$$+ \left[(1-\alpha)\,\theta_2 - \varphi_1\theta_1 - \varphi_2\theta_0\right] w_{t-3}^* + \ldots$$

$$w_t = \left(\alpha + \sum_{h=1}^{n} \varphi_h\right) w_t^* + \underbrace{(1-\alpha)\sum_{l=0}^{\infty} \theta_l w_{t-l-1}^*}_{(1-\alpha)w_{t-1}} - \varphi_1\theta_0 w_{t-2}^* - \varphi_1\theta_1 w_{t-3}^* - \ldots$$

$$w_t = \left(\alpha + \sum_{h=1}^{n} \varphi_h\right) w_t^* + (1-\alpha)\, w_{t-1} +$$
$$\underbrace{-\varphi_1\theta_0 w_{t-2}^* - \varphi_1\theta_1 w_{t-3}^* - \varphi_1\theta_2 w_{t-4}^* - \ldots - \varphi_2\theta_0 w_{t-3}^* - \varphi_2\theta_1 w_{t-4}^* - \ldots}_{-\sum_{l=1}^{n}\varphi_l\left(\sum_{l=0}^{\infty}\theta_l w_{t-1-l-h}^*\right)\Rightarrow w_{t-1-h}}$$

$$w_t = (1-\alpha)\, w_{t-1} - \sum_{l=1}^{n} \varphi_l w_{t-1-l} + \left(\alpha + \sum_{h=1}^{n} \varphi_h\right) w_t^* \tag{105}$$

Finally combining equation (104) and (105) we get an expression for a hybrid wage Phillips curve[47]:

$$w_t^* = \beta(1-\alpha)E_t w_{t+1}^* - \beta^2\varphi_1 E_t w_{t+2}^* + \left[1 - \beta\,(1-\alpha) + \beta^2\varphi_1\right](w_t - \Xi_w \mu_t^w)$$

$$w_t = (1-\alpha)\, w_{t-1} - \varphi_1 w_{t-2} + (\alpha + \varphi_1)\, w_t^* \Rightarrow w_t^* = \frac{w_t - (1-\alpha)\, w_{t-1} + \varphi_1 w_{t-2}}{\alpha + \varphi_1}$$

$$w_t = (1-\alpha)\, w_{t-1} - \varphi_1 w_{t-2} + (\alpha + \varphi_1)\left[\beta\,(1-\alpha)\right] E_t w_{t+1}^* - (\alpha + \varphi_1)\,\beta^2\varphi_1 E_t w_{t+2}^* +$$
$$+ (\alpha + \varphi_1)\left[1 - \beta\,(1-\alpha) + \beta^2\varphi_1\right] w_t - (\alpha + \varphi_1)\left[1 - \beta\,(1-\alpha) + \beta^2\varphi_1\right]\Xi_w \mu_t^w$$

[47] For the sake of simplicity we use $n = 1$, where n is the number of parameters which control the slope of the hazard function.

$$w_t = (1-\alpha)\, w_{t-1} - \varphi_1 w_{t-2} + (\alpha+\varphi_1)\left[\beta(1-\alpha)\right]\left[\frac{w_{t+1}-(1-\alpha)w_t+\varphi_1 w_{t-1}}{\alpha+\varphi_1}\right] -$$
$$-(\alpha+\varphi_1)\,\beta^2\varphi_1\left[\frac{w_{t+2}-(1-\alpha)w_{t+1}+\varphi_1 w_t}{\alpha+\varphi_1}\right] + (\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right] w_t -$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w \mu_t^w$$

$$w_t = (1-\alpha)\, w_{t-1} - \varphi_1 w_{t-2} + \left[\beta(1-\alpha)\right] w_{t+1} - \beta(1-\alpha)^2 w_t + \varphi_1\beta(1-\alpha)\, w_{t-1} - \beta^2\varphi_1 w_{t+2} +$$
$$+\beta^2\varphi_1(1-\alpha)\, w_{t+1} - \beta^2\varphi_1^2 w_t + (\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right] w_t +$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w \mu_t^w$$

$$w_t = -\beta(1-\alpha)^2 w_t - \beta^2\varphi_1^2 w_t + (\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right] w_t + (1-\alpha)\, w_{t-1} +$$
$$+\varphi_1\beta(1-\alpha)\, w_{t-1} - \varphi_1 w_{t-2} + \left[\beta(1-\alpha)\right] w_{t+1} + \beta^2\varphi_1(1-\alpha)\, w_{t+1} - \beta^2\varphi_1 w_{t+2} +$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w \mu_t^w$$

$$w_t = -\beta(1-\alpha)^2 w_t - \beta^2\varphi_1^2 w_t + \alpha\left[1-\beta(1-\alpha)\right] w_t + \alpha\beta^2\varphi_1 w_t + \varphi_1\left[1-\beta(1-\alpha)\right] w_t + \beta^2\varphi_1^2 w_t +$$
$$+(1-\alpha)\, w_{t-1} + \varphi_1\beta(1-\alpha)\, w_{t-1} - \varphi_1 w_{t-2} + \left[\beta(1-\alpha)\right] w_{t+1} + \beta^2\varphi_1(1-\alpha)\, w_{t+1} - \beta^2\varphi_1 w_{t+2} -$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w \mu_t^w$$

$$w_t - \varphi_1\left[1-\beta(1-\alpha)\right] w_t - (1-\alpha)\, w_{t-1} - \varphi_1\beta(1-\alpha)\, w_{t-1} - \alpha w_t = -\beta(1-\alpha)^2 w_t +$$
$$-\alpha\beta(1-\alpha)\, w_t + \alpha\beta^2\varphi_1 w_t - \varphi_1 w_{t-2} + \left[\beta(1-\alpha)\right] w_{t+1} + \beta^2\varphi_1(1-\alpha)\, w_{t+1} - \beta^2\varphi_1 w_{t+2} +$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w \mu_t^w$$

adding and subtracting $\varphi_1 w_{t-1}$:

$$\{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]\}w_t-(1-\alpha)w_{t-1}-\varphi_1\beta(1-\alpha)w_{t-1}+\varphi_1 w_{t-1}=-\beta(1-\alpha)^2 w_t-$$
$$-\alpha\beta(1-\alpha)w_t+\alpha\beta^2\varphi_1 w_t-\varphi_1 w_{t-2}+\varphi_1 w_{t-1}+[\beta(1-\alpha)]w_{t+1}+\beta^2\varphi_1(1-\alpha)w_{t+1}-$$
$$-\beta^2\varphi_1 w_{t+2}-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w\mu_t^w$$

$$\{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]\}w_t-\{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]\}w_{t-1}=\varphi_1(w_{t-1}-w_{t-2})-$$
$$-\beta(1-\alpha)^2 w_t-\alpha\beta(1-\alpha)w_t+\alpha\beta^2\varphi_1 w_t+[\beta(1-\alpha)]w_{t+1}+\beta^2\varphi_1(1-\alpha)w_{t+1}-\beta^2\varphi_1 w_{t+2}-$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w\mu_t^w$$

$$\{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]\}\pi_t^w=\varphi_1\pi_{t-1}^w+[\beta(1-\alpha)]w_{t+1}+\beta^2\varphi_1(1-\alpha)w_{t+1}-\beta^2\varphi_1 w_{t+2}+$$
$$-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w\mu_t^w-\underbrace{\beta(1-\alpha)^2 w_t-\alpha\beta(1-\alpha)w_t+\alpha\beta^2\varphi_1 w_t}_{\left(\beta-\alpha\beta-\beta^2\varphi_1\alpha\right)(-w_t)}$$

$$\{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]\}\pi_t^w=\varphi_1\pi_{t-1}^w+\beta w_{t+1}-\alpha\beta w_{t+1}-\beta^2\varphi_1\alpha w_{t+1}+$$
$$-\left(\beta-\alpha\beta-\beta^2\varphi_1\alpha\right)w_t-(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]\Xi_w\mu_t^w+$$
$$-\underbrace{\beta^2\varphi_1 w_{t+2}+\beta^2\varphi_1 w_{t+1}}_{-\beta^2\varphi_1\pi_{t+2}^w}$$

Finally we obtain:

$$\pi_t^w=\frac{\varphi_1}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]}\pi_{t-1}^w+\frac{\beta-\alpha\beta-\beta^2\varphi_1\alpha}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]}E_t\pi_{t+1}^w-$$
$$-\frac{\beta^2\varphi_1}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]}\pi_{t+2}^w-\frac{(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]}\Xi_w\mu_t^w$$

Let $\psi=\frac{\varphi_1}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]}$; $\kappa_w=\frac{(\alpha+\varphi_1)\left[1-\beta(1-\alpha)+\beta^2\varphi_1\right]}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]}\Xi_w$.

Since $\frac{\beta-\alpha\beta-\beta^2\varphi_1\alpha}{(1-\alpha)-\varphi_1[1-\beta(1-\alpha)]} = \beta\left[1+(1-\beta)\psi\right]$, then:

$$\pi_t^w = \psi\pi_{t-1}^w + \beta\left[1+(1-\beta)\psi\right]E_t\pi_{t+1}^w - \beta^2\psi E_t\pi_{t+2}^w - \kappa_w\mu_t^w \tag{106}$$

If we use a number of parameters equal to n the wage Phillips curve becomes:

$$\pi_t^w = \sum_{l=1}^{n}\psi_l\pi_{t-l}^w + \sum_{l=1}^{n+1}\delta_l E_t\pi_{t+l}^w - \kappa_w\mu_t^w \tag{107}$$

where:

$$\psi_l = \frac{\varphi_l + \sum_{h=l+1}^{n}\varphi_h\left[1-\beta(1-\alpha)+\sum_{k=1}^{h-1}\beta^{k+1}\varphi_k\right]}{(1-\alpha)-\sum_{h=1}^{n}\varphi_h\left[1-\beta(1-\alpha)+\sum_{k=1}^{h-1}\beta^{k+1}\varphi_k\right]}$$

$$\delta_1 = \beta\frac{(1-\alpha)-\sum_{h=1}^{n}\beta^h\varphi_h\left(\alpha+\sum_{k=1}^{h-1}\varphi_k\right)}{(1-\alpha)-\sum_{h=1}^{n}\varphi_h\left[1-\beta(1-\alpha)+\sum_{k=1}^{h-1}\beta^{k+1}\varphi_k\right]}$$

$$\delta_{l+1} = -\beta^{l+1}\frac{\varphi_l+\sum_{h=l+1}^{n}\beta^{h-1}\varphi_h\left(\alpha+\sum_{k=1}^{h-1}\varphi_k\right)}{(1-\alpha)-\sum_{h=1}^{n}\varphi_h\left[1-\beta(1-\alpha)+\sum_{k=1}^{h-1}\beta^{k+1}\varphi_k\right]}$$

$$\kappa_w = \frac{\Xi_w\left(\alpha+\sum_{h=1}^{n}\varphi_h\right)\left[1-\beta(1-\alpha)+\sum_{h=1}^{n}\beta^{h+1}\varphi_h\right]}{(1-\alpha)-\sum_{h=1}^{n}\varphi_h\left[1-\beta(1-\alpha)+\sum_{k=1}^{h-1}\beta^{k+1}\varphi_k\right]}$$

Appendix B – Robustness test

In this appendix we provide many robustness test for the estimation showed in Section 3.3. We test if our results, both for the hazard coefficients and the model comparison, are robust to change in the time series used. In particular, we use several measures for the nominal wage, GDP deflator and different filtering method. Moreover, we estimate two alternative specifications of our baseline model, by considering in one case an IS curve with no persistence, i.e. $h = 0$, in the other case a Taylor rule characterized by an $AR(2)$ interest rate. In what follows we consider the same parameters calibrations and priors used along Section 3.3.

B.1 Alternative time series

For the nominal wage we consider three alternative measures: *Average hourly earnings of production and nonsupervisory employees*, *Compensation per hour (Business sector)*, Hourly earnings (manufacturing sector). For the GDP deflator we consider the *Implicit price deflator for the nonfarm business sector*. Time series for both output and interest rate are the same used in Section 3.3. As usual the real wage is obtained by dividing nominal wage by the GDP deflator. Thus, we are able to build seven different samples as listed below:

1. SMPL1: nominal wage=average hourly earnings; deflator=GDP implicit price deflator (1967:1 - 2008:4).

2. SMPL2: nominal wage=compensation per hour business sector; deflator=GDP implicit price deflator (1960:1 - 2008:4).

3. SMPL3: nominal wage=hourly earnings manufacturing; deflator=GDP implicit price deflator (1960:1 - 2008:4).

4. SMPL4: nominal wage=compensation per hour nonfarm business sector; deflator=implicit price deflator nonfarm business sector (1960:1 - 2008:4).

5. SMPL5: nominal wage=average hourly earnings; deflator=implicit price deflator nonfarm business sector (1967:1 - 2008:4).

6. SMPL6: nominal wage=compensation per hour business sector; deflator=implicit price deflator nonfarm business sector (1960:1 - 2008:4).

7. SMPL7: nominal wage=hourly earnings manufacturing; deflator=implicit price deflator nonfarm business sector (1960:1 - 2008:4).

In Table 8 we report the Bayesian estimation of the hazard parameter for each sample considered both for the entire period (full) 1960:1 - 2008:4 and the Great Moderation (GM) 1982:1 - 2008:4.

Table 8 – Posterior mean estimation for different samples

	SMPL1		SMPL2		SMPL3		SMPL4		SMPL5		SMPL6		SMPL7	
	full	GM	full	GM	full	GM	full	GM	full	GM	full	GM	full	GM
α_p	0.01	0.04	0.02	0.06	0.02	0.05	0.01	0.04	0.01	0.04	0.01	0.05	0.01	0.04
φ_p	0.17	0.11	0.19	0.13	0.17	0.12	0.23	0.19	0.21	0.12	0.22	0.19	0.20	0.15
α_w	0.16	0.14	0.12	0.15	0.14	0.15	0.15	0.17	0.20	0.16	0.14	0.17	0.17	0.16
φ_w	0.14	0.10	0.25	0.25	0.17	0.16	0.23	0.24	0.13	0.09	0.23	0.24	0.15	0.15

Table 8 shows that our results are invariant to the time series used and that positive hazard function still holds also after the Great Moderation. In Table 9 we show the

results of the model comparison based on the Bayes factors, for each sample considered. For all the samples our model clearly outperforms the alternatives considered.

Table 9 - Log-marginal data densities (ML) and Bayes factors (BF)

		BASE	BASE(non-info)	EHL_IND	GG
DB1:					
	ML	3182.1	3179.0	3153.7	3151.6
	BF		exp [−3.1]	exp [−28.3]	exp [−30.4]
DB2:					
	ML	3609.0	3605.9	3560.0	3554.6
	BF		exp [−3.1]	exp [−49.0]	exp [−54.4]
DB3:					
	ML	3702.9	3699.9	3668.4	3663.8
	BF		exp [−3.0]	exp [−34.5]	exp [−39.1]
DB4:					
	ML	3503.2	3500.6	3465.0	3461.4
	BF		exp [−2.6]	exp [−38.2]	exp [−41.8]
DB5:					
	ML	3083.8	3081.7	3059.0	3056.7
	BF		exp [−2.1]	exp [−24.8]	exp [−27.1]
DB6:					
	ML	3496.2	3493.9	3456.8	3452.3
	BF		exp [−2.3]	exp [−39.4]	exp [−43.9]
DB7:					
	ML	3578.8	3576.5	3549.3	3545.8
	BF		exp [−2.3]	exp [−29.5]	exp [−33.0]

B.2 Alternative data filtering

We further test if our result are robust to an alternative method to extract the cyclical component. In Table 10 we show our estimations for data detrended by using the full-sample asymmetric filter proposed by Christiano and Fitzgerald (2003). Parameters calibrations and priors are the same used in Section 3.3.

Table 10 – Prior and posterior distributions

	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	1.289	0.743	1.937	1.371	0.584	2.075
γ	Gamma	2.0	0.375	2.523	1.980	2.964	2.278	1.693	2.734
h	Beta	0.6	0.2	0.911	0.868	0.948	0.906	0.861	0.951
δ_π	Normal	1.5	0.25	1.417	1.206	1.625	1.884	1.532	2.206
δ_x	Normal	0.125	0.05	0.201	0.143	0.262	0.151	0.081	0.230
ρ_r	Beta	0.6	0.2	0.817	0.785	0.848	0.852	0.820	0.881
α_p	Beta	0.132	0.1	0.023	0.001	0.050	0.067	0.007	0.138
φ_p	Normal	0.222	0.2	0.185	0.141	0.231	0.135	0.047	0.205
α_w	Beta	0.318	0.1	0.124	0.068	0.168	0.162	0.090	0.233
φ_w	Normal	0.126	0.2	0.235	0.205	0.271	0.248	0.207	0.290
ρ_a	Beta	0.5	0.2	0.798	0.727	0.869	0.839	0.771	0.896
ρ_g	Beta	0.5	0.2	0.784	0.739	0.830	0.802	0.745	0.869
ρ_ζ	Beta	0.5	0.2	0.846	0.786	0.909	0.853	0.787	0.949
σ_a	Inv. Gamma	0.01	2	0.019	0.012	0.026	0.013	0.008	0.017
σ_g	Inv. Gamma	0.01	2	0.055	0.040	0.068	0.046	0.031	0.063
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	0.027	0.023	0.012	0.033

It is clear as the estimated parameters do not change significantly we when consider an alternative detrending method. The table below shows that the results of the model comparison provided in Section 3.4 are robust to a different filter.

Table 11 - Log-marginal data densities and Bayes factors for different models

Model	Log-marginal data density	Bayes factor vs. BASE
BASE	3616.4	
BASE (non-info)	3614.3	$\exp[-2.1]$
EHLind	3571.2	$\exp[-45.2]$
GG	3566.3	$\exp[-50.1]$

B.3 Different model specifications

In what follows we provide estimations for two different model specifications: in the first case we assume a model without persistence in the IS curve (i.e. $h = 0$), whereas in the second case we consider a Taylor rule with an interest rate that follows an $AR(2)$ process.

B.3.1 Model without persistence

We test the robustness of our results to a different specification of the IS curve, assuming that there is no consumption habit (i.e. $h = 0$). The IS curve takes the following form:

$$y_t = E_t y_{t+1} - \frac{1}{\sigma}\left(i_t - E_t \pi^p_{t+1} + E_t g_{t+1} - g_t\right) \tag{108}$$

Table 12 provides the results of our estimation for both the full sample and the Great Moderation sub-sample.

Also in this case the hazard function for both the Phillips curves specifications is upward sloping and this result still holds after the Great Moderation.

Table 12 – Prior and posterior distributions

	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	3.734	2.897	4.612	3.394	2.521	4.317
γ	Gamma	2.0	0.375	2.791	2.264	3.301	2.507	1.979	3.095
δ_π	Normal	1.5	0.25	1.482	1.280	1.672	1.912	1.610	2.179
δ_x	Normal	0.125	0.05	0.228	0.157	0.282	0.168	0.107	0.235
ρ_r	Beta	0.6	0.2	0.740	0.701	0.772	0.788	0.752	0.824
α_p	Beta	0.132	0.1	0.014	0.001	0.030	0.046	0.001	0.091
φ_p	Normal	0.222	0.2	0.294	0.266	0.326	0.240	0.188	0.297
α_w	Beta	0.318	0.1	0.208	0.148	0.264	0.265	0.173	0.381
φ_w	Normal	0.126	0.2	0.278	0.257	0.301	0.258	0.218	0.294
ρ_a	Beta	0.5	0.2	0.549	0.442	0.676	0.713	0.623	0.809
ρ_g	Beta	0.5	0.2	0.875	0.848	0.900	0.892	0.858	0.925
ρ_ζ	Beta	0.5	0.2	0.836	0.769	0.899	0.828	0.751	0.917
σ_a	Inv. Gamma	0.01	2	0.015	0.012	0.020	0.010	0.008	0.013
σ_g	Inv. Gamma	0.01	2	0.033	0.027	0.039	0.027	0.021	0.033
σ_υ	Inv. Gamma	0.01	2	0.002	0.002	0.002	0.002	0.001	0.002
σ_ζ	Inv. Gamma	0.01	2	0.007	0.006	0.009	0.010	0.007	0.013

In the table below we report the results of the model comparison, when no habit persistence in assumed. Our model still clearly outperforms all the alternative considered.

Table 13 - Log-marginal data densities and Bayes factors for different models

Model	Log-marginal data density	Bayes factor vs. BASE
BASE	3473.9	
BASE (non-info)	3470.3	$\exp[-3.6]$
EHLind	3409.3	$\exp[-64.6]$
GG	3403.4	$\exp[-70.5]$

B.3.2 Taylor rule AR(2)

Here we assume that the Taylor rule is characterized by an interest rate that follows an $AR(2)$ stationary process:

$$i_t = \rho_{1r} i_{t-1} + \rho_{2r} i_{t-2} + (1 - \rho_{1r} - \rho_{2r}) (\delta_\pi \pi_t^p + \delta_x y_t) + \upsilon_t \tag{109}$$

Table 14 provides the results of our estimation for both the full sample and the Great Moderation sample.

Table 14 – Prior and posterior distributions

	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	1.180	0.677	1.726	1.238	0.598	1.865
γ	Gamma	2.0	0.375	2.599	2.093	3.036	2.358	1.833	2.864
h	Beta	0.6	0.2	0.913	0.879	0.949	0.918	0.882	0.954
δ_π	Normal	1.5	0.25	1.424	1.194	1.621	1.827	1.513	2.100
δ_x	Normal	0.125	0.05	0.217	0.150	0.274	0.171	0.095	0.238
ρ_{1r}	Beta	0.6	0.2	0.751	0.689	0.814	0.802	0.746	0.854
ρ_{2r}	Beta	0.5	0.2	0.063	0.011	0.115	0.054	0.007	0.098
α_p	Beta	0.132	0.1	0.017	0.001	0.035	0.053	0.001	0.106
φ_p	Normal	0.222	0.2	0.200	0.166	0.232	0.149	0.070	0.229
α_w	Beta	0.318	0.1	0.139	0.094	0.198	0.155	0.083	0.220
φ_w	Normal	0.126	0.2	0.226	0.187	0.258	0.239	0.192	0.285
ρ_a	Beta	0.5	0.2	0.786	0.717	0.860	0.821	0.758	0.900
ρ_g	Beta	0.5	0.2	0.777	0.726	0.824	0.801	0.726	0.858
ρ_ζ	Beta	0.5	0.2	0.801	0.728	0.858	0.795	0.708	0.876
σ_a	Inv. Gamma	0.01	2	0.020	0.014	0.028	0.015	0.010	0.020
σ_g	Inv. Gamma	0.01	2	0.053	0.040	0.068	0.051	0.033	0.068
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	0.022	0.014	0.029	0.030	0.015	0.046

Assuming an $AR(2)$ has not significant implications on the slope of the hazard functions that remains stricly positive; moreover, as showed in the table below, also under this specification, the Bayes factor still strongly favors our model.

Table 15 - Log-marginal data densities and Bayes factors for different models

Model	Log-marginal data density	Bayes factor vs. BASE
BASE	3597.8	
BASE (non-info)	3595.2	$\exp[-2.6]$
EHLind	3555.9	$\exp[-41.9]$
GG	3551.2	$\exp[-46.6]$

Appendix C – Bayesian estimation for different countries

Along this appendix we provide all details and results of the Bayesian estimations performed in Section 3.5. For all the countries we considered quarterly data and the time series used come from FRED database maintained by the Federal Reserve Bank of St. Louis. All the variables has been treated as explained in Section 3.3.1. Parameters calibrations and priors are the same used for the U.S. estimation, except for the parameters defining the hazard function. We assigned to α_p and α_w a Beta distribution centered on 0.1, with standard deviation 0.2; for φ_p and φ_w we assume a Normal distribution with mean 0.15 and standard deviation equal to 0.2.

C.1 Australia

The sample considered covers the period 1980:1 - 2008:4; for the smaller sample we consider 1994:1 - 2008:4, as the new monetary regime, identified by the inflation targeting, started in the first quarter of 1994. The following time series has been used for the estimation: the *GDP by Expenditure in Constant Prices: Total Gross Domestic Product* is used as measure of the output; the *Interbank Rates* 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 *hourly wage rate for all activities,* by the *GDP implicit price deflator*. In Table 16 we report the estimated parameters both for the full sample and for the inflation targeting sub-sample.

Table 16 – Prior and posterior distributions

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (Inflation Targeting)		
	Density	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.406	0.798	2.022	1.230	0.623	1.686
γ	Gamma	2.0	0.375	2.944	2.299	3.581	2.491	1.989	3.041
h	Beta	0.6	0.2	0.881	0.825	0.931	0.926	0.851	0.977
δ_π	Normal	1.5	0.25	1.326	1.107	1.560	1.343	1.050	1.584
δ_x	Normal	0.125	0.05	0.134	0.067	0.216	0.099	0.029	0.190
ρ_r	Beta	0.6	0.2	0.812	0.771	0.857	0.875	0.828	0.919
α_p	Beta	0.2	0.1	0.059	0.010	0.106	0.094	0.024	0.163
φ_p	Normal	0.15	0.2	0.267	0.212	0.318	0.233	0.150	0.318
α_w	Beta	0.2	0.1	0.208	0.119	0.282	0.295	0.213	0.367
φ_w	Normal	0.15	0.2	0.249	0.215	0.283	0.168	0.103	0.226
ρ_a	Beta	0.5	0.2	0.403	0.229	0.577	0.615	0.442	0.786
ρ_g	Beta	0.5	0.2	0.766	0.708	0.833	0.730	0.632	0.840
ρ_ζ	Beta	0.5	0.2	0.777	0.664	0.886	0.887	0.826	0.947
σ_a	Inv. Gamma	0.01	2	0.062	0.032	0.087	0.048	0.022	0.070
σ_g	Inv. Gamma	0.01	2	0.046	0.031	0.059	0.043	0.019	0.062
σ_υ	Inv. Gamma	0.01	2	0.002	0.002	0.003	0.001	0.001	0.001
σ_ζ	Inv. Gamma	0.01	2	0.022	0.013	0.032	0.011	0.006	0.016

C.2 United Kingdom

The sample considered covers the period 1975:1 - 2008:4; UK switched to an inflation targeting regime since 1992:4, thus we consider a smaller sample ranging from 1992:4 - 2008:4. The following time series has been used for the estimation: the *GDP by Expenditure in Constant Prices: Total Gross Domestic Product* is used as measure of the output; the *Treasury Bills* is used for the nominal interest rate. Price inflation is measured using the *Consumer price index for all items* taken in log-difference. Real wage is obtained dividing the nominal wage, measured by the *hourly earnings,* by the *CPI*. In the table below we report the results of our estimation.

Table 17 – Prior and posterior distributions

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (Inflation Targeting)		
	Density	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.475	0.808	2.305	1.319	0.637	1.916
γ	Gamma	2.0	0.375	3.269	2.570	3.953	3.140	2.380	3.798
h	Beta	0.6	0.2	0.961	0.936	0.983	0.903	0.857	0.960
δ_π	Normal	1.5	0.25	1.152	1.011	1.289	1.607	1.351	1.895
δ_x	Normal	0.125	0.05	0.074	-0.006	0.144	0.059	-0.025	0.144
ρ_r	Beta	0.6	0.2	0.896	0.873	0.923	0.876	0.841	0.911
α_p	Beta	0.2	0.1	0.057	0.013	0.112	0.070	0.019	0.115
φ_p	Normal	0.15	0.2	0.224	0.148	0.284	0.195	0.132	0.265
α_w	Beta	0.2	0.1	0.278	0.217	0.347	0.231	0.156	0.312
φ_w	Normal	0.15	0.2	0.139	0.097	0.191	0.170	0.118	0.223
ρ_a	Beta	0.5	0.2	0.904	0.845	0.956	0.818	0.715	0.912
ρ_g	Beta	0.5	0.2	0.839	0.771	0.903	0.818	0.742	0.885
ρ_ζ	Beta	0.5	0.2	0.925	0.860	0.976	0.918	0.855	0.981
σ_a	Inv. Gamma	0.01	2	0.031	0.015	0.042	0.029	0.015	0.044
σ_g	Inv. Gamma	0.01	2	0.127	0.080	0.192	0.051	0.031	0.071
σ_υ	Inv. Gamma	0.01	2	0.002	0.002	0.003	0.001	0.001	0.002
σ_ζ	Inv. Gamma	0.01	2	0.023	0.014	0.031	0.009	0.004	0.013

C.3 France

The sample considered covers the period 1970:1 - 2008:4; France joined the European Monetary Union in 1999:1, thus we consider a smaller sample ranging from 1999:1 - 2008:4. The following time series has been used for the estimation: the *GDP by Expenditure in Constant Prices: Total Gross Domestic Product* is used as measure of the output; the *Treasury Bills* is used for the nominal interest rate. Price inflation is measured using the *Consumer price index for all items* taken in log-difference. Real wage is obtained dividing the nominal wage, measured by the *hourly earnings,* by the *CPI*. Table 18 provides the results of the estimation both for the full sample and for

the European Monetary Union (EMU) sub-sample.

Table 18 – Prior and posterior distributions

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (EMU)		
	Density	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.488	0.744	2.133	1.080	0.491	1.665
γ	Gamma	2.0	0.375	2.938	2.182	3.528	2.368	1.828	3.031
h	Beta	0.6	0.2	0.948	0.918	0.977	0.914	0.851	0.988
δ_π	Normal	1.5	0.25	1.059	0.957	1.174	1.224	0.980	1.435
δ_x	Normal	0.125	0.05	0.136	0.069	0.206	0.104	0.043	0.177
ρ_r	Beta	0.6	0.2	0.898	0.879	0.916	0.837	0.770	0.896
α_p	Beta	0.2	0.1	0.071	0.017	0.114	0.104	0.033	0.178
φ_p	Normal	0.15	0.2	0.124	0.074	0.184	0.081	-0.002	0.162
α_w	Beta	0.2	0.1	0.144	0.085	0.193	0.204	0.134	0.288
φ_w	Normal	0.15	0.2	0.163	0.111	0.213	0.075	0.009	0.158
ρ_a	Beta	0.5	0.2	0.829	0.756	0.899	0.765	0.624	0.908
ρ_g	Beta	0.5	0.2	0.825	0.773	0.869	0.830	0.754	0.913
ρ_ζ	Beta	0.5	0.2	0.784	0.712	0.855	0.905	0.849	0.966
σ_a	Inv. Gamma	0.01	2	0.042	0.027	0.058	0.089	0.025	0.144
σ_g	Inv. Gamma	0.01	2	0.068	0.040	0.096	0.040	0.022	0.064
σ_υ	Inv. Gamma	0.01	2	0.001	0.001	0.001	0.001	0.001	0.001
σ_ζ	Inv. Gamma	0.01	2	0.026	0.017	0.035	0.010	0.005	0.015

C.4 Germany

The model is estimated for the period 1973:1 - 2008:4; a sub-sample is estimated considering the European Monetary Union ranging from 1999:1 - 2008:4. We use the following time series for the estimation: the *GDP by Expenditure in Constant Prices: Total Gross Domestic Product* is used as measure of the output; the *Interbank Rates* is used for the nominal interest rate. Price inflation is measured using the *Consumer price index for all items* taken in log-difference. Real wage is obtained dividing the

nominal wage, measured by the *hourly earnings,* by the *CPI*. The table below shows our estimated parameters for both the samples considered.

Table 19 – Prior and posterior distributions

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (EMU)		
	Density	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.093	0.507	1.718	1.018	0.511	1.630
γ	Gamma	2.0	0.375	2.714	2.185	3.219	1.979	1.450	2.470
h	Beta	0.6	0.2	0.921	0.891	0.961	0.868	0.805	0.935
δ_π	Normal	1.5	0.25	1.339	1.098	1.606	1.442	1.107	1.727
δ_x	Normal	0.125	0.05	0.184	0.127	0.250	0.133	0.061	0.197
ρ_r	Beta	0.6	0.2	0.810	0.771	0.850	0.843	0.791	0.908
α_p	Beta	0.2	0.1	0.042	0.008	0.075	0.083	0.024	0.156
φ_p	Normal	0.15	0.2	0.256	0.213	0.297	0.183	0.109	0.265
α_w	Beta	0.2	0.1	0.193	0.146	0.255	0.206	0.124	0.281
φ_w	Normal	0.15	0.2	0.233	0.202	0.262	0.184	0.123	0.254
ρ_a	Beta	0.5	0.2	0.489	0.353	0.628	0.620	0.438	0.803
ρ_g	Beta	0.5	0.2	0.816	0.765	0.864	0.837	0.761	0.917
ρ_ζ	Beta	0.5	0.2	0.829	0.758	0.907	0.896	0.837	0.966
σ_a	Inv. Gamma	0.01	2	0.038	0.022	0.051	0.041	0.021	0.064
σ_g	Inv. Gamma	0.01	2	0.051	0.036	0.068	0.039	0.020	0.057
σ_υ	Inv. Gamma	0.01	2	0.001	0.001	0.002	0.001	0.001	0.001
σ_ζ	Inv. Gamma	0.01	2	0.011	0.008	0.015	0.008	0.004	0.011

C.5 Italy

The model is estimated for the period 1985:1 - 2008:4; a sub-sample is estimated considering the European Monetary Union ranging from 1999:1 - 2008:4. We use the following time series for the estimation: the *GDP by Expenditure in Constant Prices: Total Gross Domestic Product* is used as measure of the output; the *Treasury Bills* is used for the nominal interest rate. Price inflation is measured using the *Consumer*

price index for all items taken in log-difference. Real wage is obtained dividing the nominal wage, measured by the *hourly earnings,* by the *CPI*. Table 20 provides the results of our estimation.

Table 20 – Prior and posterior distributions

	Prior distribution			Posterior distribution (full sample)			Posterior distribution (EMU)		
	Density	Mean	St. Dev.	Mean	5%	95%	Mean	5%	95%
σ	Gamma	1.0	0.375	1.329	0.531	2.027	1.150	0.590	1.718
γ	Gamma	2.0	0.375	2.617	2.083	3.174	2.424	1.837	3.093
h	Beta	0.6	0.2	0.948	0.920	0.975	0.926	0.868	0.974
δ_π	Normal	1.5	0.25	1.742	1.506	2.042	1.666	1.370	1.972
δ_x	Normal	0.125	0.05	0.133	0.062	0.212	0.118	0.049	0.182
ρ_r	Beta	0.6	0.2	0.848	0.818	0.878	0.816	0.756	0.869
α_p	Beta	0.2	0.1	0.042	0.009	0.068	0.078	0.015	0.135
φ_p	Normal	0.15	0.2	0.208	0.164	0.251	0.150	0.070	0.221
α_w	Beta	0.2	0.1	0.122	0.054	0.187	0.156	0.054	0.267
φ_w	Normal	0.15	0.2	0.272	0.235	0.305	0.203	0.143	0.284
ρ_a	Beta	0.5	0.2	0.730	0.638	0.830	0.831	0.747	0.932
ρ_g	Beta	0.5	0.2	0.779	0.706	0.849	0.810	0.715	0.906
ρ_ζ	Beta	0.5	0.2	0.809	0.718	0.916	0.882	0.801	0.970
σ_a	Inv. Gamma	0.01	2	0.024	0.015	0.035	0.021	0.007	0.033
σ_g	Inv. Gamma	0.01	2	0.075	0.047	0.109	0.072	0.034	0.110
σ_υ	Inv. Gamma	0.01	2	0.001	0.001	0.002	0.001	0.001	0.002
σ_ζ	Inv. Gamma	0.01	2	0.017	0.010	0.024	0.013	0.006	0.019

Appendix D – Phillips curves reduced form

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

$$\pi_t^p = 0.138\pi_{t-1}^p + 0.853E_t\pi_{t+1}^p + 0.01mc_t \tag{110}$$

$$\pi_t^w = 0.153\pi_{t-1}^p - 0.151\pi_t^p + 0.99E_t\pi_{t+1}^w - 0.01\mu_t^w \tag{111}$$

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

$$\pi_t^p = 0.144\pi_{t-1}^p + 0.847E_t\pi_{t+1}^p + 0.009mc_t \tag{112}$$

$$\pi_t^w = 0.106\pi_{t-1}^p + 0.884E_t\pi_{t+1}^w - 0.01\mu_t^w \tag{113}$$

Finally, our time-dependent specification is associated with:

$$\pi_t^p = 0.2\pi_{t-1}^p + 0.99E_t\pi_{t+1}^p - 0.196E_t\pi_{t+2}^p + 0.012mc_t \tag{114}$$

$$\pi_t^w = 0.288\pi_{t-1}^w + 0.99E_t\pi_{t+1}^w - 0.283E_t\pi_{t+2}^w - 0.007\mu_t^w \tag{115}$$

As shown in Chapter 3, the above reduced forms imply that models without time-dependent adjustment capture persistence in the wage equation by price inflation, whereas (114)-(115) 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 E – Welfare function derivation

E.1 Proof of Lemma1

Lemma1 shows the evolution of price dispersion under time-dependent pricing with positive hazard function:

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

Proof:

We define $\overline{P}_t \equiv E_i \log p_t(i)$ and $\Delta_t^p = var_i\left[\log p_t(i) - \overline{P}_{t-1}\right]$.

In Sheedy framework the price level is given by:

$$\begin{aligned}
\log p_t(i) &= \sum_{h=0}^{\infty}\theta_h \log P_{t-h}^* \\
&\Downarrow \\
\log p_t(i) &= (1-\alpha_p)\log p_{t-1}(i) - \varphi_p \log p_{t-2}(i) + \left(\alpha_p+\varphi_p\right)\log P_t^*.
\end{aligned}$$

We define a random variable $X = \log p_t(i) - \overline{P}_{t-1}$. It can be rewritten 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 $\alpha_p\overline{P}_{t-1}$ the latter becomes:

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

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

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

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

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

since $\log P_t^* = \frac{\pi_t + \alpha_p \log p_{t-1}(i) + \varphi_p \log p_{t-2}(i)}{\alpha_p + \varphi_p}$

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

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

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

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

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

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

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

Finally, we obtain:

$$\Delta_t^p = \left(\frac{1-\alpha_p}{1+\varphi_p}\right)\Delta_{t-1}^p + \left[\frac{1-\alpha_p}{\left(1+\varphi_p\right)^2\left(\alpha_p+\varphi_p\right)}\right]\pi_t^2 + \left[\frac{\alpha_p\varphi_p}{\left(1+\varphi_p\right)\left(\alpha_p+\varphi_p\right)}\right]\pi_{t-1}^2 \quad (116)$$

In the case of $\varphi_p = 0$, it encompasses Calvo price dispersion. Iterating (116) forward, the degree of price dispersion in any period $t \geq 0$ under the new policy is given by:

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

We now discount over all periods $t \geq 0$, getting:

$$\sum_{t=0}^{\infty}\beta^t\Delta_t^p = \frac{1}{1-\beta\left(\frac{1-\alpha_p}{1+\varphi_p}\right)}\sum_{t=0}^{\infty}\beta^t\left[\frac{1-\alpha_p}{\left(1+\varphi_p\right)^2\left(\alpha_p+\varphi_p\right)}\pi_t^2 + \frac{\alpha_p\varphi_p}{\left(1+\varphi_p\right)\left(\alpha_p+\varphi_p\right)}\pi_{t-1}^2\right] + t.i.p. \quad (117)$$

QED.

E.2 Final step

Finally, using Lemma 1, we insert (117) in (85) in order to express our welfare function in terms of quadratic variables:

$$W = -\frac{1}{2}E_0\sum_{t=0}^{\infty}\beta^t\left[\frac{\varepsilon_p}{\Theta}\frac{1}{1-\beta\left(\frac{1-\alpha_p}{1+\varphi_p}\right)}\left(\frac{1-\alpha_p}{\left(1+\varphi_p\right)^2\left(\alpha_p+\varphi_p\right)}\pi_t^2 + \frac{\alpha_p\varphi_p}{\left(1+\varphi_p\right)\left(\alpha_p+\varphi_p\right)}\pi_{t-1}^2\right)\right.$$
$$\left. + \left(\sigma + \frac{\gamma+\delta}{1-\delta}\right)y_t^2\right]$$

The welfare function could be rewritten as:

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

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

$$\begin{aligned}
\Omega &= \frac{\varepsilon_p}{\Theta} \frac{1}{1 - \beta \left(\frac{1-\alpha_p}{1+\varphi_p} \right)} \frac{1 - \alpha_p}{\left(1 + \varphi_p\right)^2 \left(\alpha_p + \varphi_p\right)} \\
\Gamma_1 &= \left(\frac{1 + \varphi_p}{1 - \alpha_p} \right) \alpha_p \varphi_p \\
\Gamma_2 &= \left[\frac{\varepsilon_p}{\Theta} \frac{1}{1 - \beta \left(\frac{1-\alpha_p}{1+\varphi_p} \right)} \frac{1 - \alpha_p}{\left(1 + \varphi_p\right)^2 \left(\alpha_p + \varphi_p\right)} \right]^{-1} \left(\sigma + \frac{\gamma + \delta}{1 - \delta} \right)
\end{aligned}$$

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. (2004), "Staggered Prices and Trend Inflation: Some Nuisances," *Review of Economic Dynamics*, 7(3): 642-667.

Ascari, G., and T. Ropele (2007), "Optimal Monetary Policy under Low Trend Inflation," *Journal of Monetary Economics*, 54(8): 2568-2583.

Ascari, G. and T. Ropele (2009), "Trend Inflation, Taylor Principle, and Indeterminacy," *Journal of Money, Credit and Banking*, 41(8): 1557-1584.

Ascari, G., and T. Ropele (2012), "Disinflation in a DSGE Perspective: Sacrifice Ratio or Welfare Gain Ratio?," *Journal of Economic Dynamics and Control*, 36(2): 169-182.

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.

Basu, S., J. Fernald and M. Kimball (2006), "Are Technology Improvements Contractionary?," *American Economic Review*, 96(5): 1418-1448.

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.

Blanchard, O.J. and J. Galí (2007), "Real Wage Rigidities and the New Keynesian Model," *Journal of Money, Credit and Banking*, 39(1): 35-65.

Calvo, G.A. (1983), "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3): 383-398.

Canova, F. and L. Sala (2009), "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, 56(4): 431-449.

Cecchetti, S.G. (1986), "The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines," *Journal of Econometrics*, 31(3), 255-274.

Christiano, L.J. and T.J. Fitzgerald (2003), "The Band Pass Filter," *International Economic Review*, 44(2): 435-465.

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, (1999), "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37(3): 1661-1707.

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.

Fischer, S. (1977), "Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule," *Journal of Political Economy*, 85(1): 191-205.

Fuhrer, J. and G. Moore (1995), "Inflation Persistence," *Quarterly Journal of Economics*, 109(2): 127-159.

Fuhrer, J.C. (1995), "The Persistence of Inflation and the Cost of Disinflation," *New England Economic Review*, issue Jan, 3-16.

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. (1999), "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 89(1): 249-271.

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.

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.

Kydland, F.E. and E.C. Prescott (1977), "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3): 473–491.

Lubik, T.A. and F. Schorfheide (2006), "A Bayesian Look at New Open Economy Macroeconomics," *NBER Macroeconomics Annual 2005*, M. Gertler and K. Rogoff eds. The MIT Press, Cambridge: 316-366.

Lucas, R.E. (1976), "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 19-46.

Mankiw, N.G. and R. Reis, (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117(4): 1295-1328.

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

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

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," *Monetary Policy Rules*, 57-119.

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.

Steinsson, J. (2003), "Optimal Monetary Policy in an Economy with Inflation Persistence," *Journal of Monetary Economics*, 50: 1425-1456.

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.

Taylor, J.B. (2001), "Monetary Policy Rules," *National Bureau of Economic Research Books.*

Walsh, C.E. (2010), "Monetary Theory and Policy," *Cambridge MIT Press*, Cambridge.

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

Woodford, M. (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*, B.J. Friedman and M. Woodford eds. Elsevier Science, North-Holland: 723-828.

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