FEDERAL RESERVE BANK *of* ATLANTA

Comparing New Keynesian Models in the Euro Area: A Bayesian Approach

Pau Rabanal and Juan Francisco Rubio-Ramirez

Working Paper 2003-30
October 2003

Working Paper Series

Federal Reserve Bank of Atlanta
Working Paper 2003-30
October 2003

Comparing New Keynesian Models in the Euro Area: A Bayesian Approach

Pau Rabanal, International Monetary Fund
Juan Francisco Rubio-Ramirez, Federal Reserve Bank of Atlanta

Abstract: This paper estimates and compares four versions of the sticky price New Keynesian model for the Euro area, using a Bayesian approach as described in Rabanal and Rubio-Ramírez (2003). We find that the average duration of price contracts is between four and eight quarters, similar to the one estimated in the United States, while price indexation is found to be smaller. On the other hand, average duration of wage contracts is estimated to between one and two quarters, lower than the one found for the United States, while wage indexation is higher. Finally, the marginal likelihood indicates that the sticky price and sticky wage model of Erceg, Henderson, and Levin (2002), its wage indexation variant, and the baseline sticky price model with price indexation have similar data explanation power, while it positions the baseline sticky price model of Calvo at a lower level.

JEL classification: C11, C15, E31, E32

Key words: nominal rigidities, indexation, Bayesian econometrics, model comparison

---

The authors gratefully acknowledge the Econometric Modelling Unit at the European Central Bank for providing us with the euro area data. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Pau Rabanal, International Monetary Fund, 700 19th Street, N.W., Washington, D.C. 20431, 202-623-6784, Prabanal@imf.org, or Juan Francisco Rubio-Ramirez, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, Georgia 30309, 404-498-8057, juan.rubio@atl.frb.org.

Comparing New Keynesian Models in the Euro Area: A Bayesian Approach

1 Introduction

In this paper, we use a Bayesian approach to estimate and compare the sticky price model of Calvo (1983) and three extensions, using Euro area data. The baseline New Keynesian model of Calvo has become the benchmark for analyzing monetary policy, but its fit to the data has been challenged owing to various grounds.[1] As a result, extensions have been considered to improve its fit to the data. However, the existing literature lacks a formal comparison between competing alternatives.[2]

The first extension we consider in this paper adds price indexation to the baseline model. As a result, both expectations of future inflation, as well as lagged inflation, determine current inflation. The second extension includes staggered wage contracts to the baseline model as in Erceg, Henderson, and Levin (2000). As Galí, Gertler and López-Salido (2001) point out, in a pure forward-looking model, inflation persistence is driven by the sluggish adjustment of real marginal costs. Adding sticky nominal wages delivers sticky real wages, increasing inflation persistence, which is a main shortcoming of the baseline model. Finally, we add wage indexation to the Erceg, Henderson, and Levin (2000) set up.

The paper is divided into two main parts: First, on the estimation side, for each model we combine priors and the likelihood function to obtain the posterior distributions of the structural parameters. Second, we calculate each model's marginal likelihood in order to compare their explanatory power.

Although we are not aware of any formal work comparing different New Keynesian models for the Euro area, different approaches have been used to estimate their structural parameters. Galí, Gertler and López-Salido (2001) estimate the inflation equation of the Calvo model with price indexation and find that the degree of backward looking behavior in the inflation equation is smaller than in the United States, but the degree of price stickiness is somewhat higher. Smets and Wouters (2002) estimate a dynamic general equilibrium model with nominal and real rigidities, and compare it to the fit of statistical Bayesian Vector Autorregressive (BVAR) models.

We use a Bayesian approach because of two main reasons. First, it takes advantage of the general equilibrium approach. As discussed in Leeper and Zha (2000), estimation of reduced-form equations suffers from identification problems. Second, Fernández-Villaverde and Rubio-Ramírez (2003) show that it outperforms GMM and maximum likelihood in small

[1] See Woodford (2003) for a development of monetary policy tools based in the sticky price model, and Fuhrer and Moore (1995), Chari, Kehoe and McGrattan (2000) and Mankiw and Reis (2003) for criticisms of its fit to the data.

[2] Rabanal and Rubio-Ramírez (2003) use the Bayes factor to discrminate betweeen a set of similar models using US data.

samples and, even in the case of misspecification; Bayesian estimation and model comparison are consistent.

The main results of this paper are as follows: First, we estimate an average duration of price contracts between six and eight quarters for the sticky price but flexible wage models. Second, when we introduce sticky wages, the estimated average duration of wage contracts is below two quarters and average duration of price contracts drops to around five quarters. Third, price indexation is important but smaller than in the United States, while wage indexation is larger Finally, the marginal likelihood concludes that price indexation improves the baseline model. On the other hand, it does not provides ground to order the sticky price and sticky wage model of Erceg, Henderson, and Levin (2000), its wage indexation version and the model with price indexation.

The remainder of the paper is organized as follows: In Section 2 we present the baseline sticky price model and the three extensions that we compare. In Section 3 we explain the Bayesian methodology to estimate and compare models. In Section 4 we present and discuss the results, leaving Section 5 for concluding remarks.

2 The Models

In this section we describe the four models. All four models consist of: i) a continuum of infinitely lived households, each of them selling a type of labor that is an imperfect substitute of the other types, ii) a continuum of intermediate good producers, each producing a specific good that is an imperfect substitute for the other goods, and iii) a continuum of competitive final good producers, in the spirit of Blanchard and Kiyotaki (1987). Four types of exogenous shocks are introduced: a technology shock, a monetary shock, a preference shock, and a price mark-up shock. Households have access to complete markets such that we can abstract from distributional issues. Our baseline model assumes that intermediate good producers face restrictions in the price setting process, as in Calvo (1983). We extend this baseline model in three different ways. First, we allow for indexation in prices. Second, we introduce staggered wage contracts, in the spirit of Erceg, Henderson and Levin (2000). Finally, we allow for both staggered wage contracts and indexation in wages.

Since these four models are well know in the literature we only describe the equations that describe the linear dynamics of each model.[3] These equations are obtained by taking a log-linear approximations around the steady state of the first order conditions of households, firms and resource constraints that describe the symmetric equilibrium with zero price and wage inflation rates. The rest of the section is organized as follows: First, we will describe the set of equations that are common to all models. Next, we will discuss the price and wage setting equations, which are different for each model. In what follows, the lower case variables denote log-deviations from the steady state value.

[3]An accurate description of the various models can be found in Erceg, Henderson and Levin (2000), Sbordone (2002), Smets and Wouters (2002) and Rabanal and Rubio-Ramírez (2003).

2.1 Common Equations

First, we have the Euler equation which relates output growth with the real rate of interest

$$y_t = E_t y_{t+1} - \sigma(r_t - E_t \Delta p_{t+1} + E_t g_{t+1} - g_t) \tag{1}$$

where y_t denotes output, r_t is the nominal interest rate, g_t is the preference shifter shock, p_t is the price level, and σ is the elasticity of intertemporal substitution. $E_t(.)$ denotes the expectation operator, and Δ the first difference operator.

The production function and the real marginal cost of production are, respectively:

$$y_t = a_t + (1 - \delta) n_t \tag{2}$$

$$mc_t = w_t - p_t + n_t - y_t \tag{3}$$

where a_t is a technology shock, n_t is the amount of hours worked, mc_t is the real marginal cost, and w_t is the nominal wage. δ is the capital share of output.

The desired marginal rate of substitution (mrs_t) between consumption and hours takes the form:

$$mrs_t = g_t + \frac{1}{\sigma} y_t + \gamma n_t \tag{4}$$

where γ is the inverse elasticity of labor supply with respect to real wages. Hence, the preference shifter shock affects both the consumption Euler equation, and the desired marginal rate of substitution.

We use the following specification for the Taylor rule:

$$r_t = \rho_r r_{t-1} + (1 - \rho_r) \left[\gamma_\pi \Delta p_t + \gamma_y y_t \right] + ms_t \tag{5}$$

where γ_π and γ_y are the long run responses of the monetary authority to deviations of inflation and output from their steady state values, and ms_t is the monetary shock, to be defined below. We include an interest rate smoothing parameter, ρ_r, following recent empirical work (as in Clarida, Galí and Gertler, 2000).

In order to close the model, we need the identity that links real wage growth, nominal wage growth and price inflation:

$$w_t - p_t = w_{t-1} - p_{t-1} + \Delta w_t - \Delta p_t \tag{6}$$

We specify the shocks to follow the following processes:

$$\begin{aligned} a_t &= \rho_a a_{t-1} + \varepsilon_t^a \\ g_t &= \rho_g g_{t-1} + \varepsilon_t^g \\ ms_t &= \varepsilon_t^m \\ \lambda_t &= \varepsilon_t^\lambda \end{aligned} \tag{7}$$

where each innovation ε_t^i follows a $Normal$ $(0, \sigma_i^2)$ distribution, for $i = a, g, m, \lambda$, and innovations are uncorrelated with each other.

Hence, we specify the technology shock and the preference shifter shock to follow AR(1) processes, and the monetary shock and the price mark-up shock to be *iid*. The reasons why we do not allow the two last shocks to follow AR(1) processes are: first, because the Taylor rule already includes an interest rate smoothing component. Second, because we want inflation persistence to be explained endogenously by the models, rather than inheriting the properties of the price mark-up shock.

2.2 Price and Wage Equations

The set of equations described in subsection 2.1 are common to the four models. In this subsection we describe the price and wage equations. These equations differ across models. First we explain the equations for the baseline model (BSP). Then we add price indexation (INDP). Third, we consider the model with sticky prices and sticky wages (EHL). Finally, we add wage indexation (INDW).

2.2.1 Baseline Sticky Price Model (BSP)

The pricing decision of the firm under the Calvo-type restriction delivers the following forward looking equation for price inflation (Δp_t):

$$\Delta p_t = \beta E_t \Delta p_{t+1} + \kappa_p (mc_t + \lambda_t) \tag{8}$$

where $\kappa_p = \frac{(1-\delta)(1-\theta_p\beta)(1-\theta_p)}{\theta_p(1+\delta(\bar{\varepsilon}-1))}$ and $\bar{\varepsilon} = \frac{\bar{\lambda}}{\bar{\lambda}-1}$ is the steady state value of ε, the elasticity of substitution between types of goods. λ_t is the price markup shock, θ_p is the probability of keeping prices fixed during the period, and β is the elasticity of intertemporal substitution.

Equation (8) is the so-called "New Keynesian Phillips Curve", which relates current inflation to expectations of future inflation, the real marginal cost, and the price mark-up shock. It denotes the forward looking behavior of the firms in response to the Calvo-type restriction.

With flexible wages, the usual condition that real wages equal the desired marginal rate of substitution is met:

$$w_t - p_t = mrs_t \tag{9}$$

2.2.2 Model with Sticky Prices and Price Indexation (INDP)

In this case, equation (8) is replaced by:

$$\Delta p_t = \gamma_b \Delta p_{t-1} + \gamma_f E_t \Delta p_{t+1} + \kappa'_p (mc_t + \lambda_t) \tag{8'}$$

where $\kappa'_p = \frac{\kappa_p}{1+\omega\beta}$, $\gamma_b = \frac{\omega}{1+\omega\beta}$, and $\gamma_f = \frac{\beta}{1+\omega\beta}$, and ω is the degree of price indexation. The wage setting equation remains the same (9).

2.2.3 Model with Sticky Prices and Wages (EHL)

In this case, both price and wage inflation behave in a forward looking way. The price inflation equation is given by (8). Introducing the Calvo-type wage restriction delivers the following process for nominal wage growth equation (Δw_t), that replaces (9):

$$\Delta w_t = \beta E_t \Delta w_{t+1} + \kappa_w (mrs_t - (w_t - p_t)) \tag{9'}$$

where $\kappa_w = \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+\phi\gamma)}$, θ_w is the probability of not adjusting wages in a given period, and ϕ is the elasticity of substitution between different types of labor in the production function. With staggered wage setting, it is no longer true that workers remain on their desired labor supply schedule all the time. Hence, the driving force of current nominal wage growth is expected nominal wage growth, as well as the distance between the desired marginal rate of substitution and the real wage.

2.2.4 Model with Sticky Prices, Wages and Wage Indexation (INDW)

This model extends EHL in that the nominal wage growth equation (9') incorporates indexation:

$$\Delta w_t - \alpha \Delta p_{t-1} = \beta E_t \Delta w_{t+1} - \alpha\beta \Delta p_t + \kappa_w (mrs_t - (w_t - p_t)) \tag{9''}$$

where α is the degree of wage indexation.

3 Empirical Analysis

In this section, we explain how to draw from the posterior distribution of the structural parameters and evaluate the log of the marginal likelihood of the data implied by each model. We report the mean and the standard deviation of the posterior distributions and the difference between the log of the marginal likelihoods of each model with respect to BSP.

3.1 The Data

There are several problems when choosing data for the euro area as a whole. The first problem to overtake is that, even though member countries in the European Union have converged to a unified system of national accounts, an aggregate data set for the area does not exists, since national currencies did fluctuate among themselves, and its retroactive creation is complicated. The Econometric Modeling Unit at the European Central Bank has constructed a "synthetic" data set for the euro area to overcome this problem.[4]

[4] See Fagan, Henry and Mestre (2001) for details. This dataset should not be viewed as an "official" ECB series but rather as a synthetic dataset constructed by the Econometric Modelling Unit for research purposes.

The starting date is the second dilemma we face. Obviously, the largest structural break in the euro area has been the launch of the euro in 1999. If we use the "synthetic" data we have to assume that monetary policy was also conducted in an aggregated way. The operating procedures of the different central banks were quite different during the 1980s and 1990s. Following Rabanal (2003), we choose 1984:01 as our starting date. We make the assumption that the whole euro area switched monetary policy around the same time as the United States did, and the public accepted and understood that change at the same time. Even though the conduct of monetary policy by the European central banks and the Federal Reserve was quite different, Smets and Wouters (2002), using a data from 1970:1, show that a Taylor rule would approximate the behavior of the "synthetic" European Central Bank conduct of policy quite well.

Hence, we explain the joint behavior of price inflation, real wages, interest rates, and output for the "synthetic" euro area at a quarterly frequency. The sample period is 1984:01 to 2002:03. The real variables are detrended using the Hodrick and Prescott (HP) filter, while nominal variables are treated as deviations from their unconditional mean.

3.2 The Likelihood Function

Let $\psi = \left(\sigma, \theta_p, \theta_w, \beta, \phi, \alpha, \gamma_y, \gamma_\pi, \rho_r, \delta, \bar{\lambda}, \gamma, \rho_a, \rho_g, \sigma_a, \sigma_m, \sigma_g, \sigma_\lambda\right)'$ be the vector of structural parameters, $x_t = (w_t - p_t,\ r_t,\ \Delta p_t,\ \Delta w_t,\ y_t,\ n_t,\ mc_t,\ mrs_t,\ c_t)'$ be the vector of endogenous variables, $z_t = (a_t, ms_t, g_t, \lambda_t)'$ be the vector of shocks, and $\varepsilon_t = \left(\varepsilon_t^a, \varepsilon_t^m, \varepsilon_t^g, \varepsilon_t^\lambda\right)'$ be the vector of their innovations.

The system of equations (1)-(9) can be written the following way

$$A(\psi) E_t x_{t+1} = B(\psi) x_t + C(\psi) x_{t-1} + D(\psi) z_t,$$

$$z_t = N(\psi) z_{t-1} + \varepsilon_t, \quad E(\varepsilon_t \varepsilon_t') = \Sigma(\psi).$$

We use standard solution methods for linear models with rational expectations to write the law of motion in state-space form and the Kalman filter to evaluate the likelihood of the four observable variables $d_t = (r_t, \Delta p_t, w_t - p_t, y_t)'$. We denote by $L\left(\{d_t\}_{t=1}^T | \psi\right)$ the likelihood function of $\{d_t\}_{t=1}^T$.

3.3 The Priors

We denote by $\pi(\psi)$ the prior distribution of the parameters. We present the list of the structural parameters and its associated prior distributions in Tables 1 and 2. Table 1 reflects our priors on the parameters that do not change across models, while Table 2 reflects the priors on the parameters that do change across models. Notice that prior distribution is the same used in Rabanal and Rubio-Ramírez (2003). However, our choice of priors is not that far off from the one used by Smets and Wouters (2002) for the euro area.

The inverse of the elasticity of intertemporal substitution, σ^{-1}, follows a gamma distribution. Our choice implies a prior mean of 2.5 and a prior standard deviation of 1.76. Given the wide variety of estimates for this parameter we do not see this prior as a very odd one. We also pick a gamma distribution for the average duration of prices.[5] Our selection entails that the average duration of prices has a prior mean of 3 and a prior standard deviation of 1.42. This alternative reflects the facts presented in Taylor (1999) for the United States.

Regarding the Taylor rule coefficients we select normal distributions. We set the mean of γ_π to 1.5 and that of γ_y to 0.125, which are Taylor's original guesses.[6] We also use a normal distribution for the prior of the inverse of the elasticity of the labor supply, γ, centered at 1 and with a standard deviation of 0.5. The interest rate smoothing coefficient, ρ_r, the autorregresive parameter of the technology, ρ_a, and the autorregresive parameter of preference shifter, ρ_g, have a uniform prior distribution between $[0,1)$. Finally, we opt for a prior uniform distribution between $[0,1)$ for the all standard deviations of the innovations of the stochastic shocks. The reason for this choice are twofold: First, we do not have strong prior information about the standard deviations of the innovations. Second, the lower the estimated σ_λ, the higher the estimated κ_p necessary to explain the observed inflation volatility. Since there is a negative relationship between κ_p and θ_p, the higher κ_p, the lower the estimated θ_p. Therefore, truncation of σ_λ can result in underestimation of θ_p. We want to preclude the underestimation of θ_p and be symmetric on the prior assumptions for all four standard deviations, therefore we opt for high prior upper bound on all four of then.

We imposed dogmatic priors over the parameters β, δ, ϕ and $\overline{\varepsilon}$. The reasons are as follows: First, since we do not consider capital, we have had trouble estimating β and δ. Second, there is an identification problem between the probability of the Calvo lottery, θ_p, and the mean of the price markup, $\overline{\varepsilon}$.[7] Therefore, it is not possible to identify θ_p and $\overline{\varepsilon}$ at the same time. Similarly, the same problem emerges between θ_w and ϕ. The values we use ($\beta = 0.99$, $\delta = 0.36$, $\phi = 6$ and $\overline{\varepsilon} = 6$) are quite conventional in the literature.

[Insert Table 1 here]

In Table 2, we present the priors for the parameters that differ across models. In the BSP model, wages are flexible and there is no price indexation. Therefore, we set θ_w, α and ω to zero. In the INDP model, while we maintain θ_w and α equal to zero, we choose a prior uniform distribution between 0 and 1 for the price indexation parameter, ω. In the EHL model, we set the two indexation parameters, α and ω, to zero, and we establish a gamma distribution for the prior duration of wages with mean of four quarters and standard deviation of 1.71. This choice is motivated because we expect that wage contracts to be fixed

[5] Since we need to keep the probability of the Calvo lottery between 0 and 1, we formulate the prior in terms of the parameter $1/(1-\theta_p)-1$.

[6] Taylor(1993) uses annualized Federal Funds rates and inflation data, while we use quaterly data for all series. Therefore, we would need to multiply our γ_y prior mean by four to make it comparable to Taylor's results.

[7] The slope of the Phillips curve, κ_p, is the only one containing $\overline{\varepsilon}$ and θ_p.

for a longer period of time than price contracts. The priors for the INDW model add to those of the EHL model the fact that the prior distribution for the wage indexation parameter, α, is assumed to be an uniform distribution between 0 and 1.

[Insert Table 2 here]

Finally, we censor the support of all parameters to the region where the model has a unique, stable solution. Therefore, we rule out indeterminacies due to interest rate rules that do not place enough weight on inflation.

3.4 Drawing from the Posterior and Model Comparison

Let M be the set of models that we wish to compare, where $M = \{BSP, INDP, EHL, INDW\}$. The posterior distribution of the structural parameters for each model $m \in M$ is:

$$p(\psi| \{d_t\}_{t=1}^{T}, m) \propto L(\{d_t\}_{t=1}^{T} |\psi, m)\pi(\psi|m).$$

Given our priors and the likelihood functions implied by the models, we are not able to obtain a closed-form solution for the posterior distributions. However, since we are able to evaluate both expressions numerically, we use the random walk Metropolis-Hastings algorithm, to obtain a random draw of size 500,000 from $p\left(\psi| \{d_t\}_{t=1}^{T}, m\right)$.

Having specified the likelihood function and the prior distribution, the marginal likelihood of each model is:

$$L\left(\{d_t\}_{t=1}^{T} |m\right) = \int_{\Psi} L\left(\{d_t\}_{t=1}^{T} |\psi, m\right) \pi(\psi|m)\, d\psi. \tag{10}$$

This expression involves averaging the likelihood function across the parameter space, using the prior distributions as a weighting function. Exact computation of (10) is impossible, therefore, we follow Geweke (1998) to estimate (10). We denote the estimated marginal likelihood implied by model m by $\widehat{L}\left(\{d_t\}_{t=1}^{T} |m\right)$. Once we obtain $\widehat{L}\left(\{d_t\}_{t=1}^{T} |m\right)$, we compute the Bayes factor between two distinct models n and m

$$\widehat{B}_m^n\left(\{d_t\}_{t=1}^{T}\right) = \frac{\widehat{L}\left(\{d_t\}_{t=1}^{T} |n\right)}{\widehat{L}\left(\{d_t\}_{t=1}^{T} |m\right)}.$$

As shown in Fernández-Villaverde and Rubio-Ramírez (2003), if $m^* \in M$ is the best model under the Kullback-Leibler distance, then for any other $n \in M$, $\widehat{B}_{m^*}^n\left(\{d_t\}_{t=1}^{T}\right)$ converges to zero as T increases. Hence, we focus on the Bayes factor as a tool to determine which model best explains the joint behavior of our four variables.

Both the Metropolis-Hastings algorithm and Geweke's (1998) procedure to estimate (10) are described in the appendix. In the appendix are also described all the numerical and convergence issues related to the implementation of the Metropolis-Hastings algorithm.

4 Findings

4.1 Posterior Distributions and Moments

Tables 3 and 4 present the mean and the standard deviation of the posterior distributions of the parameters for the four models.[8] For an easier discussion of the results, we present the parameters that relate directly to the price and wage equations and the Taylor rule in Table 3. We leave for Table 4 the rest of the parameters that involve preferences, technology, and the process of the underlying shocks.

[Insert Table 3 here]

The first column of Table 3 presents the estimates for the BSP model. The posterior mean of the average duration of price contracts is 6.59 quarters, while its standard deviation is 0.32.[9] We do not consider that this estimate implies a too long average duration of price contracts, and they are similar to those reported by Galí, Gertler and López-Salido (2001). The estimates of the Taylor rule are as follows: the posterior mean of coefficient on inflation is 1.30, while the posterior mean of coefficient on output is 0.26. Both posterior standard deviations are small. The interest rate smoothing posterior mean is 0.73.

The second column of Table 3 reports the results of the INDP model. The main differences are that the estimated coefficient on price indexation is 0.49, somewhat higher than what Rabanal (2003), Smets and Wouters (2002) and Galí, Gertler and López-Salido (2001) obtained for the euro area, and the estimated average duration of price contracts increases to 6.06 quarters. The estimates of the Taylor rule for the INDP model are almost identical to those obtained for the BSP model.

We present the EHL model in the third column of Table 3. The estimated average duration of price contracts is 6.01 quarters. A surprising result is the low estimated average duration of wage contracts. The average duration of wage contracts is less than two quarters, 1.23. This is puzzling because our priors indicate that we expected that wage contracts have longer average durations than price contracts.[10] The estimated Taylor rule is very close to the one obtained for models with flexible wages. The only difference is that this specification implies a higher interest rate smoothing parameters. Last column of Table 3 presents the estimates of the INDW model. The wage indexation parameter, α, is estimated to be 0.34 while price and wage average contract durations are lower (4.50 and 1.24, respectively) than in the EHL model.

[8] In order to save space, we do not plot histograms of the posterior distributions. They are available at the following URL address http://www.econ.umn.edu/~rubio/graphs2.html

[9] Our results depend on the particular values chosen for the discount factor, β, and the mean of the price markup, $\overline{\varepsilon}$. However, for a reasonable range of values for those parameters, the average duration of prices does not change significantly.

[10] In this case, there are interactions between the degree of monopolistic competition in wage setting, ϕ, and the duration of wage contracts. Playing around with the parameters suggests that it is difficult to obtain a significantly higher duration of wage contracts.

The rest of the estimated parameters are in Table 4. The posterior mean of the elasticity of intertemporal substitution, σ, extends from 0.17 to 0.24.[11] The parameter that manages the labor supply, γ, is model dependent. This reveals the fact that when agents cope with wage rigidities they cannot supply their desired amount of labor anymore. We estimate values close to 1 for the models with flexible wages (BSP and INDP) while they are closer to 2 for the models with wage stickiness (EHL and INDW). Finally, we find high and similar correlation coefficients for the technology and preference-shifter shocks.

The posterior mean for σ_λ is always larger than 18% (being 55.14% in the case of the INDP model). This result validates the choice our prior distribution for σ_λ. As a comparison, all other standard deviation estimates are lower than 5%.

We would like to remark the following facts. First, data grants support for an average duration of price contracts between four and eight quarters and a average duration of wage contracts of less than two quarters. Second, price and wage indexation are important for the euro area. Finally, the estimates on the coefficient that explains labor supply, γ, depends greatly on the underlying assumption about the wage-setting process.

4.2 Model Comparison

Last row of Table 4 reports the difference between the log marginal likelihood of each model with respect to log marginal likelihood of BSP. The results are as follows:

The first question we need to answer is: Is there price indexation in the euro area?. The log marginal likelihood difference between INDP and BSP is 13.41. Therefore, in order to choose BSP over INDP, we need a prior probability over model BSP 6.66×10^5 (= exp(13.41)) times larger than our prior probability over INDP. This evidence supports the assumption of price indexation.

The second question is: Does the inclusion of sticky wages improve the model? The log marginal likelihood difference between EHL and INDP is 1.53. This implies that we need a prior probability over INDP 4.62 times lager that our prior over EHL in order reject the fact that sticky wages improves the model. This factor is very low, so the data does not allow us to favor EHL over INDP. (Neither does it allow us to favor INDP over EHL).

The third question is: How much does wage indexation add to EHL? In this case we would only need to have a prior probability over EHL 8 (= exp(2.08)) times larger that our prior over INDW in order to choose EHL. This factor is similar to the one reported before, so, we conclude that wage indexation does not improve the ability of the EHL model to explain the data.

Therefore, we reach two conclusions: First, the marginal likelihood criterion indicates that price indexation is an important factor in the euro area. Second, once we consider price indexation, wage stickiness and wage indexation do not help to explain the data much better.

[11] Similar evidence is reported for the United States, see for instance De Jong, Ingram and Whiteman (2000) and Basu and Kimball (2000).

5 Concluding Remarks

In this paper, we have used a Bayesian approach to estimate and compare the baseline sticky price model of Calvo (1983) and three extensions, using euro area data. We have restricted ourselves to estimate simple and tractable models that are commonly used in the analysis of monetary policy. Our main results differs somewhat to what we obtained for the United States. In that paper, we found that the sticky price and wage model of Erceg, Henderson and Levin (2000) ranked best among the four. Here, we find that sticky wages in the form of Calvo staggering is rejected by the data.

These results are consistent with the evidence already flagged in Rabanal (2003), who rejects sticky wages in the euro area once habit formation in consumption is introduced. By no means we try to say that wages are flexible in Europe, but in future research it would be interesting to incorporate other factors that can affect labor market dynamics: some examples can be found in the efficiency wage literature (Felices, 2003) and in the labor market search literature (as in Trigari, 2003).

Finally, we would like to mention the importance of the marginal likelihood as a model comparison device. Classical estimation approaches do not allow us to rank a set of nonnested models based on their ability to explain the data. Here, we combine the use of dynamic general equilibrium models with powerful numerical algorithms to perform such an exercise in a Bayesian framework.

References

[1] Basu, Susanto, and Miles S. Kimball (2000), "Long-Run Labor Supply and the Elasticity of Intertemporal Substitution," mimeo, University of Michigan.

[2] Blanchard Olivier J., and Nobuhiro Kiyotaki (1987), "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review 77,* pp. 647-666.

[3] Calvo, Guillermo (1983), "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics 12*, pp.383-398.

[4] Chari, V.V., Patrick Kehoe, and Ellen McGrattan (2000), "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?," *Econometrica 68,* pp. 1151-1181.

[5] Clarida, Richard, Jordi Galí, and Mark Gertler (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics 115*, pp. 147-180.

[6] DeJong, David N., Beth F. Ingram, and Charles H. Whiteman (2000), "A Bayesian Approach to Dynamic Macroeconomics," *Journal of Econometrics 98*, pp. 203-223.

[7] Erceg, Chris J., Dale W. Henderson, and Andrew T. Levin (2000),"Optimal Monetary Policy with Staggered Wage , and Price Contracts," *Journal of Monetary Economics 46,* pp. *281-313.*

[8] Felices, Guillermo (2002), "Efficiency wages in a new Keynesian framework," Ph.D. dissertation, New York University.

[9] Fernández-Villaverde, Jesús, and Juan F. Rubio-Ramírez (2003), "Comparing Dynamic Equilibrium Economies to Data: A Bayesian Approach," *Journal of Econometrics,* forthcoming.

[10] Fuhrer, Jeffrey C., and George Moore (1995), "Inflation Persistence," *Quarterly Journal of Economics 110*, pp. 127-160.

[11] Galí, Jordi, Mark Gertler, and David López-Salido (2001), "European Inflation Dynamics," *European Economic Review 45*, pp. 1237-1270.

[12] Gelfand A.E., and D.K. Dey (1994), "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society. Series B (Methodological)* 56, pp. 501-514.

[13] Gelman, A., G.O. Roberts and W.R. Gilks (1996). "Efficiente Metropolis Jumping Rules" In J.O. Berger *et al.* (eds.) *Bayesian Statistics 5*, Oxford University Press.

[14] Geweke, John (1998), "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," *Federal Reserve Bank of Minneapolis Staff Report* 249.

[15] Leeper, Eric, and Tao Zha (2000), "Assessing Simple Policy Rules: A View from a Complete Macro Model," *Federal Reserve Bank of Atlanta Working Paper* 2000-19.

[16] Mankiw, N. Gregory, and Ricardo Reis (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve" *Quarterly Journal of Economics*, vol. 117 (4), pp. 1295-1328.

[17] Rabanal, Pau and Juan F. Rubio-Ramírez (2003) "Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach," *Federal Reserve Bank of Atlanta* Working Paper 2001–22a.

[18] Rabanal, Pau (2003), "The Cost Channel of Monetary Policy: Further Evidence for the United States and the Euro Area," *IMF Working Paper* 03/149.

[19] Sbordone, Argia (2001), "An Optimizing Model of US Wage and Price Dynamics ," Rutgers University, mimeo.

[20] Smets, Frank, and Raf Wouters (2003), "An Estimated Stochastic Dynamic General Equilibrium Model for the Euro Area," *Journal of European Economic Association*, forthcoming.

[21] Taylor, John (1993), "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Series on Public Policy 39,* pp.195-214.

[22] Taylor John (1999), "Staggered Price and wage-setting in Macroeconomics," in John Taylor , and Michael Woodford (eds.), *Handbook of Macroeconomics,* North-Holland, Elsevier, pp. 1009-1050.

[23] Trigari, Antonella (2003) "Labor Market Search, Wage Bargaining and Inflation Dynamics", New York University, mimeo.

[24] Woodford, Michael (2003) "Interest and Prices : Foundations of a Theory of Monetary Policy", Princeton University Press.

6 Appendix

6.1 The Metropolis-Hastings-Algorithm

In order to obtain a draw of size N from the posterior distribution:

1. We start with an initial value ψ_0. From this value, we evaluate $L\left(\{d_t\}_{t=1}^{T}|\psi_0\right)\pi(\psi_0)$

2. For each i,

$$\widehat{\psi}_i = \begin{cases} \widehat{\psi}_{i-1} \text{ with probability } 1-R \\ \psi_i^* \text{ with probability } R \end{cases},$$

   where $\psi_i^* = \widehat{\psi}_{i-1} + v_i$, v_i follows an iid multivariate normal distribution

   and $R = \min\left\{1, \frac{L\left(\{d_t\}_{t=1}^{T}|\psi_i^*\right)\pi(\psi_i^*)}{L\left(\{d_t\}_{t=1}^{T}|\widehat{\psi}_{i-1}\right)\pi\left(\widehat{\psi}_{i-1}\right)}\right\}$.

A final important issue is to assess the convergence of the simulated draw from the posterior distribution. In particular, it is extremely to adjust the parameters of the transition density (in the case of the random walk, the variance of the innovation term) to get an appropriate acceptance rate[12]. If the acceptance rate is very small, the chain will not visit a set large enough in any reasonable number of iterations. If the acceptance rate is very high, the chain will not tend to stay enough time in high probability regions. Gelman, Roberts and Gilks (1996) suggest that a 20% acceptance rate tends to give the best performance. We found that, in our models, an acceptance rate of around 35% outperformed different alternatives and it was the target used to adjust the variance of the proposal density. We draw chain of size 500.000 and the acceptance rates are 31.95% for BSP, 32.08% for INDP, 35.59% for EHL and 35.26% for INDW.

6.2 Obtaining the Marginal Likelihood

For each model, given a draw $\{\widehat{\psi}_i\}_{i=1}^{N}$, we build the marginal likelihood as follows. Gelfand and Dey (1994) note that for any k_m-dimensional probability density, $g(.)$, with support contained in Ψ,

$$E\left[\frac{g(\psi)}{L(\{d_t\}_{t=1}^{T}|\psi,m)\pi(\psi|m)}|\{d_t\}_{t=1}^{T},m\right] = L\left(\{d_t\}_{t=1}^{T}|m\right)^{-1}.$$

Using our draw, we can compute

$$L\left(\{d_t\}_{t=1}^{T}|m\right)^{-1} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{g(\psi_i)}{L(\{d_t\}_{t=1}^{T}|\psi_i,m)\pi(\psi_i|m)}\right].$$

[12] The acceptance rate is equal to the number of times when the chain changes position divided by the number of iterations.

As a choice of $g(.)$, we follow Geweke's (1998) and define

$$\Sigma_N = \frac{1}{N}\sum_{i=1}^{N}(\psi_i - \bar{\psi}_N)(\psi_i - \bar{\psi}_N)';$$
$$\bar{\psi}_N = \frac{1}{N}\sum_{i=1}^{N}\psi_i.$$

Then, for a given $p \in (0,1)$, define the set

$$\Psi_M = \{\psi_i : (\psi_i - \bar{\psi}_N)(\Sigma_N)^{-1}(\psi_i - \bar{\psi}_N) \leq \chi^2_{1-p}(k_m)\},$$

where $\chi^2_{1-p}(.)$ is a chi-squared distribution with degrees of freedom equal to the number of parameters in $\widehat{\psi}_i$, k_m. Note that we are taking into account the fact that the number of estimated parameters can be different for each model. Letting $I_{\Psi\cap\Psi_M}(.)$ be the indicator function of a vector of parameters belonging to the intersection $\Psi \cap \Psi_M$, we can take a truncated multivariate normal as our $g(.)$ function:

$$g(\psi) = \frac{1}{\hat{p}(2\pi)^{\frac{k}{2}}}|\Sigma_N|^{\frac{1}{2}}\exp[-0.5\Upsilon_N]I_{\Psi\cap\Psi_M}(\Psi);$$
$$\Upsilon_N = (\psi_i - \bar{\psi}_N)(\Sigma_N)^{-1}(\psi_i - \bar{\psi}_N),$$

where $\hat{p}$ is an appropriate normalizing constant. With this choice, if the posterior density is uniformly bounded away from zero on every compact set of Ψ, our computation approximates the likelihood function. With the output of the Markov chain Monte Carlo, we use the computed values of $L(\{d_t\}_{t=1}^T|\psi_i)\pi(\psi_i)$ and find its harmonic mean using the function g as a weight.

Table 1: Prior Distributions for the Common Parameters

Parameter		Mean	Std.Dev.
σ^{-1}	$\text{gamma}(2, 1.25)$	2.5	1.76
$\frac{1}{1-\theta_p} - 1$	$\text{gamma}(2, 1)$	2	1.42
γ_π	$\text{normal}(1.5, 0.25)$	1.5	0.25
γ_y	$\text{normal}(0.125, 0.125)$	0.125	0.125
γ	$\text{normal}(1, 0.5)$	1	0.5
ρ_a, ρ_g, ρ_r	$\text{uniform}[0, 1)$	0.5	0.28
$\sigma_a, \sigma_m, \sigma_g, \sigma_\lambda$	$\text{uniform}[0, 1)$	0.5	0.28
β	-	0.99	-
ϕ	-	6	-
$\overline{\varepsilon}$	-	6	-
δ	-	0.36	-

Table 2: Prior Distributions for Wage Duration and Indexation Parameters

Parameter	Model			
	BSP	INDP	EHL	INDW
$\frac{1}{1-\theta_w} - 1$	0	0	gamma$(3, 1)$	gamma$(3, 1)$
ω	0	Uniform$(0, 1)$	0	0
α	0	0	0	Uniform$(0, 1)$

Table 3: Posterior Distributions for Price- and Wage-Setting Parameters and the Taylor Rule

	BSP	INDP	EHL	INDW
Price Duration	6.50 (0.32)	8.06 (0.38)	6.01 (0.40)	4.50 (0.44)
Wage Duration	1 (–)	1 (–)	1.23 (0.08)	1.24 (0.10)
α	– (–)	– (–)	– (–)	0.34 (0.14)
ω	– (–)	0.49 (0.10)	– (–)	– (–)
γ_π	1.30 (0.12)	1.32 (0.12)	1.33 (0.14)	1.41 (0.18)
γ_y	0.26 (0.06)	0.25 (0.05)	0.29 (0.08)	0.26 (0.07)
ρ_r	0.73 (0.03)	0.72 (0.03)	0.83 (0.03)	0.82 (0.03)

Table 4: Posterior Distributions for the Remaining Parameters

	BSP	INDP	EHL	INDW
σ	0.19 (0.08)	0.18 (0.06)	0.24 (0.13)	0.17 (0.06)
γ	0.95 (0.20)	1.00 (0.20)	1.59 (0.39)	1.64 (0.33)
ρ_a	0.93 (0.01)	0.92 (0.02)	0.93 (0.02)	0.87 (0.03)
ρ_g	0.80 (0.05)	0.80 (0.05)	0.85 (0.04)	0.84 (0.04)
$\sigma_a(\%)$	0.57 (0.14)	0.56 (0.12)	0.77 (0.02)	0.96 (0.03)
$\sigma_m(\%)$	0.14 (0.01)	0.15 (0.02)	0.12 (0.01)	0.13 (0.01)
$\sigma_\lambda(\%)$	40.03 (2.30)	55.14 (2.45)	34.30 (3.51)	18.01 (3.42)
$\sigma_g(\%)$	3.69 (0.87)	3.81 (0.94)	3.74 (1.25)	4.25 (0.09)
$log(\hat{L})$	–	7.48	8.05	10.13