WP/03/149

IMF Working Paper

The Cost Channel of Monetary Policy: Further Evidence for the United States and the Euro Area

Pau Rabanal

INTERNATIONAL MONETARY FUND

 WP/03/149

IMF Working Paper

Monetary and Financial Systems Department

The Cost Channel of Monetary Policy: Further Evidence for the United States and the Euro Area

Prepared by Pau Rabanal[1]

Authorized for distribution by David Marston

July 2003

Abstract

This paper estimates the importance of the cost channel of monetary policy in a New Keynesian model of the business cycle. A model with nominal rigidities is extended by assuming that a fraction of firms need to borrow money to pay their wage bill. Hence, monetary policy tightenings increase effective unit labor costs of production, and might imply an increase in inflation. The model explains the joint dynamics of output, inflation, real wages, and interest rates, and is estimated using a Bayesian framework and data for the United States and the euro area. The main result is that cost channel effects are absent in both cases. Moreover, it is not possible to obtain a "price puzzle" type of behavior from estimated impulse responses to monetary policy shocks.

JEL Classification Numbers: C11, C15, E31, E32

Keywords: Price Puzzle, New Keynesian Models, Bayesian Methods, United States, Euro Area

Author's E-Mail Address: prabanal@imf.org

[1] I am grateful to the Econometric Modeling Unit at the European Central Bank, and especially to Alistair Deppe, for providing me with the euro area data. I would like to extend my thanks to Andreas Billmeier, Oya Celasun, Federico Ravenna, Juan Francisco Rubio-Ramírez and Robert Tchaidze for helpful comments and discussions, and to Natalie Baumer, Karen (K.C.) Craig and India Owen for editorial assistance. All remaining errors and omissions are mine.

Contents | Page

I. Introduction

What is the effect of monetary policy on prices? The conventional view suggests that monetary policy tightenings are associated with declines in output and inflation. However, one of the most controversial findings in the empirical literature on monetary policy shocks is the so-called "price puzzle," whereby a tightening of monetary policy is associated with an increase, rather than a decrease, in the price level. Two main explanations have been offered for this phenomenon: one implies that the unexpected part of monetary policy shocks is not well measured, while the other suggests that there are "cost channel" effects of monetary policy.

The first explanation suggests that vector autoregressive (VAR) models cannot measure the forward-looking component of monetary policy, and hence, do not properly measure monetary policy shocks. Suppose that, for whatever reason, the central bank expects higher inflation in the future, due to productivity shocks, oil price shocks, exchange rate developments, and the like. When the central bank increases interest rates, those shocks may have already been built into the economy, so we simultaneously observe an increase in interest rates and prices. Therefore, the price puzzle arises due to a misidentification of the unexpected component of monetary policy shocks. Sims (1992) suggested that once commodity prices are included in a VAR model, the price puzzle disappears. His explanation was that the information set available to policy makers may include variables useful in forecasting inflation that the econometrician has not considered. These series, like commodity prices, should also be used by econometricians to control for developments that have an inflationary effect outside of monetary policy.

The second explanation suggests that there is no methodological problem with a price puzzle type of behavior. On the contrary, it is indeed the cost channel of monetary policy that causes prices (or inflation) and nominal interest rates to move in the same direction after a monetary policy shock. This supply side effect of monetary policy may coexist with and, in fact, dominate the traditional, demand side, effect. When the central bank increases interest rates, some production (financing) costs increase, and an inflationary component appears in the nominal interest rate. Barth and Ramey (2001) reach this conclusion using industry level data for the United States, and show that their finding is robust even when commodity prices are introduced in the VAR.

This paper attempts to disentangle these two conflicting explanations. We specify a dynamic general equilibrium (DGE) model that allows us to identify both monetary policy shocks and cost channel effects, and we estimate the relevance of the cost channel of monetary policy, using a Bayesian approach. The use of DGE models based on staggered price and wage setting (i.e., New Keynesian models) has become increasingly popular for the analysis of monetary policy, due to their analytical tractability. At the same time, when complemented by the appropriate real rigidities, they performed well in explaining the behavior of key macro variables for the United States and the euro area.[2] Hence, in this paper, a model with nominal rigidities is extended by assuming that a fraction of firms have to borrow funds since they have to pay for their production inputs before selling their product. An inflation dynamics equation (a new Phillips curve) is then derived, explicitly incorporating the nominal interest rate as a determinant of inflation. Afterward, the elasticity of inflation to fluctuations in the nominal interest rate is estimated. By constructing a model that allows for a cost channel, we examine to what extent this is a feature of the aggregate data, and its relevance in monetary policy making.

The Bayesian approach combines prior information and the likelihood of the data to obtain the posterior distribution of the parameters. We use the Kalman filter to evaluate the likelihood function of a linear approximation of the model and a numerical algorithm (the Metropolis-Hastings algorithm) to draw from the posterior distribution. There does not exist an analytical expression for the posterior, and until recently only a very restrictive set of priors and models could be used. The recent advent of powerful Markov chain Monte Carlo (MCMC) techniques,[3] as well as fast computer processors, has removed this restriction and a more general set of models and priors can now be implemented. The model's accuracy of fit is addressed via production of the ratio of marginal likelihoods or Bayes factors. The marginal likelihood averages all possible likelihoods across the parameter space, using the prior as a weight. When conveniently weighted, it can be interpreted as the probability of observing the data under a given model.

There are several advantages of using a Bayesian approach. First, it is a flexible method that allows the researcher to introduce prior information about the model's parameters. Calibration exercises, which are used to address how well a model fits the data, lack a formal statistical foundation. Classical methods cannot accommodate even the most noncontroversial prior information about the model's parameters. Hence, ours is a more general method, with pure calibration and full-information maximum likelihood being special cases. Second, the Bayesian framework allows for consistent parameter estimation and model comparison of fundamentally misspecified models, as shown by Fernández-Villaverde and Rubio-Ramírez (2001).[4] Third, from a computational point of view, the use of prior distributions over the parameters makes the estimation algorithm more stable. This is particularly valuable when the size of the parameter space is large and only relatively small

[2] See Christiano, Eichenbaum and Evans (2001), Smets and Wouters (2002), and Rabanal and Rubio-Ramírez (2002), among others.

[3] See Geweke (1998) for a survey.

[4] In this paper, we conduct tests only between nested models, but the Bayesian approach allows us to test between nonnested models as well. See Rabanal and Rubio-Ramírez (2002).

samples of data are available.[5] Finally, the Bayesian strategy takes advantage of the general equilibrium approach: all the theoretical restrictions implied by the model for the likelihood function and the full dimension of the data are taken into account for estimation. Partial equilibrium approaches, such as the GMM method used by Galí and Gertler (1999) and Ravenna and Walsh (2003), may have identification problems.

In our view, this paper also improves upon the recent literature by estimating a New Keynesian model without conditioning on the nature of shocks. For instance, Rotemberg and Woodford (1997), Boivin and Giannoni (2003) or Christiano, Eichenbaum and Evans (2001) try to match model-based and estimated impulse-response functions to a monetary policy shock. In our view, this approach suffers from three main shortcomings: first, it relies on the identification scheme of the VAR. Second, there appears to be no evident connection between the estimates from the reduced form VAR and the parameters of the structural model. Third, in the context of this paper, if there are cost channel effects, these should be present for a wide variety of shocks, and not just monetary policy shocks.

The main results of the paper are as follows: we find that (i) the cost channel effect is small, but somewhat stronger in the United States than in the euro area, (ii) it is not possible to obtain a price puzzle type of behavior with our point estimates, (iii) estimates for all structural parameters are similar to what has been typically found in the literature, especially regarding the interest rate rule, and the average duration of price contracts, (iv) the only exception is the estimated average duration of wages, both for the United States and the euro area, which is less than two quarters. Hence, we conclude that cost channel effects are not relevant in a dynamic, stochastic, general equilibrium New Keynesian model.

The rest of the paper is organized as follows: in Section II, we present a model with nominal rigidities, which incorporates the cost channel; Section III presents the linearized version of the model, while Section IV offers the econometric methodology for parameter estimation. In Section V, the main results are discussed, while concluding remarks are left for Section VI.

II. The Model

The model consists of a now fairly standard New Keynesian (NK) model with sticky prices and wages, which will be modified to incorporate a cost channel of monetary policy. The model includes all the ingredients necessary to properly characterize output, inflation, interest rate and real wage dynamics for the U.S. and the euro area. Sticky prices are introduced in order to have short-term real effects of monetary policy. The introduction of sticky wages, as in Erceg, Henderson, and Levin (2001), is necessary for two reasons. First,

[5] Classical approaches include Kim (2000) and Ireland (2001), who estimate general equilibrium models by full-information maximum likelihood, which can be inefficient when using small samples.

because sticky prices *and* sticky nominal wages deliver sticky *real wages*, which is what we observe in the data;[6] and second, because in an optimizing model, inflation depends on the real marginal cost of production. In order to avoid jumpy inflation behavior we will need to dampen the behavior of real marginal costs of production; sticky wages, in addition to sticky prices, will help us achieve this goal.

In addition to having staggered price and wage contracts, the model includes two other features that have been shown necessary to match key macro variables in the United States and the euro area. First, we introduce habit formation in consumption. Boldrin, Christiano and Fischer (2001) use habit formation to reconcile the behavior of asset prices with the business cycle. In the context of this model, we need to assume this type of preferences to capture the output persistence that is usually absent in pure forward-looking models. Second, price indexation is introduced in the inflation equation, in order to capture the inflation persistence that we observe in the data, and that is difficult to generate with pure forward-looking models.

The structure of the goods and labor markets is monopolistic competition as in Blanchard and Kiyotaki (1987). It is assumed that both price and wage setters face a Calvo (1983) type of restriction when setting their prices and wages optimally. The model consists of: (i) a continuum of identical households, indexed by $j \in [0,1]$, each supplying a different type of labor that is an imperfect substitute for the other labor types; (ii) a continuum of intermediate goods producers, indexed by $i \in [0,1]$, each supplying a type of good that is an imperfect substitute for the other goods; and (iii) a continuum of identical final goods producers.

It is important to mention that we are ultimately interested in explaining the joint behavior of four variables. Therefore, we need four sources of stochastic shocks. Otherwise, we would have a situation where not all variables are linearly independent. This would not be a problem in terms of finding a solution to the model, but would cause problems later on because the likelihood function of the data could not be evaluated. The four shocks that the model incorporates are: monetary shock, fiscal shock, technology shock, and price markup shock.

A. Households

Each household supplies a differentiated type of labor, and is indexed by $j \in [0,1]$. They obtain utility from consuming the final good (C_t^j), and from holding real money balances (M_t^j / P_t), and obtain disutility from supplying hours of labor (N_t^j). The lifetime utility function can be expressed as:

$$E_t \sum_{\tau=0}^{\infty} \beta^{t+\tau} [u(C_{t+\tau}^j, C_{t+\tau-1}^j) + v(\frac{M_{t+\tau}^j}{P_{t+\tau}}) - w(N_{t+\tau}^j)], \tag{1}$$

[6] If we only have one of the two nominal rigidities, real wages are too volatile.

where the following functional forms are assumed:

$$u(C_t^j, C_{t-1}^j) = \frac{(C_t^j - bC_{t-1}^j)^{1-\gamma}}{1-\gamma}$$

$$v(\frac{M_t^j}{P_t}) = \frac{\left(\frac{M_t^j}{P_t}\right)^{1-\xi}}{1-\xi}$$

$$w(N_t^j) = \frac{\left(N_t^j\right)^{1-\eta}}{1-\eta}.$$

E_t denotes the rational expectations operator using information up to time t. $\beta \in [0,1]$ is the discount factor. The instant utility function at any point in time depends on the quasi-difference between the levels of consumption in the current and last period. $b \in [0,1]$ denotes the importance of the habit stock. Habit formation in consumption has been suggested as a way to reconcile the behavior of consumption and asset prices.[7] This paper includes this feature because purely forward-looking models have difficulty matching the persistence in output that we observe in the data. If b=0, then we go back to the baseline case where preferences on consumption are time separable. γ>0 is the inverse of the elasticity of intertemporal substitution, ξ>0 is the elasticity of real money balance holdings, and η>0 is the elasticity of labor supply with respect to the real wage.

Households maximize utility subject to the usual budget constraint:

$$P_t C_t^j + M_t^j + \frac{B_t^j}{R_t} = W_t^j N_t^j + M_{t-1}^j + B_{t-1}^j + T_t^j + \int_0^1 \Pi_t^j(i) di, \tag{2}$$

where P_t is the price of the final good, W_t^j is the nominal wage, and B_t^j denotes holdings of a riskless bond that costs the inverse of the gross nominal interest rate (R_t>1) and pays one unit of currency next period. T_t^j denotes nominal transfers from (or lump-sum taxes paid to) the government. The last term of the right hand side of the previous expression denotes the profits from the monopolistically competitive intermediate goods producers firms, which are ultimately owned by households.

As is well known in this class of models (see Erceg, Henderson, and Levin, 2001), we assume that there exist state-contingent securities that insure households against variations in household specific labor income. In order to keep notation simple, the structure of the

[7] See Boldrin, Christiano, and Fischer (2001).

complete asset markets is not explicitly introduced. However, it is a well known result that with complete markets, households' wealth is insured, and the consumption/savings decisions and the labor supply decision can be separated. Moreover, the consumption decision will be independent of current wealth and identical across households and states of nature. Therefore the j subscript in consumption can be safely ignored.

The first order conditions to the consumption/savings decision are:

$$\Lambda_t = (C_t - bC_{t-1})^{-\gamma} - b\beta E_t (C_{t+1} - bC_t)^{-\gamma} \tag{3}$$

$$\Lambda_t = E_t\left[\frac{\Lambda_{t+1} R_t P_t}{P_{t+1}}\right] \tag{4}$$

$$\left(\frac{M_t}{P_t}\right)^{-\xi} = \Lambda_t\left(\frac{R_t - 1}{R_t}\right), \tag{5}$$

where Λ_t is the Lagrange multiplier associated with the intertemporal budget constraint. Equation (3) simply states that the Lagrange multiplier equals the marginal utility of consumption. If there were no habit formation, we would get the usual condition that the marginal utility of consumption is decreasing. With habit formation, terms involving past and expected future consumption matter. Equation (4) reflects the intertemporal condition for holding bonds. Equation (5) reflects the money demand equation, and weights the utility of holding money with its opportunity costs.

Since the structure of the labor markets is monopolistic competition, households choose their labor supply schedule maximizing utility, and facing a downward sloping for their type of labor. Before we can derive such condition, we have to describe the technology of the intermediate and final good producers, so we leave the wage setting decision for a later section.

B. Final Good Producers

There is a continuum of final good producers, operating under perfect competition. The technology to produce the aggregate final good is:

$$Y_t = \left[\int_0^1 \left(Y_t^i\right)^{\frac{\varepsilon_t - 1}{\varepsilon_t}} di\right]^{\frac{\varepsilon_t}{\varepsilon_t - 1}}, \tag{6}$$

where $\varepsilon_t > 1$ is the elasticity of substitution between types of goods, Y_t is the final good, and Y_t^i are the intermediate goods. Since the price markup is related to the elasticity of substitution, we will have time varying price markups, as in Giannoni (2001). Profit maximizing from the final goods producers delivers the following demand for each intermediate good:

$$Y_t^i = \left(\frac{P_t^i}{P_t}\right)^{-\varepsilon_t}, \text{ for all } i \in [0,1], \tag{7}$$

where P_t, the price of the final good, is obtained from the zero profit condition in the final goods sector:

$$P_t = \left[\int_0^1 \left(P_t^i\right)^{1-\varepsilon_t} di\right]^{\frac{1}{1-\varepsilon_t}},$$

and P_t^i are the prices of intermediate goods.

C. Intermediate Goods Producers and the Cost Channel

The production function for intermediate goods producers is:

$$Y_t^i = A_t \overline{K}^{\delta} N_{i,t}^{1-\delta}, \tag{8}$$

where A_t is an economy wide technology factor and $N_{i,t}$ is the unit of effective labor input used by firm i.[8] $0<\delta<1$ is the capital share of output, and the level of capital is fixed in the short run at a level $\overline{K}$.

In order to obtain one unit of effective labor, firms employ all types of labor from households, which are aggregated the following way:

$$N_{i,t} = \left[\int_0^1 \left(N_{i,t}^j\right)^{\frac{\phi-1}{\phi}} dj\right]^{\frac{\phi}{\phi-1}}. \tag{9}$$

Each firm chooses their labor demand schedules in order to obtain the optimal labor mix. As a result, aggregating across firms we obtain the following downward sloping demand for each type of labor j:

$$N_t^j = \left(\frac{W_t^i}{W_t}\right)^{-\phi} N_t, \text{ for all j} \in [0,1], \tag{10}$$

[8] Note that the notation is different for total hours worked by a household (N_t^j) and for total hours employed by a firm ($N_{i,t}$).

where N_t is an index of aggregate labor, and the aggregate wage index is defined as:[9]

$$W_t = \left[\int_0^1 \left(W_t^j\right)^{1-\phi} dj\right]^{\frac{1}{1-\phi}}.$$

In order to explicitly include a cost channel of monetary policy, it is assumed that intermediate goods producers indexed as $i \in [0,\alpha]$ have to pay workers every period before they sell their product. These firms borrow at the riskless nominal interest rate. Hence, for a fraction α of firms the nominal wage bill is $R_t \int_0^1 W_t^j N_{i,t}^j dj$, while for the remaining 1-α the nominal wage bill is simply $\int_0^1 W_t^j N_{i,t}^j dj$. This timing of events implies that for some firms there exists a cost channel of monetary policy, since an increase in the nominal interest will increase real unit labor costs.

D. Price and Wage Setting under Staggered Contracts

Prices and wages are set by intermediate goods producers and households in a staggered way. For convenience, it is assumed that prices and wages are set at random intervals as in the model of Calvo (1983). Agents can only adjust prices or wages whenever they receive a stochastic signal to do so. The probability of receiving this signal is independent of the past history of signals and across agents. This assumption greatly simplifies the aggregation of price and wage setting decisions.

We denote by θ_p the probability of the Calvo lottery for the price setters. Then, the average duration of price contracts is $1/(1-\theta_p)$. When firms face a Calvo type restriction, they set prices maximizing the discounted sum of profits taking into account that the price that they fix today might not be reset optimally for some time. We assume that there is a degree $\omega \in [0,1]$ of indexation to last period's inflation rate (P_{t-1}/P_{t-2}), whenever firms are not allowed to reset prices optimally.

The first order condition for profit maximization is:

$$E_t \sum_{\tau=0}^{\infty} \left(\theta_p \beta\right)^{\tau} \left\{ \left[\frac{P_t^*}{P_{t+\tau}} \left(\frac{P_{t+\tau-1}}{P_{t-1}} \right)^{\omega} - \mu_{t+\tau}^p \overline{MC}_{i,t+\tau} \right] \overline{Y}_{t+\tau}^i \right\} = 0, \qquad (11)$$

[9] The aggregate wage index comes from using $\int_0^1 W_t^j N_{i,t}^j dj = W_t N_t^j$, which can be viewed as a "zero profit condition" in the labor aggregating activity.

where

$$\overline{Y}^i_{t+\tau} = \left[\frac{P^*_t}{P_{t+\tau}} \left(\frac{P_{t+\tau-1}}{P_{t-1}} \right)^{\omega} \right]^{-\varepsilon_t} Y_{t+\tau}$$

is the demand at $t+\tau$ assuming that the price set at t is not optimally reset. $\mu^p_t = \varepsilon_t / (\varepsilon_t - 1)$ is the time varying price markup. The marginal cost of production depends on the labor's share of output and the nominal interest rate, if applicable:

$$\begin{aligned} \overline{MC}_{i,t+\tau} &= \frac{W_{t+\tau} N_{i,t+\tau}}{(1-\delta) P_{t+\tau} \overline{Y}^i_{t+\tau}}, \qquad \text{for i} \in [0, \alpha] \\ \overline{MC}_{i,t+\tau} &= R_{t+\tau} \frac{W_{t+\tau} N_{i,t+\tau}}{(1-\delta) P_{t+\tau} \overline{Y}^i_{t+\tau}}, \quad \text{for i} \in [\alpha, 1]. \end{aligned} \tag{12}$$

Finally, in the symmetric equilibrium, the evolution of the price level is:

$$P_t = \left\{ \theta_p \left[P_{t-1} \left(\frac{P_{t-1}}{P_{t-2}} \right)^{\omega} \right]^{1-\varepsilon_t} + (1-\theta_p) \left(\hat{P}^*_t \right)^{1-\varepsilon_t} \right\}^{\frac{1}{1-\varepsilon_t}}, \tag{13}$$

where $\hat{P}^*_t$ is the sum of optimal prices set at time t:

$$\hat{P}^*_t = \left[\alpha \left(\widetilde{P}^*_t \right)^{1-\varepsilon_t} + (1-\alpha) \left(\widetilde{\widetilde{P}}^*_t \right)^{1-\varepsilon_t} \right]^{\frac{1}{1-\varepsilon_t}}.$$

$\widetilde{P}^*_t$ is the optimal price in the symmetric equilibrium for the "non-cost channel" firms and $\widetilde{\widetilde{P}}^*_t$ is the optimal price for the "cost channel" firms.

Households face the same restriction to set their wages. Let θ_w denote the probability of the Calvo lottery for the wage setters. The associated average duration of wage contracts is $1/(1-\theta_w)$. Therefore, households' labor supply schedule comes from choosing their wage to maximize utility facing a downward sloping demand for their type of labor. The first order condition is:

$$E_t \sum_{\tau=0}^{\infty} (\theta_w \beta)^{\tau} \left\{ \left[\Lambda_t \frac{W^*_t}{P_{t+\tau}} - \mu^w \overline{N}^j_{t+\tau} \right] \overline{N}^j_{t+\tau} \right\} = 0, \tag{14}$$

where

$$\overline{N}^j_{t+\tau} = \left[\frac{W^*_t}{W_{t+\tau}}\right]^{-\phi} N_{t+\tau}$$

is the labor demand at $t+\tau$ assuming that the wage set at t is not optimally reset. $\mu^w = \phi/(\phi-1)$ is the wage markup. If wages were fully flexible, we would obtain the intratemporal condition that real wages equal the marginal rate of substitution between consumption and hours. However, with staggered contracts, agents know that they might not be able to reset their wage in the near future. Hence they take into account current and future expected deviations between the desired marginal rate of substitution and the real wage, using the corresponding probability as a weight.

In the symmetric equilibrium, the aggregate wage level evolves as:

$$W_t = \left[\theta_w W_{t-1}^{1-\phi} + (1-\theta_w)\left(W_t^*\right)^{1-\phi}\right]^{\frac{1}{1-\phi}}.$$

E. Monetary and Fiscal Policy

As in Taylor (1993), it is assumed that the monetary authority conducts monetary policy with an interest rate rule. Then, the nominal amount of money is determined via the money demand equation. The government cannot run deficits or surpluses, so it finances transfers and other government spending (G_t) via money creation:

$$T_t + G_t = M_t - M_{t-1}. \tag{15}$$

The role of transfers in this paper is to undo the inefficiency related to the cost channel of monetary policy. In steady state, for the cost channel firms, the wage bill is higher as long as $\overline{\mathrm{R}} > 1$. Different wage bills will lead to different steady state prices and production levels. In order to avoid these effects, we assume that the government undoes this inefficiency by subsidizing the "cost channel" firms.

Autonomous government spending is assumed to be a proportion of GDP:

$$G_t = (1-\Gamma_t)Y_t$$

such that the resource constraint can be written as:

$$Y_t = C_t + G_t = C_t + (1-\Gamma_t)Y_t$$

or

$$\Gamma_t Y_t = C_t. \tag{16}$$

For a given level of output, a decrease in government spending (increase in Γ_t) will increase private consumption.

III. THE LINEARIZED MODEL

The model's dynamics are obtained by taking a log linear approximation of equations (3) to (16) around the symmetric equilibrium steady state with zero inflation. Lower case variables denote percent (log linear) deviations from the steady state value. Table 1 presents the linearized equations, which give the dynamics of the endogenous variables (λ_t, c_t, r_t, Δp_t, Δw_t, y_t, $w_t - p_t$, n_t, mc, mrs_t). The equations have been separated into four main blocks: output determination, price setting, wage setting and the monetary policy rule.

Table 1: Linearized System

Output Determination	
$\lambda_t = \frac{1}{\sigma(1-b\beta)}[bc_{t-1} - (1+b^2\beta)c_t + b\beta E_t c_{t+1}]$	Marginal Utility of Consumption
$\lambda_t = r_t - E_t\Delta p_{t+1} + E_t\lambda_{t+1}$	Euler Equation for Holding Bonds
$y_t = a_t + (1-\delta)n_t$	Production Function
$y_t = c_t - \gamma_t$	Resource Constraint
Price Setting	
$\Delta p_t = \gamma_b \Delta p_{t-1} + \gamma_f E_t \Delta p_{t+1} + \kappa_p mc_t$	New Keynesian Phillips Curve
$mc_t = w_t - p_t + n_t - y_t + \alpha r_t + \mu_t^p$	Real Marginal Cost of Production
Wage Setting	
$\Delta w_t = \beta E_t \Delta w_{t+1} + \kappa_w [mrs_t - (w_t - p_t)]$	Nominal Wage Growth
$mrs_t = \eta n_t - \lambda_t$	Marginal Rate of Substitution
$w_t - p_t = w_{t-1} - p_{t-1} + \Delta w_t - \Delta p_t$	Real Wage Identity
Interest Rate Rule	
$r_t = \rho_r r_{t-1} + (1-\rho_r)(\gamma_p \Delta p_t + \gamma_y y_t) + mp_t$	Taylor Rule

where $\sigma = 1/\gamma$ is the inverse of the elasticity of intertemporal substitution, and Δ is the first difference operator. The rest of parameters are as follows: $\gamma_b = \omega/(1+\omega\beta)$, $\gamma_f = \beta/(1+\omega\beta)$, $\kappa_p = (1-\delta)(1-\theta_p\beta)(1-\theta_p)/[\theta_p(1+\omega\beta)(1+\delta(\bar{\varepsilon}-1)]$, where $\bar{\varepsilon}$ is the steady state value of ε, and $\kappa_w = (1-\theta_w\beta)(1-\theta_w)/[\theta_w(1+\phi(\eta-1)]$.

The first block of four equations determines output. As mentioned earlier, the definition of the marginal utility of consumption (MUC) involves terms with lagged and expected future inflation. The intertemporal condition for holding bonds relates the MUC today and tomorrow with the real rate of interest. Output is produced using hours and the labor augmenting technology shock. Finally, the resource constraint relates output with private consumption and government spending.

As in Galí and Gertler (1999) and Sbordone (2001), the model-based driving force of inflation is the real marginal cost of production. Usually, measures that proxy for the position of the economy with the business cycle (such as the output gap or unemployment) are used in an ad-hoc way. The real marginal cost includes the labor's share of output. Nominal interest rates enter the definition of real marginal costs, and may have an inflationary effect. Also, the price markup shock affects the behavior of inflation. The term involving lagged inflation appears because of the backward looking indexation assumption. If ω is set equal to zero, we return to a pure forward-looking specification.

The third block denotes the wage setting process. As with the price equation, nominal wage growth depends on expected nominal wage growth and the distance between the desired marginal rate of substitution between consumption and leisure and the real wage. This block also includes the identity that links the real wage with nominal wage growth and price inflation.

Finally, the Taylor rule in linear terms reacts to deviations of inflation and output to their steady state values. $\gamma_p>1$ and $\gamma_y>0$ denote the long run responses of the nominal interest rate to inflation and output fluctuations. In addition, an interest rate smoothing component is included, following the empirical evidence in Clarida, Galí and Gertler (2000), and Smets and Wouters (2002).

The Taylor rule includes a shock, which can be subject to various interpretations. This paper interprets this shock as a monetary policy shock. The central bank uses the Taylor rule as a main guide for conducting monetary policy, but also observes more variables and indicators than the econometrician does (e.g., exchange rates, asset prices, government deficits, consumer confidence) and uses that information to "fine tune" the desired nominal interest rate. It is assumed that the monetary shock is *iid*. The dynamics of the money demand equation are ignored. This choice follows the tradition of the New Keynesian literature that assumes that nominal money plays no role, since monetary policy is conducted via an interest rate rule.

Finally, the four shocks evolve as follows:

$$\begin{aligned} a_t &= \rho_a a_{t-1} + \varepsilon_t^a \\ \gamma_t &= \rho_\gamma \gamma_{t-1} + \varepsilon_t^\gamma \\ mp_t &= \varepsilon_t^m \\ \mu_t^p &= \varepsilon_t^p . \end{aligned} \tag{17}$$

The technology and government spending shocks evolve as AR(1) processes, where $\rho_a, \rho_g \in [0,1]$ and the innovations $\varepsilon_t^a, \varepsilon_t^g, \varepsilon_t^m, \varepsilon_t^p$ follow mean zero Normal distributions. There are two reasons why the monetary and price markup shocks are assumed to be uncorrelated. The first one is that this choice reduces the number of parameters to be estimated. The second reason is that both the Taylor rule and the New Keynesian Phillips Curve equations already include lagged terms, which should be enough to induce persistence in these variables, without having to rely on autoregressive shocks.

IV. Econometric Methodology

The linearized model presented in the previous section is estimated using a Bayesian approach. This section explains how to implement this approach, using data for the United States and the euro area, and how to use the Bayes factor to assess the importance of cost channel effects. Applying a Bayesian approach to parameter estimation implies obtaining the posterior distribution of the parameters conditional on the data. From Bayes rule, we know that the posterior is proportional to the product of the likelihood function and the prior:

$$P(\psi \mid \{d_t\}_{t=1}^T) \propto L(\{d_t\}_{t=1}^T \mid \psi) P(\psi),$$

where ψ is the vector of parameters that describe the model, $\{d_t\}_{t=1}^T$ is the vector of endogenous variables, T is the sample size, $L(\{d_t\}_{t=1}^T \mid \psi)$ is the conditional likelihood function of the data on the model and the parameters, and $P(\psi)$ is the prior distribution of the parameters.[10]

Four main ingredients are necessary to implement a Bayesian framework for parameter estimation. The first one consists in the data we want to explain. Second, we need to specify the priors for the model's parameters. The third step consists of obtaining the law of motion of the (linearized) model and the evaluation of the likelihood function of the data. In this case, we have to make use of standard methods for solving linear rational expectations systems to write the law of motion in state space form, and then evaluate the likelihood function via the Kalman filter. If analytical expressions existed for the two components, it would be straightforward to obtain a tractable expression for the posterior distribution. However, the expression of the likelihood function is not analytically tractable. Hence, in the fourth step we have to rely on numerical methods to simulate the posterior distribution, obtain the relevant moments of the posterior distribution of the parameters, as well as the computation of the marginal likelihood and the Bayes factor.

---

[10] The denominator of the Bayes formula, which would be the marginal likelihood of the data for each model, is not included since it is constant with respect to the value of the parameters.

A. The Data

This paper explains the comovement of output, price inflation, real wages and nominal interest rates for the United States and the euro area. Real variables are detrended using the Hodrick and Prescott (HP) filter, while nominal variables are treated as deviations from their mean. Data sources for the United States are the following: the Federal Funds rates is used as the relevant nominal interest rate. This series was obtained through the Federal Reserve Bank of Saint Louis database (FRED). The measure of output is the "Nonfarm Business Sector Output," as published by the Bureau of Labor Statistics (BLS). The measure of prices is the associated price deflator. Finally, the measure of nominal wages is the "Hourly Compensation for the Nonfarm Business Sector," also obtained through the BLS. The choice of variables is done for comparability with previous studies. The sample period is 1984:01 to 2002:03, at a quarterly frequency. The starting point of the sample reflects the fact that many authors[11] have suggested a structural break in the conduct of monetary policy by the Federal Reserve around 1982. In particular, changes involved the switch to an interest rate targeting regime instead of targeting nonborrowed reserves, and the strong anti-inflationary stance. However, it might be the case that it took some time before this change in policy was credible and accepted by the public.

Choices for the euro area are a little bit more difficult to make. First of all, finding data for the euro area as a whole is complicated. Even though member countries in the European Union have converged to a unified system of national accounts, the construction of retroactive data suffers from many complications. The most important are that before 1999, national currencies did in fact fluctuate between themselves, and that the weights of national currencies in the euro and in the ECU are not the same. The European Central Bank Econometric Modeling Unit has constructed a synthetic dataset for the euro area, which is detailed in Fagan, Henry and Mestre (2001).[12]

The starting date for the sample period in the euro area is also controversial. The strongest structural break is in fact the launch of the euro in January 1999. The operating procedures of the Federal Reserve and the German Bundesbank were quite different during the 1980s and 1990s. At the same time, other countries like Spain or Italy did not adopt an anti-inflationary stance until much later in the sample period. We make the choice that the starting date is also 1984:01, that is, we make the assumption that the euro area switched monetary policies 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 Bundesbank

---

[11] See Clarida, Galí and Gertler (2000) and the references therein.

[12] This dataset should not be viewed as an official historical series of macro variables of the European Central bank, but rather as a synthetic dataset of its Econometric Modeling Unit.

and the Federal Reserve was quite different, Smets and Wouters (2002) show that a Taylor rule would approximate the behavior of the "synthetic" European Central Bank conduct of policy quite well.

B. Prior Distributions

This section describes the prior distributions adopted for the estimation. Let $\psi = (\psi'_1, \psi'_2, \psi'_3)'$ be the vector of all parameters that characterize the model, which is partitioned as follows: $\psi_1 = (\sigma, b, \theta_p, \omega, \theta_w, \alpha, \gamma_p, \gamma_y, \rho_r)$, $\psi_2 = (\beta, \phi, \bar{\varepsilon}, \eta, \delta)$, and $\psi_3 = (\rho_a, \rho_\gamma, \sigma_a, \sigma_\gamma, \sigma_m, \sigma_p)$. This partition is not arbitrary: while ψ_1 and ψ_2 contain the structural parameters of the model, ψ_3 contains the parameters that describe the model's exogenous shocks. The parameters in ψ_1 will be estimated, while the parameters in ψ_2 will be calibrated, using degenerate priors.

There are three main reasons to proceed this way. First, we can see by looking at the expressions in the price and wage setting equation that many parameters are not identified. The choice here is to focus on the estimation of average price and wage durations, the cost channel elasticity, and the backward-looking parameter in price inflation. Hence, some parameters need to be fixed. Second, the values of the discount factor and the capital share of output would be indirectly estimated in a model without capital, as the one used in this paper. Third, in order to reduce the dimensionality of the parameter space, the parameters that relate to the price and wage markup, and the labor supply elasticity are held fixed.

Table 2 presents the prior distributions of the parameters. The prior for the inverse elasticity of intertemporal substitution is a Gamma distribution in order to stay in positive values. In theoretical papers, this parameter is usually calibrated at a value of one to be consistent with the existence of a balanced growth path. For this parameter, empirical values in the literature are set between 2 and 4.[13] The mean of the prior is 2.5 and the standard deviation is large enough to incorporate enough uncertainty about this parameter. The habit formation in consumption parameter is a Normal distribution centered at 0.7, a value close to that suggested by Boldrin, Christiano and Fischer (2001), with a standard deviation of 0.05. The prior distribution is truncated at six standard deviations from the mean, so it can effectively take values between 0.4 and 1.

The duration of prices and wages is also a Gamma distribution. In order to keep the probabilities of the Calvo lottery between zero and one, the prior is specified in terms of "average duration minus one quarter." These priors imply an average duration of three quarters for prices, and of four quarters for wages, following the informal evidence presented in Taylor (1999). Moreover, these priors incorporate the fact that we expect wages to be

[13] See Basu and Kimball (2000).

fixed for longer periods of time than prices. The coefficient on the cost channel has a prior uniform distribution between zero and one. The same prior is used for the coefficient on backward-looking price indexation.

Table 2: Prior Distributions

Parameter		Mean	Standard Dev.
σ^{-1}	Gamma(2,1.25)	2.50	1.76
b	Normal(0.7,0.05)	0.70	0.05
$1/(1-\theta_p)-1$	Gamma(2,1)	2.00	1.42
$1/(1-\theta_w)-1$	Gamma(3,1)	3.00	1.76
ω	Uniform(0,1)	0.50	0.28
α	Uniform(0,1)	0.50	0.28
γ_p	Normal(1.5,0.25)	1.50	0.25
γ_y	Normal(0.5,0.125)	0.50	0.13
ρ_r	Uniform(0,1)	0.50	0.28
ρ_a,ρ_λ	Uniform(0,1)	0.50	0.28
$\sigma_a,\sigma_\lambda,\sigma_p,\sigma_m$	Gamma(4,0.25)	1.00	0.50

The prior distributions of the coefficients of the Taylor rule are Normal distributions, centered at Taylor's original values. The parameter space is censored to the region where the model has a unique, stable solution. Therefore, indeterminate solutions due to an interest rate rule that does not place enough weight on the coefficient of price inflation are ruled out. The interest rate smoothing parameter is allowed to take any value between zero and one, with a uniform prior. The priors of the coefficients on the autoregressive parameters are set to uniform distributions between zero and one, while the priors on the standard deviations of the shocks are Gamma distributions to stay in positive reals, with mean of 1 percent and standard deviation of 0.5 percent.

The following degenerate priors are used for the coefficients in ψ_2. The capital share of output is set to $\delta = 0.36$, and the discount factor to $\beta = 0.99$, which are values fairly standard in the literature. The values for the elasticity of substitution between types of goods and labor are set to $\phi = \bar{\varepsilon} = 6$, which implies steady state price and wage markups of 20 percent. The value for the inverse elasticity of labor supply with respect to the desired real wage is set to $\eta = 1$. This value is consistent with the estimates of Rabanal and Rubio-Ramírez (2002) for a model with staggered price and wage contracts. Below, we discuss some sensitivity analysis for these parameters.

C. The Law of Motion and the Likelihood Function

Let $x_t = \{\Delta p_t, y_t, r_t, w_t - p_t, mc_t, mrs_t, \lambda_t, n_t, c_t\}'$ be the vector of endogenous variables, $z_t = \{a_t, \gamma_t, mp_t, \mu_t^p\}$, the vector of exogenous variables, and $\zeta_t = \{\varepsilon_t^a, \varepsilon_t^\gamma, \varepsilon_t^m, \varepsilon_t^p\}$, the vector of innovations. Then, the system of equations presented in Table 1, can be written as follows:

$$
\begin{aligned}
&A(\psi_1,\psi_2)E_t x_{t+1} = B(\psi_1,\psi_2)x_t + C(\psi_1,\psi_2)x_{t-1} + D(\psi_1,\psi_2)z_t \\
&z_t = N(\psi_3)z_{t-1} + \zeta_t, E(\zeta_t\zeta_t^{'}) = \Sigma(\psi_3)
\end{aligned}
\tag{18}
$$

There are four exogenous shocks, so in order to avoid linearly dependent solutions, the likelihood function can only be written in terms of four observable variables. Let $d_t = \{\Delta p_t, y_t, r_t, w_t - p_t\}'$ be the vector of observable variables that we wish to explain. Then, the law of motion of the previous system can be obtained by using standard methods for solving linear rational expectations models, as in Uhlig (1999):

$$
\begin{aligned}
&\begin{bmatrix} x_t \\ z_t \end{bmatrix} = T(\psi)\begin{bmatrix} x_{t-1} \\ z_{t-1} \end{bmatrix} + U(\psi)\zeta_t, E(\zeta_t\zeta_t^{'}) = \Sigma(\psi_3) \\
&d_t = W(\psi)\begin{bmatrix} x_t \\ z_t \end{bmatrix}
\end{aligned}
\tag{19}
$$

such that the solution to the system is written in state-space form. The first equation in (19) is the *transition* equation, while the second equation in (19) is the *measurement* equation. Finally, we can evaluate the likelihood function of the observable data conditional on the parameters $L(\{d_t\}_{t=1}^T \mid \psi)$, by applying the Kalman filter.[14]

D. Drawing from the Posterior and Computing the Bayes Factors

It becomes clear from the previous two subsections that it is not going to be possible to obtain an analytical solution for the posterior distribution. However, since we can evaluate numerically both the prior distribution and the likelihood function, we can apply a numerical algorithm to obtain a draw from the posterior distribution. Geweke (1998) discusses efficient methods to obtain draws from an unknown posterior distribution that can be evaluated numerically. The development of fast speed processors in recent years has removed the time

---

[14] See Hamilton (1994), Ireland (2001) and DeJong, Ingram, and Whiteman (2000) for a detailed explanation.

constraint that many of these methods entail. Not too long ago, only well known, standard distributions could be evaluated or simulated.

In order to obtain a random draw of size N from the posterior distribution, a random walk Markov Chain using the Metropolis-Hastings algorithm is generated. The algorithm is implemented as follows:

1. Start with an initial value (ψ^0). From that value, evaluate the product $L(\{d_t\}_{t=1}^T \mid \psi^0)P(\psi^0)$.
2. For each i:

$$\psi^i = \psi^{i-1} \text{ with probability } 1\text{-}R$$
$$\psi^i = \psi^{i,*} \text{ with probability } R$$

where $\psi^{i,*} = \psi^{i-1} + v^i$, v^i is follows a multivariate Normal distribution, and $R = \min\{1, \frac{L(\{d_t\}_{t=1}^T \mid \psi^{i,*})P(\psi^{i,*})}{L(\{d_t\}_{t=1}^T \mid \psi^{i-1})P(\psi^{i-1})}\}$

The idea for this algorithm is that, regardless of the starting value, more draws will be accepted from the regions of the parameter space where the posterior density is high. At the same time, areas of the posterior support with low density (the tails of the distribution) are less represented, but will eventually be visited. The variance-covariance matrix of v^i is adopted such that the random draw has some desirable time series properties.[15]

Once we have obtained a random draw, for each model it is possible to obtain the marginal likelihood as follows:

$$L(\{d_t\}_{t=1}^T) = \int_{\psi \in \Psi} L(\{d_t\}_{t=1}^T \mid \psi)P(\psi)d\psi \tag{20}$$

The marginal likelihood averages all possible likelihoods across the parameter space, using the prior as a weight. When conveniently weighted, it can be interpreted as the probability of observing the data under a given model. The main shortcoming of the marginal likelihood is that it depends on the priors chosen by the researcher.

The marginal likelihood can be used to compare nested, as well as nonnested models.[16] In the context of this paper, where nested models are compared, it is important to stress that the

[15] See the discussion in Fernández-Villaverde and Rubio-Ramírez (2001).

[16] Rabanal and Rubio-Ramírez (2002) compare nonnested versions of the New Keynesian model. Smets and Wouters (2002) compare a general equilibrium model with statistical (BVAR) models.

marginal likelihood already takes into account that the size of the parameter space for different models can be different. Hence, more complicated models will not necessarily rank better than simpler models, if the extra parameterization is unimportant.[17] This is so because the marginal likelihood visits all the regions of the parameter space, and takes the average of both large and small values of the likelihood function.

Multiple integration is required to compute the marginal likelihood, making the exact calculation impossible. We follow Geweke (1998) to obtain an estimate from the random draw of the posterior distribution. Then, for two different models (1 and 2), the tool to compare them is the Bayes factor:

$$B_{12}(\{d_t\}_{t=1}^{T}) = \frac{\hat{L}(\{d_t\}_{t=1}^{T} \mid \text{model} = 1)}{\hat{L}(\{d_t\}_{t=1}^{T} \mid \text{model} = 2)}.$$

Note that the Bayes factor tells us how we would update our priors on which model is closer to the true one after observing the data. Therefore, the most preferred model when compared pair wise is the one that has the highest marginal likelihood. Moreover, it can be shown that as the sample size increases, $B_{12}(\{d_t\}_{t=1}^{T})$ goes to zero if model 2 is the "best" model.

V. Results

This section presents the results obtained by estimating the model presented in section II. In order to obtain the posterior distributions, a random draw of size 250,000 was obtained using the Metropolis-Hastings algorithm.[18] In the first part of this section, we comment on the mean and the standard deviation of the model's parameters posteriors. In the second part of this section, we comment on the estimated impulse responses to monetary and policy shocks, and discuss under which circumstances the price puzzle would arise.

A. Posterior Distributions of the Parameters

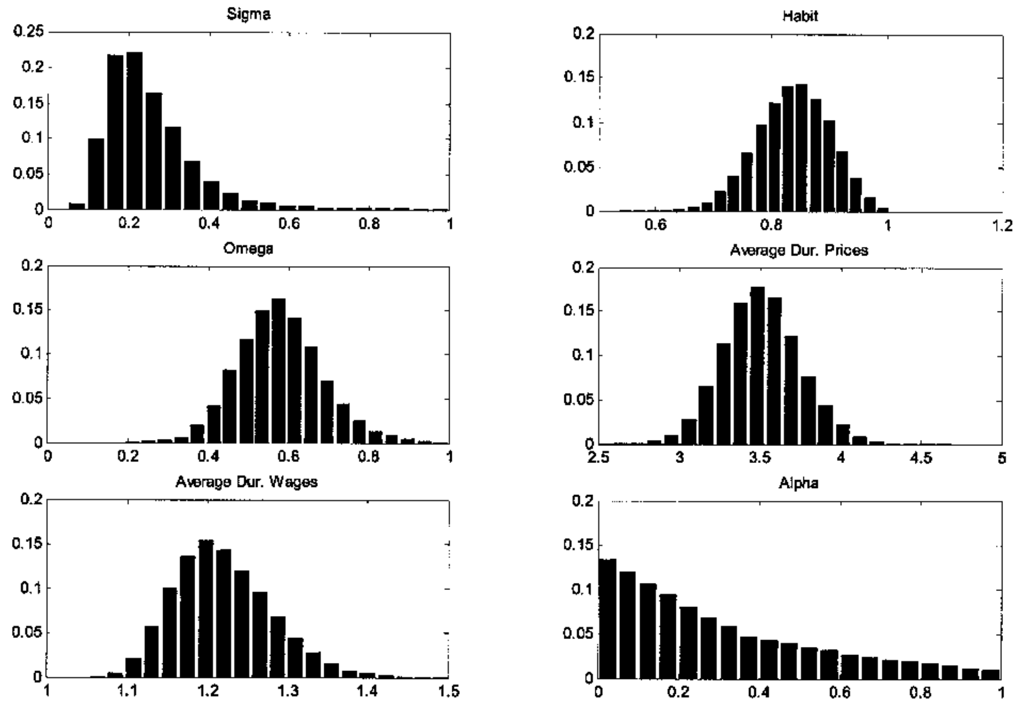
Figure 1 shows the posterior distribution for selected parameters of the model using data for the United States, while Table 3 contains the mean and standard deviation of the posterior distribution of all the model's parameters.

---

[17] In terms of performing maximum likelihood, it is true that the additional parameterization, even if unimportant, cannot deliver a smaller maximum value.

[18] As is common in this type of Markov Chain Monte Carlo methods, the initial 5 percent of draws was discarded, and the variance-covariance matrix of the perturbation term in the algorithm was adjusted such that the acceptance rate lies between 20 and 40 percent.

Figure 1: Posterior Distributions, United States



Source: Author's estimates.

The first two columns of Table 3 show the estimates for the model presented in section 2. The estimates for the inverse elasticity of substitution suggest that it is much smaller than one, which would be the value consistent with a balanced growth path. The estimate we obtain is similar to that of Rabanal and Rubio-Ramírez (2002) and others.[19] The parameter on habit formation in consumption is higher (0.85) than previous values found in the literature. The fact that the we do not incorporate capital accumulation and investment rigidities makes the estimate of this parameter higher than Christiano, Eichenbaum and Evans (2001) or Smets and Wouters (2002), because output persistence in this model only comes from the consumption side.

The proportion of firms that keep prices fixed for a given period delivers a mean posterior average duration of price contracts of between three and four quarters. This value is somewhat higher than previous estimates (usually in the range of two to three quarters), but still implies a reasonable duration of average interval price changes.

[19] See also Basu and Kimball (2000).

Table 3: Posterior Distributions, U.S.

Parameter	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
σ	0.30	0.16	0.33	0.19	0.30	0.11	0.31	0.09
b	0.85	0.06	0.86	0.05	0.88	0.05	0.89	0.04
Average Price Duration	3.51	0.24	3.41	0.22	3.21	0.22	3.12	0.19
Average Wage Duration	1.21	0.06	1.21	0.05	1.00	-	1.00	-
ω	0.58	0.08	0.61	0.10	0.53	0.14	0.55	0.11
α	0.24	0.19	0.00	-	0.22	0.11	0.00	-
γ_p	1.61	0.15	1.58	0.14	1.69	0.18	1.75	0.12
γ_y	0.34	0.06	0.35	0.06	0.34	0.06	0.35	0.05
ρ_r	0.65	0.03	0.65	0.03	0.62	0.03	0.63	0.04
ρ_a	0.55	0.07	0.56	0.06	0.69	0.04	0.69	0.04
ρ_g	0.81	0.04	0.81	0.04	0.82	0.03	0.84	0.03
σ_m	0.20	0.02	0.20	0.02	0.20	0.02	0.20	0.02
σ_p	7.88	1.05	7.58	0.92	6.47	0.76	6.15	0.68
σ_a	1.14	0.23	1.15	0.22	0.57	0.06	0.57	0.06
σ_g	0.67	0.06	0.66	0.06	0.65	0.06	0.66	0.06
$\log(\hat{L})$	943.29		953.02		936.25		952.74	

Source: Author's estimates.

The result for the average duration of wage contracts is surprisingly low, just 1.2 quarters. It is surprising because our prior specification already assumes that wages are set for longer periods of time than prices are. By looking at the expression of κ_w, we can see that the elasticity of substitution between types of labor, φ, interacts greatly with average wage duration. However, for reasonable wage markups, it is not possible to obtain much higher mean wage durations. An explanation of this result can be stated as follows: the average nominal wage series includes all wages in the economy—that is, wages set in multiperiod contracts as well as wages for temporary positions, which are in effect for very short periods of time. Hence, it may well be that the average duration of wage contracts is effectively that low.

The proportion of firms that follow a backward indexing rule is about one half, which means that in the hybrid equation of inflation a weight of one-third is given to the backward-looking inflation part, and two-thirds to expected inflation. The coefficient that effectively measures the cost channel gives us a point estimate of 0.24, with a standard deviation of 0.19. When taking a look at the posterior, we can see that it places more weight on smaller values of α. Hence, given our priors, the data do not seem to be that informative about this parameter. Putting it in another way, the prior and the posterior look quite similar because the likelihood function is relatively flat with respect to the value of α.

The coefficients on the Taylor rule are quite similar to what the literature traditionally assumes, with point estimates of around 1.6 for the reaction of the Taylor rule to deviations of inflation around its steady state value, and a value of around 0.35 for the reaction to deviations in the output gap. The coefficient interest rate smoothing is also high, with a point estimate of 0.65. Finally, the estimates for the shock processes suggest that the autoregressive component of the shocks is lower than usually assumed. This result can be attributed to the fact that, in the 1984–2002 period, the macroeconomic series for the United States exhibit less volatility and persistence than in the 1960–1984 period, and also because habit formation and backward-looking inflation imply significant departures from pure forward-looking behavior, and hence increase the endogenous propagation of exogenous shocks.

In the next two columns of Table 3, the estimates for the constrained model (i.e. $\alpha = 0$) are presented. The estimates for the rest of parameters do not change significantly with respect to the previous case. So the question remains to identify the additional relevance of the cost channel of monetary policy. The log marginal likelihood for the model with cost channel is 953.02, while the log marginal likelihood for the model constrained to $\alpha = 0$ is 943.29. The log Bayes factor in favor of the constrained model is 9.69, a result for which Kass and Raftery (1995) suggest that there is "decisive" evidence for the constrained model vis-à-vis the model with an explicit cost channel. Since we obtain evidence favoring the constrained model, we can conclude that the introduction of the cost channel parameter is unimportant.

Since the sticky wage component is estimated to be fairly, low, we reestimated the model assuming wages are flexible. In that case, we assume that $\theta_w = 0$, which means that the relevant wage setting equation becomes:

$$w_t - p_t = mrs_t = \eta n_t - \lambda_t .$$

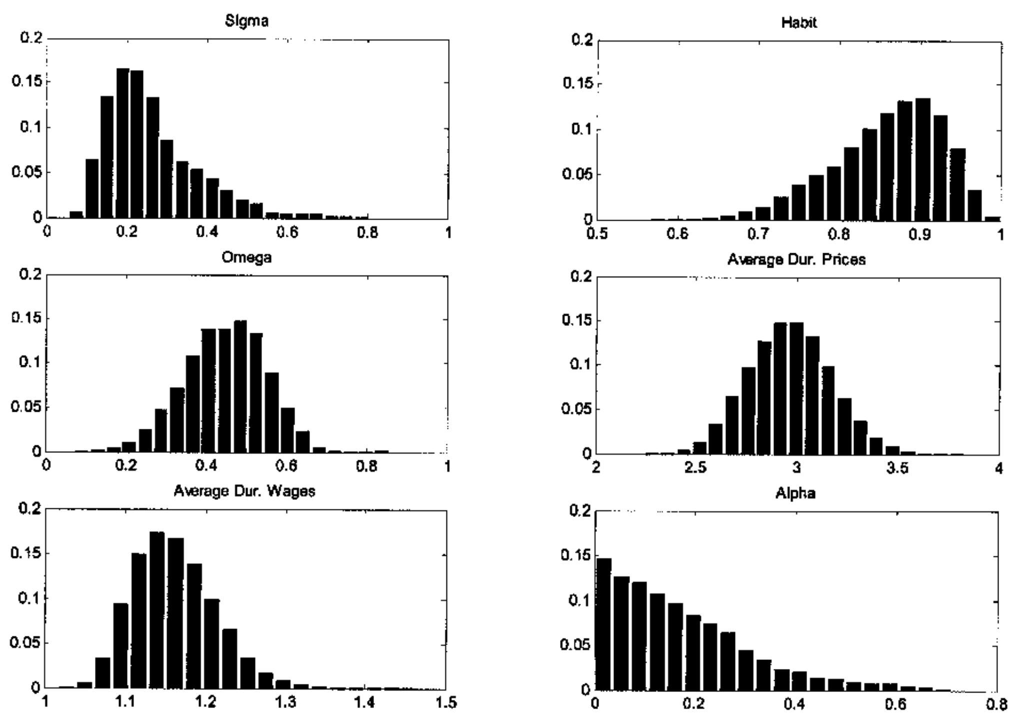
Rabanal and Rubio-Ramírez (2002) found that, in a pure forward-looking model, sticky wages were an important addition to the sticky price model. This was the case because sticky nominal wages **and** sticky prices deliver **sticky real wages**, which is what we observe in the data. However, with habit formation in consumption, real wages are already persistent through the effect of persistent consumption in the marginal rate of substitution.

Columns 5–8 present the analogous estimates to columns 1–4 with flexible wages. The most significant change is that the autoregressive component of the technology shock is estimated to be higher, with a posterior mean of 0.70. This is an indication that once the wage rigidity is removed, more exogenous persistence needs to be fed into the model to match the data. The parameter measuring the cost channel of monetary policy is estimated to be even smaller, with a posterior mean of 0.14. The remaining parameters do not change significantly. The log Bayes factor suggests that the constrained model is also better, when flexible wages are considered. Adding the cost channel simply implies an unimportant overparametrization.

We can also compare the models with sticky and flexible wages using the Bayes factor. The log Bayes factor (with cost channel) is about 7, while the log Bayes factor when the cost channel is eliminated is around 10. According to Kass and Raftery (1995), these magnitudes suggest there is "very strong" evidence of one model over the other. However, we feel that this result should be studied further, since the Bayes factor depends on the prior. It is an important result because it suggests that, once we introduce habit formation in consumption, and backward-looking inflation in the New Keynesian model, we only need small amounts of nominal wage stickiness to match the data.

Figure 2 shows the posterior distribution for selected parameters of the model using data for the euro area, while Table 4 contains the relevant moments of the posterior distributions for all the estimated parameters.

Figure 2: Posterior Distributions, Euro Area



Source: Author's estimates.

Table 4: Posterior Distributions, Euro Area

Parameter	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
σ	0.26	0.11	0.30	0.15	0.37	0.11	0.38	0.13
b	0.86	0.06	0.87	0.06	0.91	0.03	0.91	0.04
Average Price Duration	2.96	0.21	2.92	0.19	2.77	0.20	2.69	0.19
Average Wage Duration	1.16	0.05	1.15	0.05	1.00	-	1.00	-
ω	0.44	0.11	0.46	0.11	0.54	0.13	0.49	0.09
α	0.17	0.13	0.00	-	0.10	0.09	0.00	-
γ_p	1.49	0.14	1.48	0.12	1.54	0.12	1.53	0.17
γ_y	0.34	0.08	0.35	0.08	0.37	0.06	0.36	0.07
ρ_r	0.75	0.03	0.75	0.03	0.74	0.04	0.73	0.03
ρ_a	0.65	0.06	0.67	0.06	0.77	0.04	0.76	0.04
ρ_g	0.77	0.04	0.78	0.04	0.81	0.03	0.82	0.03
σ_m	0.14	0.02	0.14	0.02	0.15	0.02	0.15	0.02
σ_p	6.25	0.88	6.09	0.84	5.67	0.78	5.28	0.68
σ_a	0.98	0.22	0.93	0.21	0.47	0.07	0.47	0.07
σ_g	0.63	0.06	0.64	0.06	0.63	0.06	0.64	0.06
$\log(\hat{L})$	807.19		832.86		812.08		826.14	

Source: Author's estimates.

The parameter on the elasticity of intertemporal substitution is estimated at a value of 0.26, which is in line with what we obtained for the U.S., and the parameter on habit formation in consumption has a posterior mean of 0.86, which is also similar to the U.S. case.

The average duration of price contracts has a posterior mean of three periods, which suggests a smaller degree of stickiness in the euro area as a whole. Similar to the U.S. case, we obtain a surprisingly low average duration of wage contracts, of less than two quarters. A first main difference with respect to the U.S. is that the price indexation parameter is smaller: in this case we obtain a point estimate of 0.44, suggesting that the degree of backward-looking behavior in the euro area is less important. This result is exactly the same that Smets and Wouters (2002) obtain for the euro area.

A second main difference with the estimates previously obtained for the United States is that the point estimate for the cost channel parameter is even lower, with a mean posterior of 0.17. When taking a look at the posterior, we can see that actually a lot of mass is placed close to zero, and almost no mass above 0.5. Hence, in this case we can say that the likelihood function (the data) give overriding evidence that the cost effect channel of monetary policy in Europe is quantitatively very small. A third element of difference with the United States are the coefficients of the Taylor rule. We obtain a smaller coefficient for the reaction of the nominal interest rate with respect to fluctuations in inflation. At the same time, the interest rate smoothing parameter is higher for the euro area.

When reestimating the model without the cost channel parameter, we obtain the expected result that coefficients do not change much. The log marginal likelihood for the model with cost channel is 807.19, while the log marginal likelihood for Bayes factor for the constrained model is 832.06. So, in this case the evidence is even stronger to reject the model with cost channel than in the U.S. case, with the interpretation that we are simply overparametrizing the model.

When we cancel the sticky wage channel and reestimate the model, we obtain higher estimates for the elasticity of intertemporal substitution, and the habit formation parameter. Also, the AR coefficients of the shocks are estimated to be higher. Hence, once again, the lack of endogenous persistence due to sticky wages is substituted by higher exogenous persistence in the shocks. The log Bayes factor with flexible wages (14.06) suggests once again that the cost channel parameter is not important.

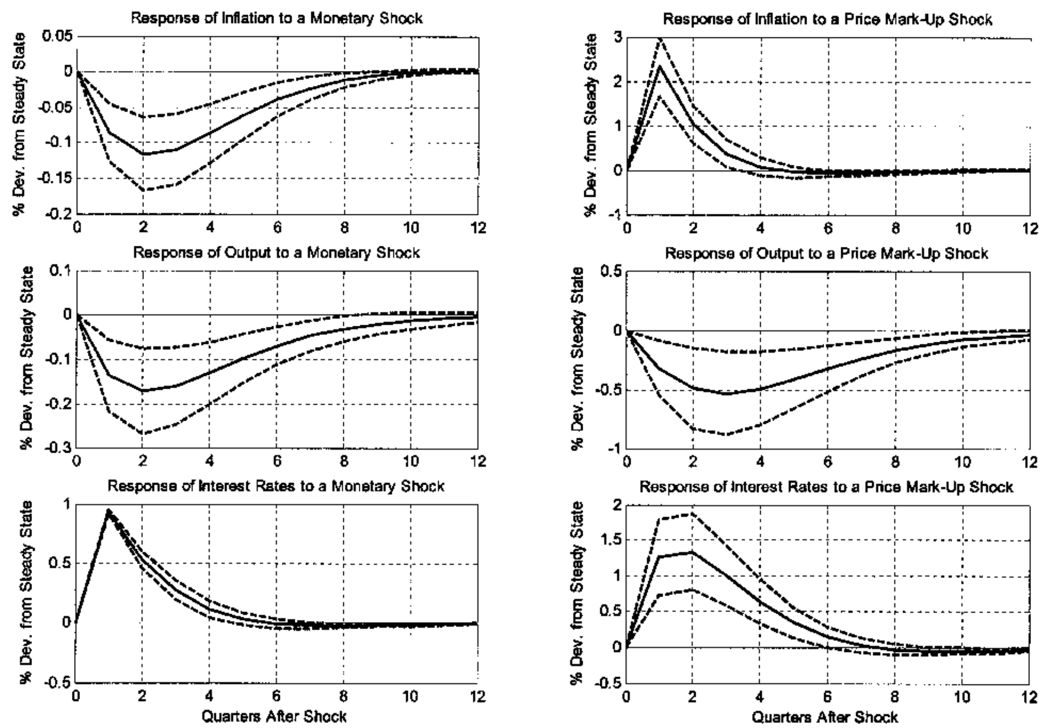
Finally, when we compare the flexible wage model and the staggered wage setting model, we obtain contradictory results: with a cost channel in place, the sticky wage model is rejected. With no cost channel in place, the flexible wage model is rejected. We find it difficult to provide some intuition for this result, but it clearly suggests that more investigation is necessary in order to study the role of sticky wages in the sticky price model, once other rigidities are in place.

To conclude this subsection, we comment on some robustness results. First, we should mention that we experimented with various parameterizations of the coefficients on the price and wage markup. In particular, we experimented with markups of 10 percent for both prices and wages, and the result was only marginal changes in the average duration of price and wage contracts. We also used a range of values for the parameter that governs the elasticity of the labor supply with respect to the desired real wage, η. With values ranging from 0 to 3, we did not find significant changes in the posterior mean estimates.

B. Analysis of Impulse Responses

Figure 3 displays the impulse responses of inflation, interest rates and output to a monetary policy, and a price markup shock using the posterior mean of the vector of parameters, estimated for the United States.[20] The solid line shows the mean posterior impulse response functions, and the dashed line shows two standard deviation posterior bands.

Figure 3: Impulse Responses to Monetary and Price Markup Shocks



Source: Author's estimates.

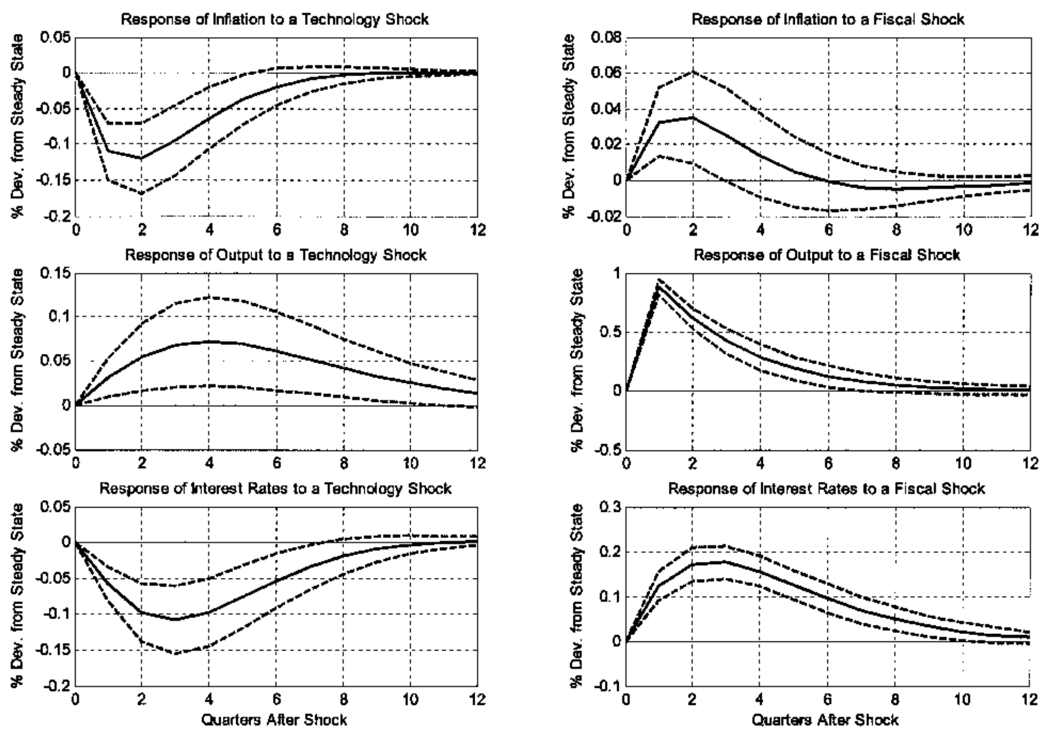
We can observe that, with the estimated parameter values, it is not possible to obtain a price puzzle type of behavior. When the nominal interest rate increases by 100 (annualized) basis points, both inflation and output remain below their steady state values for several quarters. Hence, the estimated value for α is not enough to induce a comovement in the same direction of inflation and interest rates. A price markup shock has a strong positive impact on inflation, as we would expect. The effect is not very long lived, as inflation returns to its steady state value fairly quickly. The shock has a less important quantitative effect on output, but it is however more persistent. Output declines following the monetary policy tightening associated to the increased inflation rate. By comparing the reaction to the two shocks, we can see how it is possible that by confusing or not properly identifying one shock with the

[20] The model is simulated 250,000 times using the random draw from the Metropolis-Hastings algorithm, and the mean and the standard deviation for the posterior impulse responses are computed.

other, we might be attributing cost channel or inflationary effects of monetary policy when they are truly absent.

Figure 4 presents impulse response to technology and fiscal shocks. With a technology shock, inflation decreases, since less units of effective input are needed to produce the same amount of output. Output exhibits the usual behavior in models with nominal rigidities: it stays at the steady state value when the shock hits the economy, and slowly increases over time, exhibiting a hump-shaped response and high persistence. Fiscal shocks have a strong positive effect on output, and a somewhat more moderate effect on inflation.

Figure 4: Impulse Responses to Technology and Fiscal Shocks

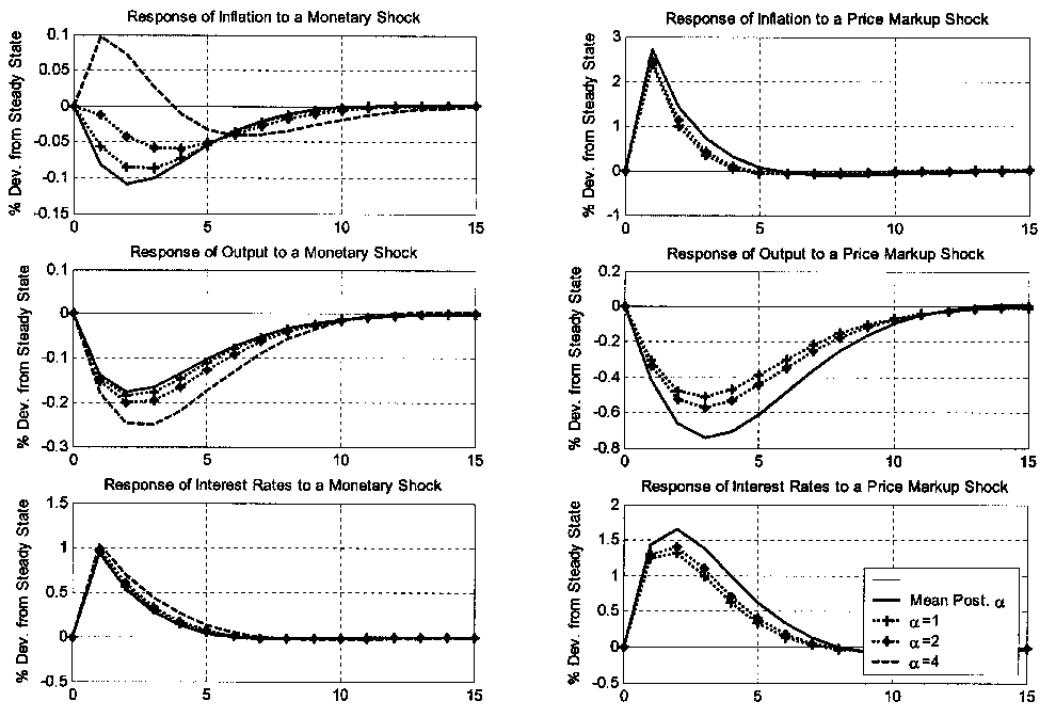


Source: Author's estimates.

C. Generating the Price Puzzle through Cost Channel Effects

After reviewing the impulse responses of the previous subsection, the question arises of under which parameterization, in particular for α, would a cost channel effect of monetary policy arise? Figure 5 shows the result of performing the following exercise: holding all parameters at their estimated mean posterior values, and increasing the value of α. At the estimated mean posterior value, the impulse response exhibits the usual effect: output and inflation remain below their steady state values for several periods. When we assume that all firms are subject to the cost channel effect, $\alpha = 1$, the reaction of inflation is less pronounced but still qualitatively the same.

Figure 5: The Price Puzzle



Source: Author's estimates.

In order to see how far we have to go to generate a significant cost channel effect, we experiment with values of α larger than one. In this context, α can no longer be seen as a proportion of firms. It can be reinterpreted as all firms having to borrow money to pay their wage bill, and α relates the effective interest rate (i_t^{eff}) that firms have to pay to the short-term risk free interest rate:[21]

$$(1+i_t^{eff}) = (1+i_t)^{\alpha}.$$

Under this interpretation, α still represents the structural elasticity of the inflation rate to changes in the nominal interest rate in the New Keynesian Phillips curve. With $\alpha = 2$ (meaning that firms borrow at an interest rate twice that of the Federal Funds rate), it is still

---

[21] Otherwise, we can think that if a fraction $\hat{\alpha}$ of firms need to borrow funds, at a rate that is τ times the Federal Funds rate, then the overall elasticity is $\alpha = \hat{\alpha}\tau$. Of course, it is not possible to estimate those parameters separately, but it is another way to justify an α greater than one.

not possible to generate a cost channel effect. With higher values, such as $\alpha = 4$, we finally obtain the result that inflation and interest rates move in the same direction. However, this value seems to be quite high, and our estimation procedure does not get close to it.[22] In fact, as we showed in a previous subsection, the estimated value both for the United States and the euro area is close to zero.

VI. Concluding Remarks

In this paper we have estimated a small-scale macroeconomic model that accounts for the joint behavior of output, real wages, interest rates, and inflation. We have incorporated an explicit cost channel of monetary policy, whereby increases in the nominal interest rate directly increase the inflation rate. Using a Bayesian framework, we have estimated that particular elasticity for the United States and for the euro area. We have found in both cases that it is not quantitatively significant. Hence, we have to conclude that the cost channel is not present in aggregate data and not relevant for monetary policymaking.

At the same time, we have obtained plausible estimates for the model's parameters, similar to what other researchers have found in the literature using other econometric and modeling approaches. If anything, we should mention that the estimate of average duration of wage contracts is surprisingly low. A novel result is that, once we take into account habit formation in consumption and backward-looking behavior in inflation, staggered wage setting might not be that important after all. However, we feel that this result deserves further study.

Barth and Ramey (2001) suggest that the cost channel does occur when using industry level data. However, their estimates point at stronger cost channel effects in the pre-1979 period. Hence, the fact that we do not detect those cost channel effects in the post-1984 period should not be viewed as a direct contradiction of their results. An interesting way to reconcile the findings of this paper and theirs would be to estimate a dynamic general equilibrium model using Bayesian methods and sectoral data.

[22] Ravenna and Walsh (2003) estimate values for α of up to 11. Under this interpretation, if the U.S. Federal Funds rate was around 6 percent by end 2001, firms would be borrowing at an effective rate of 66 percent, which by any standards seems far too high.

REFERENCES

Barth, M. and V. Ramey, 2001, "The Cost Channel of Monetary Transmission," in *NBER Macroeconomics Annual 2001,* ed. by Ben S. Bernanke and Kenneth Rogoff (Cambridge, Massachussets: National Bureau of Economic Research).

Basu, S. and M. Kimball, 2000, "Long-Run Labor Supply and the Elasticity of Intertemporal Substitution," (unpublished; Ann Arbor, Michigan: University of Michigan).

Blanchard, O. and N. Kiyotaki, 1987, "Monopolistic Competition and the Effects of Aggregate Demand," *American Economic Review*, Vol. 77, No. 4, pp. 647–666.

Boivin, J. and M. Giannoni, 2003, "Has Monetary Policy Become More Effective?," NBER Working Paper No. 9459 (Cambridge, Massachussets: National Bureau of Economic Research).

Boldrin, M., L. Christiano, and J. Fischer, 2001, "Habit Persistence, Assets Returns and the Business Cycle," *American Economic Review,* Vol. 91, No. 1, pp. 149–166.

Calvo, G., 1983, "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics,* Vol. 12, No. 3, pp 383–398.

Clarida, R., J. Galí, and M. Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, Vol. 115, No. 1, pp. 147–180.

Christiano, L., M. Eichenbaum, and C. Evans, 2001, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," NBER Working Paper No. 8403 (Cambridge, Massachussets: National Bureau of Economic Research).

DeJong, D.N., B.F. Ingram, and C.H. Whiteman, 2000, "A Bayesian Approach to Dynamic Macroeconomics," *Journal of Econometrics*, Vol. 98, No. 2, pp. 203–223.

Erceg, C., D. Henderson, and A. Levin, 2001, "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics*, Vol. 46, No. 2, pp. 281–313.

Fagan, G., J. Henry, and R. Mestre, 2001, "An Area-Wide Model (AWM) for the Euro Area," European Central Bank Working Paper No. 42 (Frankfurt: European Central Bank).

Fernández-Villaverde, J., and J.F. Rubio-Ramírez, 2001, "Comparing Dynamic Equilibrium Models to Data," PIER Working Paper 01–037 (Philadelphia, Pennsylvania: University of Pennsylvania).

Galí, J. and M. Gertler, 1999, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, Vol. 44, No. 2, pp. 195–222.

Geweke, J., 1998, "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," Federal Reserve Bank of Minneapolis Staff Report 249 (Minneapolis, Minnesota: Federal Reserve of Minneapolis).

Giannoni, M., 2001, "Robust Optimal Monetary Policy in a Forward-Looking Model with Parameter and Shock Uncertainty," (unpublished; New York: Federal Reserve Bank of New York).

Hamilton, J., 1994, *Time Series Analysis* (Princeton, New Jersey: Princeton University Press).

Ireland, P., 2001, "Sticky-Price Models of the Business Cycle: Specification and Stability," *Journal of Monetary Economics*, Vol. 47, No. 1, pp. 3–18.

Kass, R. and A. Raftery, 1995, "Bayes Factors," *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 773–795.

Kim, J., 2000, "Constructing and Estimating a Realistic Optimizing Model of Monetary Policy," *Journal of Monetary Economics*, Vol. 45, No. 2, pp. 329–359.

Rabanal, P. and J.F. Rubio-Ramírez, 2002, "Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach," revised version of Federal Reserve Bank of Atlanta Working Paper 2001–22a (Atlanta, Georgia: Federal Reserve Bank of Atlanta).

Ravenna, F. and C. Walsh, 2003, "The Cost Channel in a New Keynesian Model: Evidence and Implications," (unpublished; Santa Cruz, California: University of California).

Rotemberg, J. and M. Woodford, 1997, "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," in *NBER Macroeconomics Annual 1997*, ed. by Ben S. Bernanke and Julio Rotemberg (Cambridge, Massachussets: National Bureau of Economic Research).

Sbordone, A., 2001, "An Optimizing Model of U.S. Wage and Price Dynamics," (unpublished; New Brunswick, New Jersey: Rutgers University).

Sims, C., 1992, "Interpreting the Macroeconomic Time Series Facts: the Effects of Monetary Policy," *European Economic Review*, Vol. 36, No. 5, pp. 975–1000.

Smets, F. and R. Wouters, 2002, "An Estimated Stochastic Dynamic General Equilibrium Model for the Euro Area," European Central Bank Working Paper No. 171 (Frankfurt: European Central Bank).

Taylor, J., 1993, "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Series on Public Policy,* Vol. 39, pp. 195–214.

Taylor, J., 1999, "Staggered Price and Wage-Setting in Macroeconomics," in *Handbook of Macroeconomics*, ed. by John Taylor and Michael Woodford (Amsterdam: Elsevier Science).

Uhlig, H., 1999, "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon and A. Scott (Oxford, United Kingdom: Oxford University Press).