Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Monetary Economics 52 (2005) 1183–1197
Journal of MONETARY ECONOMICS
www.elsevier.com/locate/jme

Do expected future marginal costs drive inflation dynamics?

Argia M. Sbordone*,[1]

Argia M. Sbordone, Domestic Research, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045, USA

Available online 27 September 2005

Abstract

This article discusses a more general interpretation of the two-step minimum distance estimation procedure proposed in Sbordone (2002). The estimator is again applied to a version of the New Keynesian Phillips curve, where inflation dynamics are driven by the expected evolution of marginal costs. The article clarifies econometric issues, addresses concerns about uncertainty and model misspecification raised in recent studies, and assesses the robustness of previous results. While confirming the importance of forward-looking terms in accounting for inflation dynamics, it suggests how the methodology can be applied to extend the analysis of inflation to a multivariate setting.

JEL classification: E31; E32

Keywords: Inflation; New Keynesian pricing; Marginal costs

1. Introduction

The standard pricing assumption in real business cycle models implies a constant markup of prices over marginal cost, and hence an inflation rate equal to the rate of growth of average nominal marginal cost. These predictions are at odd with the data: in particular, US inflation is less volatile than marginal costs. However, by introducing

*Tel.: (212) 720 6810; fax: (212) 720 1844.
E-mail address: argia.sbordone@ny.frb.org.
[1]The opinions expressed are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System.

0304-3932/$ - see front matter
doi:10.1016/j.jmoneco.2005.08.010

nominal price rigidities it is possible to explain cyclical markup variations, and hence to generate an inflation path whose volatility is like that observed in the data.

The widely used Calvo model of staggered pricing (Calvo, 1983) implies an equilibrium pricing condition that, in log-linearized form, links current inflation to expected future inflation and current real marginal cost[2]

$$\pi_t = \beta \mathrm{E}_t \pi_{t+1} + \zeta s_t + \eta_t. \tag{1.1}$$

Here s_t is the (log of) average real marginal cost in the economy, the parameter β is a discount factor, and ζ is a nonlinear function of the relevant structural parameters: $\zeta = (1-\alpha)(1-\alpha\beta)/\alpha(1+\theta\omega)$. θ is the elasticity of substitution among differentiated goods, ω is the elasticity of firms' marginal costs to their own output,[3] and α is the percentage of prices that are not reset optimally at time t. The degree of price inertia is measured by $1/1-\alpha$.[4] The error term η_t is included to account for fluctuations in the desired mark-up, or for other forms of misspecification of the equation;[5] throughout this article it is assumed to be a mean zero, serially uncorrelated stochastic process.[6]

This model has been generalized in a number of ways to be able to generate additional inflation inertia. Here I follow Christiano et al. (2005) by assuming that firms that are not selected to reset prices through the Calvo random drawing are nonetheless allowed to index their current price to past inflation, and I assume that they do so by some fraction $\varrho \in [0, 1]$. The solution of the model in this case[7] is

$$\pi_t - \varrho\pi_{t-1} = \beta(\mathrm{E}_t\pi_{t+1} - \varrho\pi_t) + \zeta s_t + \eta_t, \tag{1.2}$$

which nests Eq. (1.1) (the case of $\varrho = 0$), and, in the opposite case of full indexation ($\varrho = 1$), as considered in Christiano et al. (2005), implies an expectational equation in the *rate of growth* of inflation. This generalized equation has the same form as the 'hybrid model' of Gali and Gertler (1999), when rewritten as

$$\pi_t = \frac{\varrho}{1+\beta\varrho}\pi_{t-1} + \frac{\beta}{1+\beta\varrho}\mathrm{E}_t\pi_{t+1} + \frac{\zeta}{1+\beta\varrho}s_t + \widetilde{\eta}_t, \tag{1.3}$$

or

$$\pi_t = \gamma^b\pi_{t-1} + \gamma^f\mathrm{E}_t\pi_{t+1} + \overline{\zeta}s_t + \widetilde{\eta}_t. \tag{1.4}$$

In this expression γ^b and γ^f can be interpreted as the weights, respectively, on 'backward-' and 'forward-looking' components of inflation. Iterating forward, Eq. (1.2) gives a present

[2]A detailed derivation of this equation can be found in Woodford (2002, ch. 3).

[3]The presence of this term is due to the further assumption of firm-specific capital. This term alters the mapping between the parameter ζ and the frequency of price adjustment, as discussed in Sbordone (2002), making a low estimate of ζ consistent with a reasonable degrees of price stickiness.

[4]The variables are expressed in log deviation from steady state values. If the log-linearization is around a zero steady state inflation, the log deviation of inflation can be measured by its actual value. Under the assumption that real wage and productivity share the same long run trend, the log deviation of the labor share can also be measured by its actual value. In the data, we will see below that stationarity may require a slight transformation of the share.

[5]This was suggested by Rotemberg and Woodford (1999). In Steinsson (2002) the error represents exogenous variations in the elasticity of substitution; in Giannoni (2000) it represents time varying tax distorsions.

[6]In my (2002) paper, I examined the degree to which the data could be fit by a model with no error term. Here, instead, an explicit hypothesis about the nature of the error term allows to address various issues such as a possible simultaneous-equations bias.

[7]A detailed derivation of this expression can be found in Woodford (2003, ch.3).

value relationship, where inflation is a function of lagged inflation and expected future real marginal costs:

$$\pi_t = \varrho\pi_{t-1} + \zeta \sum_{j=0}^{\infty} \beta^j \mathrm{E}_t s_{t+j} + v_t. \tag{1.5}$$

The empirical evaluation of Eq. (1.1), or its generalized form (1.2), known as the New Keynesian Phillips curve (NKPC), has generated a great deal of debate, as the papers presented in this volume testify. Gali and Gertler (1999) pioneered an approach to estimation based on the Euler equation (1.4), which raised a lot of discussion about the appropriateness of the use of GMM estimation. Gali, Gertler and Lopez-Salido respond in this issue to most of the criticisms to their approach.

Sbordone (2002) proposed an alternative two-step procedure based on the empirical evaluation of the closed form solution (1.5) in its restricted form ($\varrho = 0$), and in this paper I wish to clarify this methodology, assess the robustness of my previous results, and evaluate some of the criticisms raised in the other articles in this volume.

2. Estimating the closed-form solution

My (2002) paper proposed to estimate the basic NKPC specification, Eq. (1.1), by matching actual inflation dynamics to the inflation path predicted by the Calvo model, taking as given the dynamics of *nominal* marginal cost, denoted here as mc_t. I assumed that the model held exactly ($\eta_t = 0$), and solved the model forward to obtain a predicted path of prices as function of expected future nominal marginal cost:

$$p_t = \lambda_1 p_{t-1} + (1-\lambda_1)(1-\lambda_2^{-1}) \sum_{j=0}^{\infty} \lambda_2^{-j} \mathrm{E}_t mc_{t+j}. \tag{2.1}$$

The parameters λ_1 and λ_2 in (2.1), roots of the characteristic equation $P(\lambda) = \beta\lambda^2 - (1+\beta+\zeta)\lambda + 1 = 0$, are non linear combinations of the structural parameters β and ζ. The proposal was to evaluate this pricing model along the lines of Campbell and Shiller's (1987) evaluation of present value relationships in finance.

For such evaluation I first assumed that appropriate conditions held to guarantee proportionality of average and marginal costs: the unobservable marginal cost could then appropriately be proxied by a measure of unit labor cost. Furthermore, to express the relationship in terms of stationary variables, I transformed (2.1) into a relation between the price/unit labor cost ratio and the rate of growth of unit labor costs (respectively $p_t - u_t$ and Δu_t, with variables expressed in logs). The empirical evaluation was in two steps: first the estimation of an unrestricted vector autoregression model to forecast unit labor costs; then, taking as given this forecast, the estimation of the parameters of the structural model by minimizing the distance between the path of the price/ulc ratio implied by the model and the actual dynamic path of the data (given the path of unit labor cost, a predicted path for the price/ulc ratio then implies a path for inflation as well).

More specifically, assuming that all information at time t about current and future values of the rate of growth of unit labor cost Δu_t could be summarized by a vector of variables Z_t, where $\{Z_t\}$ is a stationary Markov process,

$$Z_t = AZ_{t-1} + G\varepsilon_{zt}, \tag{2.2}$$

and $E_{t-1}(\varepsilon_{zt+j}) = 0$, for all $j \geqslant 0$,[8] the infinite sum of future expected unit labor cost could be computed as[9]

$$\sum_{j=0}^{\infty} \lambda_2^{-j} E_t \Delta u_{t+j} = \sum_{j=0}^{\infty} e_u' \lambda_2^{-j} A^j Z_t = e_u'(I - \lambda_2^{-j} A)^{-1} Z_t.$$

Letting $B = (I - \lambda_2^{-j} A)^{-1}$, and noting that the price/unit labor cost ratio is the inverse of the labor share (so that $p_t - u_t = -s_t$), the solution (2.1) could be written as

$$s_t = \lambda_1 s_{t-1} + \Delta u_t - (1 - \lambda_1) e_u' B Z_t - \eta_t. \tag{2.3}$$

Denoting by $s_t^m(\psi, A)$ the path of the labor share predicted by the model, and by $\varepsilon_t^s = s_t - s_t^m(\psi, A)$ the distance between actual and predicted paths, under the null that the model is true we have that

$$E(\varepsilon_t^s) = E(s_t - s_t^m(\psi_0, A_0)) = 0, \tag{2.4}$$

where ψ_0, A_0 denote true parameter values. With this notation, the proposed two-step estimator involved first the estimation of the system (2.2), and then the estimation of the vector $\psi = (\beta, \zeta)'$ by

$$\widehat{\psi}_1 = \arg\min \, var(\widehat{\varepsilon}_t^s) \tag{2.5}$$

where $\widehat{\varepsilon}_t^s = s_t - s_t^m(\psi, \widehat{A})$, and $\widehat{A}$ is a consistent estimator of the elements of A. The theoretical (or fundamental) inflation rate was then derived as

$$\pi_t^m = -(s_t^m(\widehat{\psi}_1, \widehat{A}) - s_{t-1}^m(\widehat{\psi}_1, \widehat{A})) + \Delta u_t.$$

This approach to estimating inflation dynamics provided, I believe, an approach to the empirical assessment of Phillips curve relationships which was novel in two respects. First, it focused on the relationship between the dynamics of prices and the dynamics of marginal costs, as opposed to the relationships between inflation and output gap. This choice was motivated by the observation that the Calvo model of optimizing firms with staggered prices makes predictions *only* about the dynamic relation between prices and marginal cost. In order to get an empirical Phillips curve specification in terms of output gap one needs further theoretical assumptions, both about how marginal costs are related to output, and about how to construct a theory-based measure of potential output.[10] The choice of marginal cost as forcing variable was at the same time independently made by Gali and Gertler (1999), who similarly proxied marginal costs with unit labor costs.

The second novelty was in the estimation procedure. The paper focused on the estimation of the present value relationship between prices and marginal costs implied by the optimizing model, and applied a two-step estimation procedure. As described above, following Campbell and Shiller's tests of the present value theory of stock price determination, the first step involved estimating an auxiliary forecasting model to generate predictions of the future values of the forcing variable—the growth of nominal marginal

[8]Note that any autoregressive process of order k can be expressed in this form through a suitable definition of the vector Z_t and the matrices A and G.

[9]The vector e_x denotes a selection vector for variable x (a unit vector with 1 in the position corresponding to x and zero otherwise).

[10]One can obviously simply interpret marginal costs as a particular measure of output gap: given the uncertainty in the estimation of output gap, and the difficulties of constructing a truly theory-based measure (this is attempted, however, by Nelson and Neiss, 2005) this is a convenient measure.

costs in my application. The second step involved estimating the parameters of the structural model, conditional on the forecasting model estimated in the first step, using a distance estimator.

The results were quite striking. For Eq. (1.1), calibrating the discount parameter β, the estimate of the coefficient ζ was positive and statistically significant. Its value was consistent with price rigidity lasting 3 to 4 quarters, in line with survey-based evidence. Moreover, the dynamics of predicted inflation were very close to the actual inflation dynamics, and the model allowed to reproduce quite closely the serial correlation of the data.

These results depended of course upon the correct specification both of the structural model and of the auxiliary forecasting system. The (2002) paper considered only the purely forward-looking specification of the structural model (the case of $\varrho = 0$), but checked the sensitivity of the results to two modifications of the measure of marginal cost, the presence of adjustment costs for labor, and the case of a technology with overhead labor. Both modifications altered the specification of the structural model by adding other forcing variables in the inflation equation.

To address the second problem, I considered alternative forecasting systems for unit labor costs, both excluding the price/unit labor cost from the system, and including additional variables; in all cases the qualitative results of the model remained valid. Finally, I showed that the ability of the model to track inflation dynamics was worsened when excluding the forward looking terms in (2.1), and concluded that this component appears to be important for explaining the dynamics of prices.[11]

Both specification issues receive a more critical assessment in the contributions of this volume, and in this article I take the opportunity to comment further on them. While I leave the issues related to the robustness of GMM estimates to the reappraisal by Gali et al. (2005), I try instead to offer some perspective on the issue of uncertainty raised by Kurmann (2005), which hinges upon the specification of the auxiliary forecasting model; on the issue of whether the forward-looking component in the inflation dynamics is insignificant, as claimed by Rudd and Whelan (2005); and on the issue of what should be the appropriate proxy for marginal cost, as discussed by Batini et al. (2005).

3. Appraisal of the criticisms

As the summary of my approach in the previous section shows, the analysis in my 2002 paper treated *nominal* marginal costs as the forcing variable in a model of price dynamics. Most of the subsequent literature on the NKPC, and all the papers in this issue of the *JME*, take instead *real* marginal cost as the forcing variable, and estimate directly an inflation equation of the form (1.1) or (1.5). Since the choice of the most appropriate forcing variable is not the focus of this article, I will conduct my discussion in terms of this latter version of the model. The application of the methodology discussed above to this case involves defining a distance function directly in terms of inflation paths $\varepsilon_t^\pi = \pi_t - \pi_t^m(\psi)$, and its minimization is conditional on a forecasting model for *real*

[11] My conclusions were in line with those reached by Gali and Gertler, who estimated the model parameters on a specification of the type (1.4), which explicitly includes backward and forward-looking inflation terms.

marginal cost. In what follows the forecasting model has the same form (2.2), and the vector Z_t includes at least inflation and a measure of real marginal cost (the labor share).

3.1. Kurmann's critique

Kurmann's (2005) paper analyzes the robustness of fit of the New Keynesian model to alternative specifications of the forecasting VAR, and alternative values for the weight of the forward-looking component. His criticism is directed to the fit of what Gali and Gertler call 'fundamental inflation', which is the inflation path derived from the solution of the Calvo model. To construct this path they use the coefficients estimated for the Euler equation (1.4) by a standard GMM procedure, and the forecast of real marginal costs implied by a separately estimated vector autoregression model.

Kurmann constructs the path of inflation from the restricted form of (1.5),

$$\pi_t = \zeta \sum_{j=0}^{\infty} \beta^j \mathrm{E}_t s_{t+j}, \tag{3.1}$$

setting $\zeta = .035$ (a 'benchmark' value among those estimated by Gali and Gertler, 1999), and $\beta = 1$. He constructs the forecasts s_{t+j} from the VAR model estimated by Gali and Gertler, a bivariate model with four lags of π_t and s_t. The goodness of fit is measured by the relative standard deviation of predicted vs. actual inflation, $\sigma_{\pi^m}/\sigma_\pi$, and the correlation between predicted and actual inflation $\rho(\pi^m, \pi)$.

The criticism raised by Kurmann is that while the point estimates of these statistics indicate an impressive fit to both the dynamic path and the volatility of inflation, there is a lot of uncertainty surrounding them.

First, he argues that, even assuming that the forecasting model is correct, in the sense of containing all the variables that help to forecast the expected future value of the labor share, the model is merely estimated, and treating the estimated parameters as true population values leads to underestimate the uncertainty surrounding the estimated inflation statistics. Specifically, he shows that the confidence interval around the two estimated statistics is quite large, due to the uncertainty of the estimated VAR coefficients. Second, he shows that the point estimates themselves are highly sensitive to the specification of the forecasting model that one chooses. Finally, he discusses the sensitivity of the predicted inflation dynamics to the degree of price stickiness implied by the assumed value of the coefficient ζ.

Kurmann's paper aims at showing that the evidence provided by graphs of the fundamental inflation and by point estimates of standard deviation and correlation statistics is misleading, because it hides the uncertainty of the estimated forecasting model that is used to construct the predicted path of inflation.

The question of the uncertainty in the VAR estimation is particularly relevant to the estimation method discussed above, since it uses the auxiliary VAR not just for the construction of the model's predicted path, but also as a crucial step of the estimation procedure. Next section addresses therefore the issue of uncertainty by revisiting the two-step estimation procedure in a way that shows how to take into account the imprecision of the first step estimation. At the same time it tries to clarify the relation between this approach and the instrumental variables approach used by Gali and Gertler (1999) to estimate the Euler equation of the Calvo model.

3.2. The distance estimator reinterpreted

As noted above, my proposed two-step distance estimator was based on Campbell and Shiller's procedure. This analogy can perhaps be better illustrated by giving a slightly different interpretation to the distance ε_t^π, namely by viewing it as a measure of the *restrictions* imposed by the structural model on the parameters of the forecasting process.[12] For this interpretation one should observe that, by definition, the vector of forecasting variables Z_t includes current inflation and the labor share, so that we can write, with an appropriate definition of selection vectors e_π and e_s,

$$\pi_t = e_\pi' Z_t, \quad \text{and} \quad s_t = e_s' Z_t. \tag{3.2}$$

Then, using (2.2), the infinite sum of expected future values of the labor share that appears in the solution (3.1) is computed as

$$\sum_{j=0}^{\infty} \beta^j \mathrm{E}_t s_{t+j} = e_s'(I - \beta A)^{-1} Z_t,$$

so that the solution (3.1) can be written as

$$e_\pi' Z_t = \zeta e_s'(I - \beta A)^{-1} Z_t.$$

Under the null that the model is a good representation of the data, this equality must hold for every Z_t; hence it must be true that

$$e_\pi' - \zeta e_s'(I - \beta A)^{-1} = 0. \tag{3.3}$$

This expression defines a $(1 \times 2p)$ vector of restrictions on the elements of the matrix A that characterizes the process (2.2). These cross-equation restrictions between the parameters of the structural model and those of the driving process Z_t represent, in the words of Hansen and Sargent (1980), the 'hallmark' of rational expectations models. For the present value model of stock prices, with no free parameters, Campbell and Shiller proposed a Wald test of these restrictions. The distance estimator that I proposed can be interpreted as an 'unweighted' measure of these restrictions. Denoting the restrictions (3.3) as a vector function $F(\psi, A)$, for a consistent estimate $\widehat{A}$ of the matrix A, it is the case that

$$F(\psi, \widehat{A})' \equiv e_\pi' - \zeta e_s'(I - \beta \widehat{A})^{-1} \tag{3.4}$$

converges to 0, and the proposed estimator $\widehat{\psi}$ is the vector that minimizes the square of the function $F(\psi, \widehat{A})$. Under this interpretation, one may also modify the proposed estimator to minimize a 'weighted' function of the restrictions, by giving a higher weight in the objective function, for example, to those elements of $\widehat{A}$ which are estimated more precisely. This can be done by weighting the quadratic function with the covariance matrix of the restricted parameters $\widehat{A}$. The weighted estimator is then defined as

$$\widehat{\psi}_2 = \arg\min[F(\psi, \widehat{A})' \Sigma_A^{-1} F(\psi, \widehat{A})] \tag{3.5}$$

where Σ_A is a matrix with appropriately selected elements of the estimated variance–covariance matrix of $\widehat{A}$.

To summarize, my proposed approach to estimate the present value form of the Calvo model of inflation dynamics is a two-step distance estimator that exploits an 'auxiliary'

[12] The estimator in this form is applied to a two-variable model of price and wage dynamics in Sbordone (2003).

autoregressive representation of the data. The estimator may take two forms. In (2.5) the objective function to minimize is the variance of the distance between model and data, which is an unweighted quadratic form of this distance, while in (3.5), the objective function is similarly a (possibly weighted) quadratic form of a distance function representing the restrictions that the model solution imposes on the parameters of the auxiliary VAR.

The first interpretation emphasizes the role of the auxiliary VAR process as a forecasting process from which to compute the expected future values of the forcing variables. In the second interpretation, the VAR provides an *unrestricted* representation of the data, against which to compare the restrictions imposed by the structural model.

The analogy has thus far been illustrated for the case in which the Calvo model holds exactly. More generally, when the inflation equation includes an error term, as in specification (1.1) and, further, when it also includes a term in lagged inflation, as in specification (1.2), the model solution is

$$\pi_t^m = \varrho \pi_{t-1} + \zeta e_s'(I - \beta A)^{-1} Z_t + \beta \eta_t.$$

In this more general case the vector of structural parameters is redefined as $\psi = (\varrho, \beta, \zeta)'$, and minimizing the distance function $\varepsilon_t^\pi(\psi)$ requires some assumption about the stochastic term η_t. If one assumes that $\mathrm{E}(\eta_t | Z_{t-1}) = 0$, and furthermore that η_t is serially uncorrelated, the estimator $\widehat{\psi}_1$ in (2.5) can be redefined by replacing the moment condition analogous to (2.4) with a conditional expectation

$$\mathrm{E}(\varepsilon_t^\pi | Z_{t-1}) = \mathrm{E}(\pi_t - \pi_t^m(\psi_0, A_0) | Z_{t-1}) = 0. \tag{3.6}$$

Using the auxiliary VAR to construct the projection of π_t and Z_t on Z_{t-1}, (3.6) becomes

$$e_\pi' A Z_{t-1} - \varrho e_\pi' Z_{t-1} - \zeta e_s'(I - \beta A)^{-1} A Z_{t-1} = 0, \tag{3.7}$$

and one can then define a minimum distance estimator for ψ as in (2.5), with the appropriate redefinition of $\widehat{\varepsilon}_t^\pi$. Similarly, the estimator $\widehat{\psi}_2$ in (3.5) would be based on an analogous redefinition of the function F, which is now given by the orthogonality conditions (3.7). Since these conditions must hold for every Z_{t-1}, it must be the case that

$$e_\pi' A - \varrho e_\pi' - \zeta e_s'(I - \beta A)^{-1} A = 0. \tag{3.8}$$

The function F in (3.4) is then replaced by the left hand side of (3.8), with A replaced by its consistent estimate $\widehat{A}$, and the estimator of ψ is again defined as $\widehat{\psi}_2$ in (3.5).

3.3. Relation with the GMM approach

Gali and Gertler (1999) estimate the baseline inflation model of (1.1) with a seemingly different empirical procedure. Instead of estimating the closed form solution of the model (as discussed here), they define the error in expectations

$$v_{t+1} = \pi_{t+1} - \mathrm{E}_t \pi_{t+1}$$

and, substituting actual for expected value of future inflation in the model, obtain

$$\pi_t = \beta(\pi_{t+1} - v_{t+1}) + \zeta s_t,$$

or

$$\beta v_{t+1} = \beta \pi_{t+1} - \pi_t + \zeta s_t.$$

From the definition of rational expectations, the surprise in inflation at $t+1$ is unforecastable given the information set at time t, I_t: $\mathrm{E}(v_{t+1}|I_t) = 0$, or

$$\mathrm{E}[(\pi_t - \beta\pi_{t+1} - \zeta s_t)|I_t] = 0. \tag{3.9}$$

Gali and Gertler's estimation of the parameter vector ψ exploits this orthogonality condition in a traditional GMM context. They observe that the orthogonality condition implies that any vector of variables X_{t-j} which is in the information set I_t should be *uncorrelated* with the expectational error: this implies a set of moment conditions based on the *unconditional covariance* of v_{t+1} and X_{t-j}. They therefore define a vector function

$$H(\psi, w_t) = (\pi_t - \beta\pi_{t+1} - \zeta s_t)X_{t-j},$$

where $w_t = (\pi_t, \pi_{t+1}, s_t, X_{t-j})'$, and use the orthogonality conditions $\mathrm{E}(H(\psi, w_t)) = 0$ for estimation. They then proceed with textbook GMM estimation: given T observations on the vector of variables w_t, the parameter vector ψ is estimated as the vector that minimizes the sample equivalent of the orthogonality conditions, for an appropriate weighting matrix.

Now suppose that $X_{t-j} = Z_{t-1}$; this amounts to choosing as instruments the variables that optimally forecast Z_t. Then it is easy to see the relationship between this estimator and the distance estimators proposed above. Taking conditional expectation of (3.9) one gets

$$\mathrm{E}(\pi_t|Z_{t-1}) - \beta\mathrm{E}(\pi_{t+1}|Z_{t-1}) - \zeta\mathrm{E}(s_t|Z_{t-1}) = 0,$$

which, using the auxiliary VAR to compute the projections, gives

$$e'_\pi AZ_{t-1} - \beta e'_\pi A^2 Z_{t-1} - \zeta e'_s AZ_{t-1} = 0.$$

Hence we have

$$e'_\pi A - \beta e'_\pi A^2 - \zeta e'_s A = 0. \tag{3.10}$$

A distance estimator of the kind I proposed, but based on restrictions (3.10), would be exploiting similar orthogonality conditions as a GMM estimator (conditional expectations instead of unconditional covariances), where the instrument set is chosen to be the set of predetermined variables of the 'auxiliary' VAR.[13] In this context, the issue raised by Kurmann of the uncertainty in the estimate of the first step VAR would boil down to the issue of the choice of variables in X_{t-j}; insignificant VAR coefficients imply that those variables are weak instruments.

There is an important difference, however, between the restrictions exploited by the GMM approach described above, and the distance estimator of my formulation. These restrictions are stated in the *infinite horizon* form—conditions (3.8), which is based on the projection of all future values of the forcing variables that appear in the closed form solution. The GMM restrictions are instead stated in the *single period* form—conditions (3.10), using the projection of one period-ahead inflation onto the variables in Z_{t-1}; as such, they are a *non linear* transformation of conditions (3.8), obtained by postmultiplying

[13]This is the interpretation that Li (2003) gives to her estimate of the New Keynesian Phillips curve. While the orthogonality conditions appear the same, however, the distance $\widetilde{F}(\psi, \widehat{A})$ is not a sample mean, as in the method of moments estimation.

them by $(I - \beta A)$ (and using the fact that $(I - \beta A)^{-1}A = A(I - \beta A)^{-1}$). How this affects inference is a matter to be explored.[14]

3.4. Accounting for the VAR uncertainty

Whichever interpretation is given to the two-step distance estimator, a proper account should be given to the uncertainty associated with the first-step estimate of the autoregressive parameters. While one can easily derive appropriately corrected asymptotic standard errors,[15] in my application of this estimator to a two-variable model (Sbordone, 2003) I use instead a small sample approach, which is in the same spirit of Kurmann's assessment of the significance of the statistics of the Calvo model.

Specifically, I use the empirical distribution of the parameter vector $\alpha_\pi \equiv e'_\pi A$, draw from it N samples $\alpha_{\pi i}$ $(i = 1, \ldots, N)$, and for each of those I compute a minimum distance estimate $\widehat{\psi}_i$ of the vector of structural parameters ψ. I then compute the sample variance of the estimated $\widehat{\psi}_i$, and report its square root as the standard error.[16]

Furthermore, for each $\widehat{\psi}_i$ I compute a value of the distance function $\widehat{F}_i$, and from this generated sample I compute the covariance matrix of $\widehat{F}, \Sigma_F$. I use the last to compute a Wald statistic, $Q = F(\widehat{\psi})'\Sigma_F^{-1}F(\widehat{\psi})$, where $F(\widehat{\psi})$ is the value of the distance evaluated at the optimal parameter values, and use this statistic to evaluate the overall restrictions imposed by the model on the VAR structure.

3.5. Model misspecifications

The other papers in this volume address the issue of potential misspecifications of the basic NKPC model (1.1), which take the form of omitted variable problems. I already considered the general specification (1.5), which allows lagged inflation to affect current inflation directly, beyond its possible role in forecasting the labor share. Batini et al. (2005) extend the NKPC to the case of an open economy, and consider the role of material input prices, and of foreign competition. Furthermore, they allow for employment adjustment costs, which imply that both current and future employment enter the specification of their inflation equation.

Such modifications can be interpreted as corrections to the labor share in order to reach an appropriate measure of the real marginal cost.[17] For example, when facing labor adjustment costs employers may vary the effort margin: in this case an appropriate measure of labor input should include a measure of effort. But if effort depends on how hours are expected to grow, compared to actual hours, the marginal cost would differ from the average labor cost (or labor share) by such a difference. In this particular case, the theoretical real marginal cost that drives inflation dynamics is no more equal to the labor

[14] This issue was raised by Campbell and Shiller, and has been discussed by others as well. See for e.g. Lafontaine and White (1986).

[15] These involve the derivative of the model solution with respect to the second stage parameters, and the covariance matrix of the VAR parameters (an appendix is available from the author).

[16] N is set to 500 in this calculation.

[17] For an extensive discussion of how to construct suitable measures of marginal cost see Rotemberg and Woodford (1999).

share, but it is better approximated as follows

$$rmc_t = s_t + \delta_0(\mathrm{d}h_t - \delta_1 \mathrm{E}_t\, \mathrm{d}h_{t+1}), \tag{3.11}$$

where the term in brackets represent the expected deviation of future hours growth from current growth, and the coefficient δ_0 measures the curvature of the adjustment cost function.[18] When substituted in the pricing equation, this expression leads to an equation similar to the one obtained by Batini et al. (2005). A closed form solution for inflation dynamics of the form (3.1) is obtained by computing the forecast into the infinite future of the deviation of hours from the value expected one period ahead. In this case the two-step estimation approach requires, in the first step, the estimation of a VAR model extended to include hours of work: this allows to construct the hours forecast that appears in (3.11). In the second step, the function $F()$ is appropriately redefined to reflect the modification to the labor share as a measure of marginal costs.

4. Selected results

Table 1 reports some results obtained by applying the described methodology to estimating various specifications of the pricing model. The baseline unrestricted representation of the data is a VAR in inflation and labor share, with three lags ($p = 3$). Z_t is an mp-vector containing the current and $(p-1)$ lags of all elements of y_t, where $y_t = [\pi_t \quad \widetilde{s}_t]'$, and $\widetilde{s}_t$ is a measure of the labor share, transformed to obtain stationarity.[19] The parameters of the matrix A are estimated by OLS, and the consistent estimate of the covariance matrix of its relevant elements α_π $(\equiv e'_\pi A)$ is $\widehat{\Sigma}_{\alpha_\pi}$. The weighting matrix in the distance estimator is set equal to $\mathrm{diag}(\widehat{\Sigma}_{\alpha_\pi})$, which, given the interpretation of $F(\psi, A)$ as a set of restrictions on the parameters of the inflation equation, downweights the parameters which are estimated with higher uncertainty. The discount factor β is calibrated in all specifications to the value of .99.

The data cover the period 1951:1–2002:1 (a slightly longer period than that used by Kurmann, 2005); both the hypotheses that inflation has no predictive power for the labor share and that labor share has no predictive power for inflation can be rejected at standard confidence levels.[20]

The 'inertia' coefficient ζ, is, as we saw, a combination of various structural parameters. In the pure forward-looking model (row 1) the estimate of ξ is statistically significant, and corresponds to price rigidity of about 10 to 12 months.[21] π^m indicates the inflation series predicted by the model, and two statistics measure the approximation of predicted to actual inflation: the ratio of the standard deviations, $\sigma_{\pi^m}/\sigma_\pi$, in column 4, and the

[18]The model of labor hoarding that generates this result is developed in my previous work (Sbordone, 1996). The parameter δ_1 depends on the steady state value of the discount factor, and on the growth rate of hours and wages.

[19]The labor share is the ratio of real wage to productivity. I use instead the variable $\widetilde{s} = w - aq$ (with $a = .9558$), which eliminates the downward trend of the ratio which characterizes the data in the 1990s.

[20]The F-value of a Granger causality test of the predictive power of inflation for the labor share has a p-value of .051, and that for the predictive power of labor share for inflation has a p-value of .029. These results contrast with those of Kurmann, who finds absence of Granger causality, but are not due to the different sample length. One possible explanation is Kurmann's overparametrization.

[21]As I discuss in Sbordone (2002), the estimated ζ allows inference about the average time between price adjustments, providing one calibrates the capital share, and the parameter θ which drives the elasticity of demand. The numbers that I report are obtained using inflation measured on a quarterly basis.

Table 1
Parameter estimates and moments[a]

	ϱ	ζ	δ	$\sigma_{\pi^m}/\sigma_\pi$	corr(π^m, π)	Q-test
Pure forward-looking model		.025		.765	.919	2.25
		(.013)				[.89]
Generalized model	.224	.017		.721	.903	3.42
	(.15)	(.010)				[.75]
Excluding forward-looking terms	.488	.079		.432	.557	14.96
	(.08)	(.04)				[.02]
Adding labor adjustment costs	.254	.016	.046	.792	.824	6.27
	(.09)	(.010)	(.07)			[.73]
With endogenous labor share	.226	.026		.710	.905	23.55
	(.103)	(.006)				[.79]

[a]Standard errors in parentheses, p-values in square brackets.

correlation coefficient, corr(π^m, π), in columns 5. Both measures show a quite high degree of approximation. Moreover, as the statistic Q in the last column shows, the restrictions imposed by the model on the VAR are not rejected at standard significance levels.

The second row considers the role of lagged inflation. I find that π_{t-1} enters significantly the equation, and its inclusion reduces to some extent the size of the estimated coefficient of the forward-looking component, as would be expected in the case of an omitted variable problem. The fit in the other dimensions is similar. Note that in models with lagged inflation π_t^m is constructed sequentially, starting from the actual value of inflation in period 0 (1951:4 in the sample used here): $\pi_t^m = \widehat{\varrho}\pi_{t-1}^m + \widehat{\zeta}e_s'(I - \beta A)^{-1}Z_t$. Given the initial value of inflation, this series describes the evolution of inflation *implied* by the Calvo model, which depends at any point in time on the realization in the previous period (according to the model), on the current value of real marginal cost, and on a forecast of its future realizations. The statistics reported in cols. 4 and 5 measure how the volatility of this implied series compares to the volatility of actual inflation, and how close the dynamic evolution of the theoretical and the actual inflation series are.

The results obtained for the generalized model allow to evaluate the relative importance of backward vs. forward-looking components, an issue addressed by Rudd and Whelan (2005). As expression (1.3) shows, the weight on the backward-looking component is $\gamma^b = \varrho/(1 + \beta\varrho)$; the estimates reported in the table imply that γ^b is approximately .18, while the corresponding weight for the forward-looking component, $\gamma^f = \beta/(1 + \beta\varrho)$, is approximately .82. These values are consistent with the results of Gali and Gertler (1999) and Gali et al. (2001, 2005), and show that, even if one may reject the purely forward-looking version of the NKPC in favor of an equation including lagged inflation, the forward-looking component remains quantitatively more relevant. Note that these estimates, like those obtained by Rudd and Whelan (2005), are computed from the closed-form solution of the model, but differ from theirs in the way in which expected future values of the labor share are treated. While they proxy expected future values with realized values, I compute expected values as VAR forecast, using two different VAR specifications. They argue, however, that the forward-looking component doesn't add much to the explanation of inflation dynamics.

How do they reach this conclusion? First, it should be noted that they do not provide a structural interpretation of their lag polynomial, and are therefore not able to map their estimates of the lagged inflation coefficient into the weight of the backward-looking component in a model such as (1.2). Second, they seem primarily interested in comparing the purely forward-looking version of the NKPC with a univariate autoregressive model of inflation: although they find that the forward-looking terms in a generalized equation are statistically significant (Table 2, case B of their paper), they argue that they are quantitatively unimportant, and they do not significantly reduce the explanatory power of the own-lagged inflation terms.

By contrast, my benchmark is not a purely autoregressive model of inflation, but an unrestricted bivariate representation of inflation and labor share, and I ask whether the New Keynesian Phillips curve is a structural model that can provide an explanation for the inertial behavior of the data. In the generalized form of the NKPC model, lagged inflation derives from the assumption of partial indexation, which can be justified in the context of the micro foundations of the model, typically by information costs associated with reoptimization. Furthermore, to evaluate the importance of the forward-looking terms, I ask what would be the fit of the model were the forward looking component be set to zero. In the context of the closed-form solution of the generalized Calvo model, this amounts to setting to zero all but the contemporaneous value of the labor share: it is *not* equivalent to estimate an autoregressive model of inflation.

When estimating the model under this constraint (the results are on the third row of Table 1), I obtain coefficient estimates on lagged inflation and on labor share both significantly higher than in the case in which expected future marginal costs are allowed to matter: this can be interpreted as evidence that there is an omitted variable problem in this specification. More importantly, though, I obtain in this case a much poorer approximation of the inflation dynamics, as the statistics presented in the table show. The reverse restriction, which sets the backward-looking component to zero, gives the purely forward-looking pricing equation (3.1) that, as we already saw, provides instead quite a good approximation of the dynamics of inflation.

Augmenting the model to allow for labor adjustment costs does not improve the fit of the model. When real marginal cost is specified as in (3.11), the coefficient that measures the curvature of the adjustment cost function is not statistically significant, whether I measure labor by hours or employment.[22] This difference from the result of Batini et al. (2005), however, may be due to the tighter specification of the adjustment cost adopted here, or may reflect some structural difference between US and UK, and it is certainly worth further investigation.

Finally, to show how one can go further with this methodology into the specification of marginal costs, I report in the last line of the table estimates of the parameters ϱ and ζ from my (2003) study where I analyze both inflation and wage dynamics. The structural model considered in that study adds to inflation a second equation describing the evolution of the labor share, derived from a model of wage setting with staggered contracts. The two-equation model therefore imposes a number of additional restrictions on the time series representation of wages and prices which are exploited for estimation. The parameter vector ψ in this case is six-dimensional, and includes parameters describing the intertemporal rate of substitution between leisure and consumption and the degree of wage indexation (though I report here only the estimates of the parameters of the inflation

[22]Row 4 in the table reports results using hours. In the estimation I set $\delta_1 = 1$, and report in the table the estimate of the coefficient of the hours term in the projected inflation, which is $\delta = \zeta\delta_0$.

equation). The elements of the distance function F include in this case restrictions on the VAR parameters of both the inflation and the labor share equations. As the reported estimates show, the endogenization of the wage process allows a sharper estimate of the coefficients of the inflation process, and otherwise confirms the single-equation results.

A final footnote on another point discussed by Kurmann (2005): the inertia coefficient, in all of the estimates presented here, when allowing forward looking terms, ranges between .017 and .026. These results are consistent with the estimates obtained by Gali–Gertler (1999), or Batini et al. (2005), for example, and they also imply a degree of inertia in prices similar to that reported by survey estimates. This implies that parametrizations of ζ as high as the value .08 chosen by Kurmann (2005) do not appear to be supported by the data.

5. Conclusion

In this paper I discuss the two-step estimation procedure used in Sbordone (2002), give a more general interpretation to it, and present some additional results on the estimation of the New Keynesian Phillips curve.

I show that under this more general interpretation, the auxiliary forecasting model on which the procedure relies (the first step of the estimation) is an unrestricted representation of the data, against which to test the model. While the uncertainty of the first estimation stage, discussed by Kurmann (2005), can be taken into account within the procedure itself, issues about the VAR modeling, like the preliminary stationarity-inducing transformations, the size of the model and the lag length, and the time invariance of the structure still remain to be addressed.[23] And ultimately only an increase in the precision of the VAR estimates can reduce the uncertainty surrounding the derived statistics that Kurmann documented so thoroughly.

The partial-information estimation strategy that I discussed has the advantage of relying on a small number of restrictions (in the case analyzed here, those specific to the inflation dynamics) which must hold in every model that incorporates the same form of inflation dynamics. Moreover, as the application to the model of price and wage dynamics shows, one can sequentially endogenize variables that are initially modeled only with an unrestricted time series model.

What does all this imply for the empirical assessment of the Calvo model of inflation dynamics? I would argue that the pricing model explored here is a good representation of the data, and price stickiness of this kind is a valid hypothesis to incorporate into more complete models for business cycle and policy analysis. In particular, the forward-looking terms are quite important in explaining the dynamics of inflation: while it is possible to reproduce the dynamics of inflation fairly well with a purely forward-looking model, eliminating instead the dependence on expected future values of the labor share significantly worsens the overall fit.

The validity of this pricing model, however, does not necessarily imply a relation between inflation and *output* of the form generally referred to as the NKPC. What has emerged from the copious empirical research on inflation dynamics, in my opinion, is that

[23] I address the issue of the structural invariance of the Calvo parameters in joint work with Tim Cogley (Cogley and Sbordone (2005)). We estimate an unrestricted time series model of inflation with drifting parameters, and investigate the issue of whether the parameters of the Calvo model are invariant to instability in trend inflation.

a full understanding of the Phillips curve can in fact be reached only through an understanding of the dynamics of labor costs, and how these relate to output dynamics. And this is where future empirical research should be focused.

References

Batini, N., Jackson, B., Nickell, S., 2005. An Open Economy New Keynesian Phillips Curve for the U.K., Journal of Monetary Economics, this issue, doi:10.1016/j.jmoneco.2005.08.003.

Calvo, G., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12 (3), 383–398.

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062–1088.

Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy, 113, 1–45.

Cogley, T., Sbordone, A.M., 2005. A search for a structural Phillips curve. Federal Reserve Bank of New York Staff Reports No. 203.

Gali, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. Journal of Monetary Economics 44, 195–222.

Gali, J., Gertler, M., Lopez-Salido, J.D., 2001. European Inflation Dynamics. European Economic Review 45 (7), 1237–1270.

Gali, J., Gertler, M., Lopez-Salido, J.D., 2005. Robustness of the estimates of the hybrid New Keynesian Phillips curve. Journal of Monetary Economics, this issue, doi:10.1016/j.jmoneco.2005.08.005.

Hansen, L.P., Sargent T.J., 1980. Formulating and estimating dynamic linear rational expectations models. Journal of Economic Dynamics and Control 2, 7–46.

Kurmann, A., 2005. Quantifying the uncertainty about a forward-looking New Keynesian pricing model. Journal of Monetary Economics, this issue, doi:10.1016/j.jmoneco.2005.08.004.

Lafontaine, F., White, K.J., 1986. Obtaining any wald statistic you want. Economic Letters 21, 35–40.

Li, H., 2003. Testing alternative theories of aggregate supply. Princeton University, unpublished.

Nelson, E., Neiss, K., 2005. Inflation dynamics, marginal cost and the output gap: evidence from three countries. Journal of Money, Credit and Banking 37 (7).

Rotemberg, J.J., Woodford, M., 1999. The cyclical behavior of prices and costs. In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, vol. 1B. North-Holland, Amsterdam.

Rudd, J., Whelan, K., 2005. New tests of the New-Keynesian Phillips curve. Journal of Monetary Economics, this issue, doi:10.1016/j.jmoneco.2005.08.006.

Sbordone, A.M., 1996. Cyclical productivity in a model of labor hoarding. Journal of Monetary Economics 38, 331–361.

Sbordone, A.M., 2002. Price and unit labor costs: a new test of price stickiness. Journal of Monetary Economics 49, 265–292.

Sbordone, A.M., 2003. A Limited Information Approach to the Simultaneous Estimation of Wage and Price Dynamics. Rutgers University, wp # 2003-23.

Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press.