

Empirical Comparison of Sticky Price and Sticky Information Models

Oleg Korenok*

November 2004,

Revised September 2005

Abstract: The goal of this paper is to provide a fair empirical comparison of two alternative explanations of the relationship between aggregate price and output. We compare the empirical performance of the sticky price and the Mankiw and Reis (2002) sticky information models. We put both models in a similar analytical form and use the same data set on unit labor cost and aggregate prices in the U.S. after WWII to evaluate the models. We use the Bayesian full information likelihood approach for parameter estimation, uncertainty evaluation, and model comparison. Statistical comparison of the two non-nested models and estimates of the empirical encompassing model lead to the same result - the sticky information model is dominated by the sticky price model.

JEL classification: E12, E3, C32

Keywords: sticky price, sticky information, model selection, full information likelihood, Bayesian model comparison

* Oleg Korenok (okorenok@vcu.edu): Department of Economics, Virginia Commonwealth University, 1015 Floyd Avenue, Box 84-4000, Richmond, VA 23284. The author owes a great many thanks to Bruce Mizrach and Argia Sbordone for numerous useful comments and guidance on an earlier draft of this paper. The author wish to thank Roberto Chang, Kevin Grier, Oleksiy Kryvtsov, Carol S. Lehr, Florian Pelgrin, Laura Razzolini, Ricardo Reis, Norman R. Swanson, Greg Tkacz, Hiroki Tsurumi, Christian Zimmermann and seminar participants at Bank of Canada, Bowdoin College, Rutgers University, Virginia Commonwealth University, University of Connecticut and University of Oklahoma.

1 Introduction

The goal of this paper is to provide a fair empirical comparison of two alternative explanations of the relationship between aggregate price and output. We compare the empirical performance of the sticky price model and the Mankiw and Reis (2002) sticky information model. Statistical comparison of the two non-nested models and estimates of the empirical encompassing model lead to the same result - the sticky information model, is dominated by the sticky price model.

The literature on the New Keynesian Phillips curve has identified several ways in which the standard sticky price model does not adequately model the relationship between aggregate prices and output. Ball (1994) found that the model implies that announced credible disinflation causes booms rather than recessions. Fuhrer and Moore (1995) showed that the sticky price model falls short of explaining inflation persistence in US. Finally, Mankiw and Reis (2002) noted that the model has trouble explaining why shocks to monetary policy have delayed and gradual effects on inflation.

Mankiw and Reis (2002) propose a sticky information model that can potentially address the failures of the New Keynesian Phillips curve. They posit that information about macroeconomic conditions spreads slowly because of information acquisition or re-optimization costs. Compared to the standard sticky price model, prices in this setup are always readjusted, but decisions about prices are not always based on the latest available information. The model is representative of the wider class of Rational Inattention models developed in Phelps (1970), Lucas (1973), Sims (2003) and Woodford (2003).

To compare these two alternative models, we utilize a theoretical relation between aggregate prices and unit labor cost that allows us to leave unspecified household preferences, wage setting and money demand. We introduce then a modeling approach to nest the sticky price and the sticky information models within a single empirical framework. We use a single-step estimation method that provides consistent estimates of adjustment speeds and reliable confidence bands that enable us to reject flexible price hypothesis.

To estimate information and price stickiness, we follow the approach, advocated by Gali and Gertler (1999) and Sbordone (2002), and use the theoretical relation between aggregate prices and unit labor cost. For estimation purposes, we transform the relation into its stationary equivalent: the relation between real marginal cost, measured by unit labor cost/price ratio, and changes in unit labor cost. Our methodology differs in several respects from previous attempts to estimate

information stickiness. First, unlike Mankiw and Reis (2002), the relation between unit labor cost and prices is derived from a firm's optimization problem. Second, it does not require taking a stand on household preferences or wage setting as in Khan and Zhu (2002) in order to relate inflation and the output gap. Finally, as shown in Sbordone (2002), our conclusions from this approach are quite robust to alternative specifications of the production function, which specifies the relation between unobserved marginal cost and unit labor cost.

We use the method of undetermined coefficients to solve both models. The same solution technique enables us to construct and to estimate an empirical model that encompasses the sticky price and sticky information models as special cases, even though two theoretical models are non-nested. This encompassing model facilitates the models comparison. Our solution methodology differs from previous attempt by Sbordone (2002) to estimate price stickiness. She employs the Campbell and Shiller (1988) approach to find a rational expectations solution for the sticky price model. Unfortunately, such approach can not be applied to the sticky information model because of complex expectations formation.

We use a Bayesian version of the full information likelihood approach, which enables us not only to estimate parameters and the statistical uncertainty around them, but also to compare directly non-nested sticky information and sticky price models. Full information likelihood approach was advocated by Fuhrer and Moore (1995)¹ for empirical evaluation of the sticky price model and hypothesis testing. Our estimation methodology differs from previous studies by Mankiw and Reis (2001), Sbordone (2002), and Khan and Zhu (2002), who use a two-step approach. They estimate the driving processes in the first step and then, in the second step, they incorporate it into a linear rational expectations solution for process of interest (unemployment, unit labor cost/price ratio or the output gap respectively). Parameters, estimated in the first step, are taken as given in the second step. Thus, the measure of uncertainty around the parameter of information or price stickiness is inaccurate.

Our main result is that a formal statistical comparison of the two models favors the sticky price model. Summarizing other estimation results, the estimate of information stickiness implies that agents in the economy revise information about every 15 months. With 95% probability, agents revise information sets in the interval between 12 and 18 months. The estimate of price stickiness implies that the average time between price changes is 10 months. With a 95% probability, agents

¹See also Linde (2002) on comparison of full likelihood and limited likelihood methods in sticky price model framework

change prices in the interval between 9 and 11 months. The estimates are robust to reasonable permutations of prior distributions, estimation method, and sample size.

The rest of the paper is organized as follows. The optimization based dynamics of real marginal cost and the rational expectation solutions are described in Section 2. We explain the Bayesian full likelihood estimation of information and price stickiness models in Section 3. The methodology for models comparison is described in Section 4. The data and estimation results are presented and discussed in Section 5. The results of the comparison of the sticky price and sticky information models are gathered in Section 6. Finally, concluding remarks are given in Section 7.

2 Optimization Based Dynamics of Marginal Cost

In this section, we derive the optimization based dynamics of marginal cost for the sticky price and sticky information models. Also we solve out expectations in both models using the method of undetermined coefficients.

We assume a continuum of monopolistically competitive firms, indexed by i , which produce differentiated goods. we also assume that firms operate with a technology,

$$Y_{i,t} = H_{i,t}^a, \quad (1)$$

where output, $Y_{i,t}$, is firm specific; labor, $H_{i,t}$, is the only factor of production and is unique for each agent, and $a \in [0, 1]$.

The demand curve for product i can be represented as:

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t} \right)^{-\theta} Y_t, \quad (2)$$

where P_{it} is the optimization based price choice of the firm i , P_t is the index of prices, Y_t is the aggregator function defined as $Y_t = \left(\int_0^1 Y_{i,t}^{\frac{\theta-1}{\theta}} di \right)^{\frac{\theta}{1-\theta}}$, and θ is the Dixit-Stiglitz elasticity of substitution among differentiated goods.

A supplier of good i does not believe that his price decisions can affect either the aggregate demand Y_t or the aggregate price P_t because good i is a very small part of households' consumption. Thus the supplier chooses the price $P_{i,t}$, taking aggregate demand and aggregate price as given.

Optimization then involves setting a price,

$$\frac{P_{i,t}}{P_t} = \mu S_{i,t}, \quad (3)$$

where $S_{i,t}$ is real marginal cost of supplier and $\mu = \frac{\theta}{\theta - 1}$ is a markup. Taking into account the firm technology (1), the real marginal cost of the individual firm can be written as

$$S_{i,t} = \frac{1}{a} \frac{W_t H_{i,t}}{P_t Y_{i,t}}, \quad (4)$$

where W_t is a labor compensation on a common, perfectly competitive labor market. Assuming a specific aggregator function for labor input and using the demand curve, we can rewrite (4) as:

$$S_{i,t} = S_t \left(\left(\frac{P_{i,t}}{P_t} \right)^{-\theta} \right)^{\frac{1-a}{a}}, \quad (5)$$

where $S_t = \frac{1}{a} \frac{W_t H_t}{P_t Y_t}$ is the *average* level of real marginal costs in the economy. This expression shows that the real marginal cost of the firm i does not need to be the same as the average. In the model the fact that firms charge different prices determines their different levels of sales, and hence firm's different level of marginal costs.

Combining the profit maximizing rule (3) with the marginal cost (5) yields a relation between the average level of marginal cost in economy, individual firm price choices and the aggregate price level,

$$\frac{P_{i,t}}{P_t} = \mu S_t \left(\left(\frac{P_{i,t}}{P_t} \right)^{-\theta} \right)^{\frac{1-a}{a}}. \quad (6)$$

I can rewrite (6) as:

$$p_{i,t} - p_t = s_t - \frac{\theta(1-a)}{a} (p_{i,t} - p_t), \quad (7)$$

where we use lower case variables to denote the log transformation. Here and below, we ignore constants because we focus on deviations from the mean in empirical applications. (7) can be solved for the individual price in terms of the aggregate price and marginal cost:

$$p_{i,t} = \alpha s_t + p_t, \quad (8)$$

where $\alpha = \frac{a}{a + \theta(1-a)}$, $\alpha < 1$ when $\theta > 0$. We interpret (8) as a positive relation between desired

relative price of the individual firm and real marginal cost. The relation is a direct consequence of profit maximizing rule for the individual firm.

In the above discussion, all firms choose the price of good i each period. The price choice is independent of prices that were charged in the past. The choice is based on the full information about current demand and cost. These conditions characterize the perfectly flexible price model. Now we want to consider two deviations from the perfectly flexible model: one, where the fraction of firms have to keep the last period price, the sticky price model, and the other, where the fraction of firms have to use past or outdated information, the sticky information model.

2.1 Sticky Price Model

I follow Calvo (1983) and assume that in every quarter a fraction of firms λ_p can set a new price independently of the past history of price changes. Other firms have to keep the last period price. The expected time between price updates is therefore $\frac{1}{\lambda_p}$ quarters. The model nests the monopolistically competitive economy with perfectly flexible prices in the case $\lambda_p = 1$.

Because the opportunity to change price is random, a firm wants to set its price equal to an average of desired prices until next adjustment:

$$x_{i,t} = \lambda_p \sum_{j=0}^{\infty} (1 - \lambda_p)^j E_t p_{i,t+j}, \quad (9)$$

where $x_{i,t}$ is the price level set by a firm when it has such possibility. Note that index for each individual firm can be dropped since all firms according to (8) make similar choices.

The overall price level is the average of firms who were able to adjust price and those who were not

$$p_t = \lambda_p x_t + (1 - \lambda_p) p_{t-1}. \quad (10)$$

The system of three equations (the firm's desired and actual price (8) and (9), and the overall price level (10)) can be solved for the relation between inflation and real marginal costs

$$\pi_t = \frac{1}{\xi} s_t + E_t \pi_{t+1}, \quad (11)$$

where $\xi = \frac{(1-\lambda_p)}{\alpha \lambda_p^2}$ is a nonlinear function of structural parameters. Equation (11) implies that current inflation can rise because increase in cost or agent's expectation of high inflation in the future.

From the definition of S_t and the definition of average unit labor cost it follows that real marginal cost is proportional to unit labor cost/price ratio,

$$s_t = ulc_t - p_t. \quad (12)$$

Using (12), we follow Sbordone (2002) and derive the relation between the real marginal cost and changes in unit labor cost,

$$s_t = z_1 s_{t-1} + \Delta ulc_t - (1 - z_1) \sum_{j=0}^{\infty} z_2^{-j} E_t \Delta ulc_{t+j}, \quad (13)$$

where z_1 and z_2 are the real roots of the characteristic polynomial of the difference equation in p_t , $p(z) = z^2 - (2 + \frac{1}{\xi})z + 1 = 0$, with $0 < z_1 < 1 < z_2$. (13) describes the process for marginal cost in the sticky price model as a weighted average of the past marginal costs, current and expected future changes in unit labor cost.

2.2 Sticky Information Model

As in the flexible price model, Mankiw and Reis (2002) assume that an individual firm makes a price adjustment each period. The novelty of the approach is the assumption that the information used for decision-making is not necessarily the current one. Only a fraction of firms λ_{inf} uses current information in pricing decisions, while the remaining fraction $1 - \lambda_{inf}$ uses past or outdated information. The expected time between price updates is therefore $\frac{1}{\lambda_{inf}}$ quarters. The model nests the monopolistically competitive economy with perfectly flexible prices when $\lambda_{inf} = 1$.

One can think about the sticky information model as a variant of the sticky price model, where a fraction of firms that are unable to set prices optimally instead of keeping the last period price use a more complex updating scheme. Firms, when they have an opportunity to set prices optimally, solve not only for optimal current prices but also for the infinite path of future prices based on available information. When they are not able to set prices optimally, firms set the price to the appropriate value from this solution set.

Each period, a representative firm sets the price to

$$x_{i,t}^j = E_{t-j} p_{i,t}, \quad (14)$$

where j represents the latest period when the firm updated its information set. As in the sticky

price model index for each individual firm can be dropped since all firms according to (8) make similar choices. The overall price in this case is an average among the firms who update information sets in different periods,

$$p_t = \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^j x_t^j. \quad (15)$$

Combining (8) and (14) with (15) yields,

$$p_t = \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^j E_{t-j}(\alpha s_t + p_t). \quad (16)$$

Note that the structure of expectations here is different from the usual backward iteration of expectations: expectations are formed for the current t value of the variable at the different $t - j$ periods in the past (for example $E_{t-j}p_t$).

With some algebra (see the Appendix for details), (16) can be transformed into an expression that describes the dynamics of real marginal cost,

$$s_t = \eta \Delta ulc_t + \eta s_{t-1} + \eta \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^j E_{t-j-1}((1 - \alpha) s_t - (1 - \alpha) s_{t-1} - \Delta ulc_t), \quad (17)$$

where $\eta = \frac{1 - \lambda_{inf}}{1 - \lambda_{inf}(1 - \alpha)} < 1$.

Dynamics of (17) is comparable to the dynamics of real marginal cost in the sticky price model. In both models, it depends on a weighted average of past realizations of real marginal cost, current changes in unit labor cost, and expectations. This similarity enables us to compare the sticky price and sticky information models on the basis of observed series for the unit labor cost/price ratio and the first difference in unit labor cost.

The principal difference between the sticky price and sticky information models is in expectations formation. In the sticky price model, agents have an opportunity to set desired price on rare occasions. As a result, they have to take into account future costs during periods when they are not able to set price. The expectations term, which is the weighted average of current expectations about future changes in unit labor cost, is the way agents take into account future costs. In the sticky information model, on the other hand, agents can maximize profit at each period. Though, occasionally, by assumption, they are not able to use current information in maximization. As a result, agents use the best measure of current information available to them: expectations about current realizations of necessary variables made in the past. In particular, they use the weighted

average of expectations about current and previous values of real marginal cost and current change in unit labor cost made in the past.

2.3 Rational Expectations Solution

I solve out for unobserved expectations by assuming agents are rational and using the method of undetermined coefficients.

Assume that Δulc_t follows a general linear process with the MA representation,

$$\Delta ulc_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k}, \quad (18)$$

where ψ_k , $k = 0, 1, 2, \dots$, is a sequence of parameters and ϵ_t are serially uncorrelated random variables with zero mean.

For the sticky price model, a solution for the stochastic process s_t satisfies (13) and (18). One can guess that the solution has the following general form:

$$s_t = \sum_{k=0}^{\infty} \gamma_k \epsilon_{t-k}. \quad (19)$$

Finding the solution for s_t is equivalent to finding γ_k that satisfy (13) and (18). To solve for γ_k , we substitute s_t and s_{t-1} from (19) and Δulc_t from (18) in (13). By moving (18) forward and taking expectations with respect to information in period t , we substitute expectations of future differences of unit labor cost as well. The substitutions result in

$$\sum_{k=0}^{\infty} \gamma_k \epsilon_{t-k} = z_1 \sum_{k=0}^{\infty} \gamma_k \epsilon_{t-1-k} + \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k} - (1 - z_1) \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} z_2^{-j} \psi_{k+j} \epsilon_{t-k}. \quad (20)$$

By equating coefficients of $\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots$ on both sides of (20) we solve for γ_k in terms of z_1, z_2 and ψ_k ,

$$\begin{aligned} \gamma_0 &= \psi_0 - (1 - z_1) \sum_{j=0}^{\infty} z_2^{-j} \psi_j, \\ \gamma_k &= z_1 \gamma_{k-1} + \psi_k - (1 - z_1) \sum_{j=k}^{\infty} z_2^{k-j} \psi_j, \\ k &= 1, 2, \dots \end{aligned} \quad (21)$$

The forward looking nature of prices and therefore unit labor cost/price ratio results in the solution

for γ_k which depends on future ψ_k .

For the sticky information model, a solution for the stochastic process s_t has to satisfy (17) and (18). We make a similar guess that (19) is the solution. I substitute s_t and s_{t-1} from (19) and Δulc_t from (18) in (17). Finally, we can substitute expectations with respect to information in period $t - j - 1$ in (17) by taking appropriate expectations of (18) and (19). The substitutions result in:

$$\begin{aligned} \sum_{k=0}^{\infty} \gamma_k \epsilon_{t-k} &= \eta \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k} + \eta \sum_{k=0}^{\infty} \gamma_k \epsilon_{t-k-1} + \\ &\eta \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^j \left((1 - \alpha) \sum_{k=j+1}^{\infty} \gamma_k \epsilon_{t-k} - (1 - \alpha) \sum_{k=j}^{\infty} \gamma_k \epsilon_{t-k-1} - \sum_{k=j+1}^{\infty} \psi_k \epsilon_{t-k} \right). \end{aligned} \quad (22)$$

By equating coefficients of $\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots$ on both sides of equality we solve for γ_k in terms of η, λ_i and ψ_k :

$$\begin{aligned} \gamma_0 &= \eta \psi_0 \\ \gamma_k &= \frac{\eta (1 - \lambda_{inf})^k}{1 - \eta (1 - \alpha) (1 - (1 - \lambda_{inf})^k)} \psi_k + \frac{\eta (1 - (1 - \alpha) (1 - (1 - \lambda_{inf})^k))}{1 - \eta (1 - \alpha) (1 - (1 - \lambda_{inf})^k)} \gamma_{k-1} \\ k &= 1, 2, \dots \end{aligned} \quad (23)$$

The difference in RE solutions for the two models can be traced to the difference in profit maximization problems. In the sticky price model solution (21) γ_k depends not only on past and current values of ψ_k , but also on all future values. Future values appear because agents have to take into account future costs when they are not able to set the price. In the sticky information model RE solution (23), γ_k depends only on past and current values of ψ_k because agents can maximize profit at each period using only current and past information.

The method of undetermined coefficients enables us to achieve two goals. First, we find the rational expectations solution of the sticky information model which is not feasible in the Campbell and Shiller approach. Second, we describe the dynamics of real marginal cost and changes in unit labor cost in a similar analytical form. The similar analytical form, in turn, allows us to construct the empirical encompassing model that nests the sticky information and sticky price models, even though two models are non-nested.

3 Estimation

In this section, we describe a Bayesian full information likelihood approach for estimation and model comparison. The posterior distribution, the main object of interest in Bayesian analysis, is the product of a likelihood and a prior:

$$p(\boldsymbol{\theta} | data) = \pi(\boldsymbol{\theta}) L(\boldsymbol{\theta}; data),$$

where $\boldsymbol{\theta}$ is a vector of parameters. We use the general representation for (19), together with exogenous process Δulc_t , (18), to form the likelihood function. Then we formulate priors for structural parameters in the model.

3.1 Likelihood Function

I start by assuming that the driving process, Δulc_t , has a stationary AR(q) representation. We truncate the infinite MA representation of s_t at k_{max} and add an error term v_t to compensate for truncation. The error term can be also justified by the fact that the description of driving process is potentially incomplete. One can think of other shocks to inflation which were not included in the original theoretical model.

The processes for the change in unit labor cost and marginal cost together can be written as:

$$\begin{aligned} s_t &= \sum_{k=0}^{k_{max}} \gamma_k \epsilon_{t-k} + v_t, \\ \Delta ulc_t &= \sum_{j=1}^q \rho_j \Delta ulc_{t-j} + \epsilon_t, \end{aligned} \tag{24}$$

where k_{max} is a truncation point. From the second expression it follows that

$$\epsilon_{t-k} = \Delta ulc_{t-k} - \sum_{j=1}^q \rho_j \Delta ulc_{t-k-j},$$

which, when substituted into the first expression, results in

$$s_t = \left(\sum_{k=0}^{k_{max}} \gamma_k \left(\Delta ulc_{t-k} - \sum_{j=1}^q \rho_j \Delta ulc_{t-k-j} \right) \right) + v_t.$$

Let $\boldsymbol{\Lambda}$ be a vector of structural parameters of interest: $\boldsymbol{\Lambda} = \boldsymbol{\xi}$ in the sticky price model and

$\mathbf{\Lambda} = (\lambda_{inf}, \alpha)'$ in the sticky information model. From the rational expectations solution (21) and (23), it follows that the coefficients of marginal cost process, γ_i , are functions of structural parameters, $\mathbf{\Lambda}$, and coefficients of the driving process, $\boldsymbol{\psi} = [\psi_0, \psi_1, \dots]$, where ψ_i are taken from (18). It is easy to solve for ψ_i as a function of $\boldsymbol{\rho}$, $\boldsymbol{\rho} = [\rho_1, \rho_2, \dots]$, i.e. in the AR(1) case, $\psi_i = \rho_1^i$.

Taking into account the above transformations, (24) can be written as:

$$\begin{aligned} s_t &= \left(\sum_{k=0}^{k_{max}} \gamma_k(\mathbf{\Lambda}, \boldsymbol{\rho}) (\Delta ulc_{t-k} - \sum_{j=1}^q \rho_j \Delta ulc_{t-k-j}) \right) + v_t, \\ \Delta ulc_t &= \sum_{j=1}^q \rho_j \Delta ulc_{t-j} + \epsilon_t. \end{aligned} \quad (25)$$

The coefficients $\gamma_k(\mathbf{\Lambda}, \boldsymbol{\rho})$ are different for two models. The analytical form for $\gamma_k(\mathbf{\Lambda}, \boldsymbol{\rho})$ is a generalization of (21) and (23).

Error terms (ϵ_t, v_t) are independent by construction. We assume that they are draws from a multivariate normal distribution with zero mean and variance covariance matrix $\boldsymbol{\Sigma}$. We collect all parameters of the model to be estimated in a single vector $\boldsymbol{\theta} = (\boldsymbol{\Sigma}, \boldsymbol{\rho}, \mathbf{\Lambda})'$. The likelihood function then can be written, apart from a constant, as

$$L(\boldsymbol{\theta}; data) \propto |\boldsymbol{\Sigma}|^{(-1/2T)} \exp(-1/2 \text{tr } \boldsymbol{\Sigma}^{-1} \mathbf{Q}), \quad (26)$$

where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{v}' \mathbf{v} & \mathbf{v}' \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon}' \mathbf{v} & \boldsymbol{\epsilon}' \boldsymbol{\epsilon} \end{bmatrix}.$$

I use the notation \mathbf{v} for the vector of v_t and $\boldsymbol{\epsilon}$ for the vector of ϵ_t . Both vectors can be solved from (25).

3.2 Prior Selection

The prior for information stickiness is chosen to be non-informative, e.i. it is Uniform on the range from zero to one, which covers all range of possible values for information stickiness. We chose to center prior for ξ at Sbordone's (2002) estimate, but with a wide enough variance to include the possibility of perfectly flexible prices, $\lambda_p = 1$, which would imply $\xi = 0$.

A prior for $\boldsymbol{\rho}$ was chosen to be normal with a large variance and centered at zero. A bivariate inverted Wishart distribution is chosen as a prior for joint distribution of parameters $\sigma_v, \sigma_{v,\epsilon}$ and

σ_ϵ . The information about prior distributions for all parameters is summarized in Table 1.

3.3 Models Comparison

In this section we discuss a Bayesian comparison of non-nested sticky price and sticky information models. We also describe construction of the empirical encompassing model that nests two models.

The problem is to compare two parametric models for the data, defined by a probability density function $L(data|\boldsymbol{\theta}_{sp})$ and $L(data|\boldsymbol{\theta}_{si})$, where $\boldsymbol{\theta}_{sp} = (\xi, \boldsymbol{\Sigma}, \boldsymbol{\rho})$ are parameters in the sticky price model and $\boldsymbol{\theta}_{si} = ((\lambda_{inf}, \alpha)', \boldsymbol{\Sigma}, \boldsymbol{\rho})$ are parameters in the sticky information model. Then the full parameter family space is $\Theta_{sp} \cup \Theta_{si}$. Its prior can be constructed as probability π that space is Θ_{sp} and $1 - \pi$ that space is Θ_{si} . Then the posterior probability that the sticky price model is true has the following form:

$$p_{sp} = \frac{\pi \int_{\boldsymbol{\theta}_{sp} \in \Theta_{sp}} L(data|\boldsymbol{\theta}_{sp})\pi(\boldsymbol{\theta}_{sp})d\boldsymbol{\theta}_{sp}}{\pi \int_{\boldsymbol{\theta}_{sp} \in \Theta_{sp}} L(data|\boldsymbol{\theta}_{sp})\pi(\boldsymbol{\theta}_{sp})d\boldsymbol{\theta}_{sp} + (1 - \pi) \int_{\boldsymbol{\theta}_{si} \in \Theta_{si}} L(data|\boldsymbol{\theta}_{si})\pi(\boldsymbol{\theta}_{si})d\boldsymbol{\theta}_{si}},$$

and the posterior probability that the sticky information model is true is equal to $p_{si} = 1 - p_{sp}$.

The main problem in evaluating the posterior probability is the estimation of marginalized likelihood. Marginalized likelihood is $\int_{\boldsymbol{\theta} \in \Theta} L(data|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = m(data)$ and one can use the basic likelihood identity to write

$$m(data) = \frac{L(data|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{p(\boldsymbol{\theta}|data)},$$

where the numerator is a product of the likelihood and the prior, and the denominator is the posterior density of $\boldsymbol{\theta}$. To calculate the marginal likelihood, we use the Gelfand and Dey (1994) procedure modified and implemented by Geweke (1999).² Note that, to calculate the posterior probability that one of the models is true, the methodology does not require the models to be nested.

There is an alternative to the direct comparison of non-nested models. This alternative comparison is possible because rational expectations solutions for both models have similar analytical form. We construct the empirical encompassing model that nests the sticky price and sticky information models by introducing a weighting parameter ω and defining a new set of parameters γ_k^* as:

$$\gamma_k^* = \omega\gamma_k(\lambda_{inf}, \alpha, \boldsymbol{\rho}) + (1 - \omega)\gamma_k(\xi, \boldsymbol{\rho}), \quad (27)$$

²The program **mlike** is used which can be found at www.econ.umn.edu/~bacc

and replace with them γ_k in (25). The resulting model encompasses the sticky price model as a special case when $\omega = 0$. The model encompasses also the sticky information model when $\omega = 1$.

To estimate the nested model, we extend the vector of parameters of the likelihood function by combining structural parameters for both models, $\Lambda^* = (\lambda_{inf}, \alpha, \xi)'$ and adding the weighting parameter ω , $\theta = (\Sigma, \rho, \Lambda^*, \omega)'$. A uniform prior is chosen for the weighting parameter over the interval $[0, 1]$. If the estimate of weighting parameter ω is not significantly different from zero, we interpret the result as empirical support to the sticky price model. If the parameter is not significantly different from one, the result can be interpreted as an empirical support to the sticky information model. Otherwise result is inconclusive.

4 Data and Estimation Results

The unit labor cost/price ratio is calculated from the historical quarterly observations of the aggregate price level (GDP deflator, FRED database) and unit labor cost series (US Department of Labor) for the sample from 1960:1 to 2002:1.

Based on the Schwartz information criteria, we choose an AR(2) representation for the difference in unit labor cost. By inverting the estimated coefficients of the AR(2), ρ , one can solve for the MA coefficients, ψ_i . We truncate the MA representation of unit labor cost/price ratio at $k_{max} = 20$. The choice of k_{max} is motivated by two reasons: by the accuracy of solution, and by the fact that with a finite sample we want to keep the number at reasonable level. At $k = 20$ the estimates of coefficients for the unit labor cost/price ratio, γ_i , are close to zero.

A detailed description of the Markov Chain Monte Carlo (MCMC) algorithm is given in the appendix A.2. The length of MCMC chain is 15,000 draws with 5,000 burn. To monitor convergence we use standard techniques, such as the Kolmogorov-Smirnov test, the ZG test, and convergence of chains from different starting points.

The discussion of empirical results in a Bayesian analysis is based on comparing the prior and posterior distributions of the structural parameters. If the data are informative, it should change our beliefs about the parameters. Tables 2, 3 and 5 summarize these comparisons. We compare the mean of the prior distribution in column 2 with the mean of the posterior distribution in column 4. Prior and posterior standard deviations can be found in column 3 and 5. In addition, we report the 95% highest posterior density interval (HPDI) in columns 6 and 7 which are used for hypothesis testing.

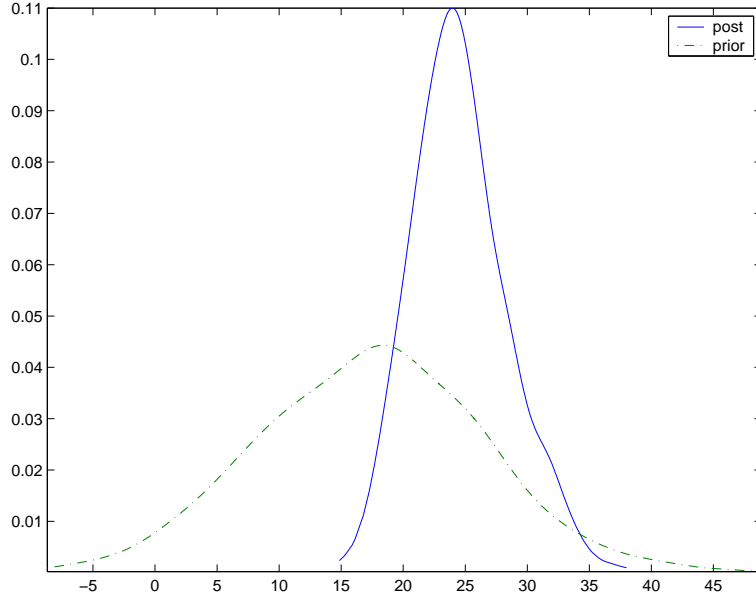


Figure 1: Posterior and Prior Distributions of ξ .

4.1 Sticky Price Model

Panel A of the Table 2 presents our estimates of the sticky price model.

The main parameter of interest in the sticky price model is ξ , row 3 of panel A. The experiment updates beliefs for the location and dispersion of the parameter. The mean of the posterior, 24.54, is somewhat higher than the prior mean, 17.2. The posterior standard deviation, 3.80, is twice lower than the prior standard deviation, 9. A similar observation follows from comparing the prior density and the estimate of the posterior density for ξ in Figure 1. The lower and upper bounds of 95% HPDI interval are 17.61 and 32.2. The 95% HPDI for ξ does not include zero, which implies that λ_p is different from one. From this observation, we conclude that perfectly flexible price hypothesis is rejected.

The parameter ξ is a nonlinear combination of structural parameters a , θ and λ_p , which can not be separately identified. We follow Sbordone (2002) and assume that the share of labor $a = 0.75$ and the Dixit-Stiglitz elasticity of substitution among differentiated goods $\theta = 6$ (which implies the average value of markup of 1.2). Then, using definition of ξ , the implied price stickiness, λ_p , can be calculated. The average implied price stickiness is 0.29, with 95% probability interval between 0.26 and 0.34. The implied price stickiness estimate translates into the average time between price changes which is equal to 10 months and the 95% interval is between 9 and 11 months. The average estimated time between price adjustments is somewhat higher than Sbordone's estimate, 9 months,

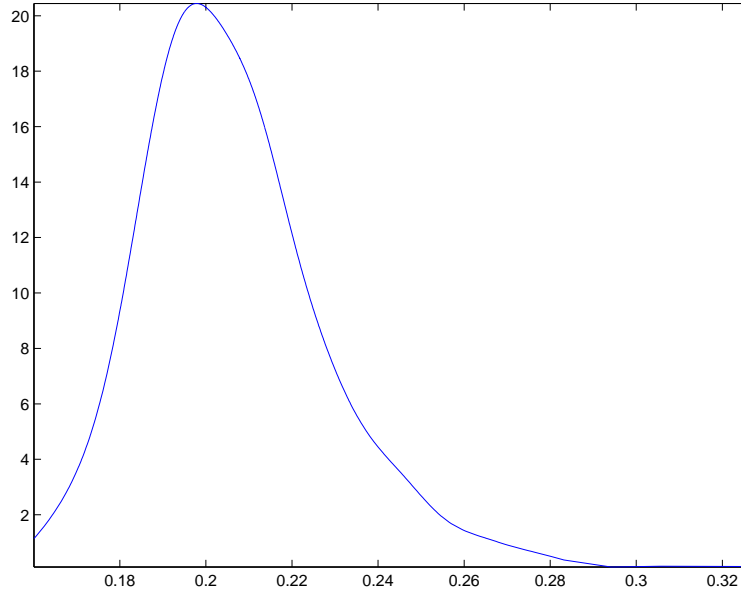


Figure 2: Posterior Distribution of Information Stickiness.

though not significantly, as the HPDI includes 9 months.

Our estimates for the average length that prices remain sticky is similar to previous ones. The benefit of our approach is improvement in evaluation of uncertainty. Sbordone's estimate is half of mine, but this estimate takes the parameters of driving process ρ_i as given. This introduces downward bias in the estimate of standard deviation.³ Having unbiased full information likelihood estimate of standard deviation, we are more confident rejecting the economy with perfectly flexible prices.

4.2 Sticky Information Model

Panel B of the Table 2 presents the summary of the prior and posterior distributions of the sticky information model parameters.

The main object of interest is the parameter of information stickiness, λ_{inf} , row 3 of panel B. The experiment updates beliefs about location and dispersion of information stickiness. The posterior mean 0.21 is half the prior mean of 0.5. The posterior standard deviation, 0.02, is 10 times lower than the prior standard deviation. Similar observations follow from the plot of marginal posterior density of the parameter λ_{inf} , Figure 2.⁴ The economic interpretation of this result is

³See Pagan (1984) for the formal asymptotic argument in linear setup. The author shows that the estimate of variance is inconsistent and biased downward.

⁴I do not plot the prior density since it is a trivial horizontal line over the region $[0,1]$.

that on average agents in economy update information once in every five quarters. The posterior mean for information stickiness is very close to the value assumed by Mankiw and Reis (2002), 0.25, or estimated by Mankiw and Reis(2001) and Khan and Zhu (2002).

The estimate of information stickiness is significantly different from an economy with perfectly flexible prices. The boundaries of the 95% highest posterior density interval are far from 1. The HPDI values imply that with 95% probability agents update their information set from once every year to once every year and a half. Despite the difference in estimation methods and the data used, the estimates of standard errors are close to Khan and Zhu (2002).

The likelihood is less informative about the second structural parameter α , the fourth row of panel B. As with information stickiness, beliefs for both location and dispersion are updated. The posterior mean increases to 0.84 from the prior mean 0.5. The posterior standard deviation is twice lower than the prior standard deviation. The HPDI indicates that the plausible range for α is between 0.60 and 1, which is twice smaller than the prior range. The estimated range for α implies a high value of the labor share a , or a low value of the Dixit-Stiglitz elasticity of substitution among differentiated goods θ , which in turn implies a high markup.

To summarize, even though our estimates of information stickiness are close to the previous estimates, an important benefit of our approach is improvement in uncertainty evaluation: we are more confident rejecting the economy with perfectly flexible prices.

4.3 Robustness Check

To estimate the sticky price or sticky information model, previous empirical studies used different solution and estimation methods. In order to check the robustness of our results and facilitate comparison, we replicate their results with our variables and the same data sample.

4.3.1 Alternative Solution to RE

Sbordone (2002) applied the Campbell and Shiller methodology to solve out expectations. In particular, she estimated the driving process in companion form, $Z_t = AZ_{t-1} + \nu_t$, where Δulc_t is the first element of vector Z_t . The forward looking expectations of the first element, $E_t \Delta ulc_{t+j}$, equal to $e_1 A^j Z_t$, where e_1 is a unit vector with a first element equal to 1 and all the other elements

equal to 0. In this case, the infinite sum in (13), reduces to

$$\sum_{j=0}^{\infty} z_2^{-j} E_t \Delta ulc_{t+j} = e_1 (I - z_2^{-1} A)^{-1} Z_t.$$

The empirical estimates for alternative solutions of the sticky price model differ very little as it is clear from comparing panel A and panel C of the Table 2. To construct panel C, we use the AR(2) process for difference in unit labor cost to substitute expectations in the equation of marginal cost, then we estimate resulted system of equations via Bayesian likelihood method described in subsection 3.1. The estimates of ρ_i , $\sigma_{1,1}$, $\sigma_{1,2}$, and $\sigma_{2,2}$ are quite close. The main difference is an increase in the mean estimate of ξ . The increase is not statistically significant: the mean estimates are inside the 95% HPDI for both panels. We conclude that the estimates of the level of price stickiness appears robust to solution method.

Unfortunately, the same method can not be applied to the sticky information model, (17), because in addition to expectations of the driving process, it contains expectations of the real marginal cost, the process for which a solution is needed.

4.3.2 Two-step vs. One-step Estimation

To estimate the parameters of information or price stickiness, Sbordone (2002), Mankiw and Reis (2001) and Khan and Zhu (2002) used a two step procedure. In the first step, the authors estimated a statistical model for a driving process. It is equivalent to estimating the second expression in (24). In the second step, they estimated structural parameters assuming that the estimates from the first step are given. The uncertainty estimate is biased downward because by taking the first step estimates as given one ignores the uncertainty associated with them. To evaluate the size of the increase for our sample we estimate structural parameters taking values of OLS estimates of ρ_i as given.

Table 3 reports the parameter estimates. The data support that the two stage estimates understate the uncertainty of the parameter estimates. The standard errors and HPDIs in Table 3 for parameters λ_i , α and ξ are smaller than the full likelihood estimates in Table 2. The reduction of uncertainty for information stickiness is not substantial. On the other hand, the standard deviation of ξ reduces by 24% from 3.8 to 2.95.

5 Comparing Sticky Price and Sticky Information Models

Table 4 summarizes the comparison of the sticky price and sticky information models. Panel A reports the estimates from a full information likelihood approach, our baseline estimation technique. The sticky price model dominates the sticky information model empirically. The log of the marginal likelihood of the sticky price model is 820. It is significantly higher than the log of marginal likelihood of the sticky information model, 813. The result is robust both across different sample sizes and reasonable permutations of prior distributions.

In addition to the discussed models, panel A also reports the values of marginal likelihoods for the less restrictive models of (25). In the first model, row 4, γ_i are estimated as unconstrained reduced form coefficients. In the second model, row 5, we keep the same lag structure for Δulc_t , but to estimate coefficients we use Δulc_{t-k} as exogenous variables instead of $\Delta ulc_{t-k} - \sum_{j=1}^q \rho_j \Delta ulc_{t-k-j}$. The model with reduced form γ_i and the restricted VAR model are dominated by theoretical models. They have lower marginal likelihood than any of the theoretical models, 805 and 802 respectively. This observation suggests that theoretical models give the best performance, when unit labor cost/price ratio dynamics is explained only by the changes in unit labor cost.

I also compare models estimated in two stages. Panel B of Table 4 reports the estimates. Comparing panel A and panel B, one can see that the sticky price versus sticky information ordering does not change. The sticky price model still has higher marginal likelihood than the sticky information model. The observation allows us to conclude that the sticky price dominance is robust to alternative estimation methods. The rank of solution methods, within the sticky price model, does change. The change in rank suggests that there is no clear dominance between the solution methods.

The alternative way to compare models is to estimate the empirical encompassing model that nests the sticky price and sticky information models as special cases. These estimates support our previous finding that the sticky price model dominates the sticky information model. Table 5 reports the estimates. From comparing the prior and posterior distributions for weighting parameter ω in row 6 of the Table, one can see that the mean decreases from 0.50 to 0.18 and the standard deviation falls by more than 50% from 0.29 to 0.11. The weighting parameter is not significantly different from zero: the left boundary of the 95% HPDI is equal to 0.0003. The same observations follow from Figure 3, which reports the full posterior distribution of the weighting parameter. One can see that a lot of posterior density is concentrated near zero.

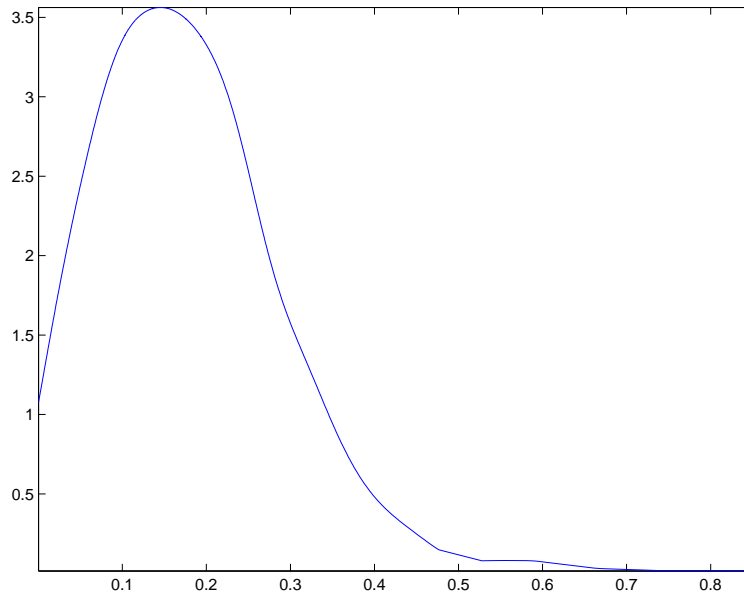


Figure 3: Posterior Distribution of ω .

6 Conclusions

The goal of this paper is to provide a fair empirical comparison of two alternative explanations of the relationship between aggregate price and output. We compare the empirical performance of the sticky price and sticky information models. We put both models in a similar analytical form and use the same data on unit labor cost and aggregate prices to evaluate the models. A Bayesian full information likelihood approach is used for parameter estimation, uncertainty evaluation, and model comparison.

I find that the posterior of the information stickiness parameter is informative and centered at 0.2 which implies that agents update their information set every five quarters. With 95% probability agents revise the information set in the interval between 4 and 6 quarters. Despite differences in methodology and data set, the estimates are close to estimates of Mankiw and Reis (2001) and Khan and Zhu (2002). The posterior of the price stickiness parameter is also informative. We corroborate earlier findings that the average time between price changes is less than a year. With a 95% probability the average time is between 9 and 11 months. Estimates for both models are robust to reasonable permutations of prior, estimation method or sample size. A full information approach enables us to compute proper estimates of standard errors. Using these unbiased estimates, we still reject the hypothesis of perfectly flexible prices for both models. Statistical comparison of the two non-nested models and estimates of the empirical encompassing model lead to the same result -

the sticky information model, is dominated by the sticky price model in a 40 year sample period that included four major disinflation episodes.

Results suggest a possibility of similar finding for information rigidities within bigger general equilibrium models. In Korenok and Swanson (2005), we extend comparison to models with specified consumer preferences and money demand. We emphasize that the theoretical appeal of the sticky information model is limited only to the special case of anticipated persistent disinflation. The theoretical responses to unanticipated, monetary policy shocks for the sticky price and sticky information models are similar.

A Appendix

A.1 Equation for inflation

I start with (16). I substitute (12) into (16) to get

$$p_t = \lambda_i \sum_{j=0}^{\infty} (1 - \lambda_{inf})^j E_{t-j}(\alpha ulc_t + (1 - \alpha)p_t), \quad (\text{A1})$$

the expression that describes the aggregate price level as a weighted average of past expectations of current value of unit labor cost and current price. Taking out the first term and redefining the summation index,

$$p_t = \lambda_{inf} ((1 - \alpha)p_t + \alpha ulc_t) + \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^{(j+1)} E_{t-1-j}(\alpha ulc_t + (1 - \alpha)p_t), \quad (\text{A2})$$

which can be rearranged in the following form,

$$\frac{1}{(1 - \lambda_{inf})} (p_t - \lambda_{inf} ((1 - \alpha)p_t + \alpha ulc_t)) = \lambda_{inf} \sum_{j=0}^{\infty} E_{t-1-j}(\alpha ulc_t + (1 - \alpha)p_t) (1 - \lambda_{inf})^j. \quad (\text{A3})$$

For the period $t - 1$ price level can be written as,

$$p_{t-1} = \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^j E_{t-1-j}((1 - \alpha)p_{t-1} + \alpha ulc_{t-1}), \quad (\text{A4})$$

Subtracting (A4) from (A2) we obtain,

$$\begin{aligned} \pi_t &= \lambda_{inf} ((1 - \alpha)p_t + ulc_t) + \lambda_{inf} \sum_{j=0}^{\infty} (1 - \lambda_{inf})^{(j+1)} E_{t-1-j}(\alpha ulc_t + (1 - \alpha)p_t) \\ &\quad - \lambda_{inf} \left(\sum_{j=0}^{\infty} (1 - \lambda_{inf})^j E_{t-1-j}((1 - \alpha)p_{t-1} + \alpha ulc_{t-1}) \right). \end{aligned} \quad (\text{A5})$$

I rearrange terms in (A5) and define $\Delta ulc_t = ulc_t - ulc_{t-1}$:

$$\begin{aligned} \pi_t &= \lambda_{inf} (\alpha ulc_t + (1 - \alpha)p_t) + \sum_{j=0}^{\infty} \lambda_{inf} (1 - \lambda_{inf})^j E_{t-1-j}((1 - \alpha)\pi_t + \alpha \Delta ulc_t) \\ &\quad - \lambda_{inf}^2 \sum_{j=0}^{\infty} E_{t-1-j}(\alpha ulc_t + (1 - \alpha)p_t) (1 - \lambda_{inf})^j. \end{aligned} \quad (\text{A6})$$

Substituting into the last term of (A6) from (A3),

$$\pi_t = \frac{\lambda_{inf} \alpha}{1 - \lambda_{inf}} (ulc_t - p_t) + \sum_{j=0}^{\infty} \lambda_{inf} (1 - \lambda_{inf})^j E_{t-1-j} (\alpha \Delta ulc_t + (1 - \alpha) \pi_t). \quad (\text{A7})$$

Using the relation between unit labor cost, unit labor cost/price ratio and inflation, we get the equation (17) in the text.

A.2 The Gibbs-sampling algorithm

The Gibbs-sampling algorithm consists of three steps:

1. Conditional on the values of $(\mathbf{\Sigma}, \boldsymbol{\rho})'$, generate $\mathbf{\Lambda}$ using the random walk Metropolis Hastings algorithm:
 - (a) Set initial values of $\mathbf{\Lambda}^0$. Initial values can be set arbitrarily, but we set $\mathbf{\Lambda}^0$ as means of prior distributions for structural parameters of the model.
 - (b) Draw parameters $\mathbf{\Lambda}^{i+1}$ from the following generating function:

$$g(\mathbf{\Lambda}^{i+1} | \mathbf{\Lambda}^i) \sim N(\mathbf{\Lambda}^i, \mathbf{\Sigma}_{\mathbf{\Lambda}}),$$

where $\mathbf{\Sigma}_{\mathbf{\Lambda}}$ is a variance-covariance matrix of the 1,000 previous draws. I update it after each 1,000 draws. Accept the new draw $\mathbf{\Lambda}^{i+1}$ with the acceptance probability $\alpha(\mathbf{\Lambda}^i, \mathbf{\Lambda}^{i+1})$ which is defined as

$$\alpha(\mathbf{\Lambda}^i, \mathbf{\Lambda}^{i+1}) = \frac{p(\mathbf{\Lambda}^{i+1}, \boldsymbol{\rho}, \mathbf{\Sigma} | data)}{p(\mathbf{\Lambda}^i, \boldsymbol{\rho}, \mathbf{\Sigma} | data)}.$$

Note that there is no ratio of generating functions since we use the random walk Metropolis-Hasting algorithm.

2. Conditional on the values of $(\mathbf{\Sigma}, \mathbf{\Lambda})'$ generate $\boldsymbol{\rho}$ using the Metropolis Hastings algorithm. It is convenient to rewrite the likelihood (26) using the fact that the inverse covariance matrix can be written as:

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{\sigma_{v,\epsilon}}{\sigma_\epsilon^2} & 1 \end{bmatrix} \begin{bmatrix} \left(\sigma_v^2 - \frac{\sigma_{v,\epsilon}^2}{\sigma_{\epsilon_2}} \right)^{-1} & 0 \\ 0 & \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{\sigma_{v,\epsilon}}{\sigma_\epsilon^2} \\ 0 & 1 \end{bmatrix}.$$

Then, we restate the likelihood,

$$\begin{aligned}
L(\sigma_v, \sigma_{v,\epsilon}, \sigma_\epsilon, \boldsymbol{\rho}, \boldsymbol{\Lambda}) &= \left(\sigma_v^2 - \frac{\sigma_{v,\epsilon}^2}{\sigma_\epsilon^2} \right)^{-1/2T} \\
&\times \exp \left(-1/2 \operatorname{tr} \left(\sigma_v^2 - \frac{\sigma_{v,\epsilon}^2}{\sigma_\epsilon^2} \right)^{-1} \left(\mathbf{v} - \frac{\sigma_{v,\epsilon}}{\sigma_\epsilon^2} \boldsymbol{\epsilon} \right)' \left(\mathbf{v} - \frac{\sigma_{v,\epsilon}}{\sigma_\epsilon^2} \boldsymbol{\epsilon} \right) \right) \\
&\times (\sigma_\epsilon^2)^{-1/2} \exp \left(-1/2 \operatorname{tr} \frac{\boldsymbol{\epsilon}' \boldsymbol{\epsilon}}{\sigma_\epsilon^2} \right).
\end{aligned} \tag{A8}$$

I use second part of the likelihood to generate a vector of parameters $\boldsymbol{\rho}$,

$$p(\boldsymbol{\rho} | \boldsymbol{\Lambda}, \boldsymbol{\Sigma}, \text{data}) = \pi(\boldsymbol{\rho}) (\sigma_\epsilon^2)^{-1/2} \exp \left(-1/2 \operatorname{tr} \frac{\boldsymbol{\epsilon}' \boldsymbol{\epsilon}}{\sigma_\epsilon^2} \right).$$

This choice of random walk generating density improves the convergence properties of the algorithm significantly. The rest of step 2 is equivalent to step 1.

3. Conditional on the values of $(\boldsymbol{\rho}, \boldsymbol{\Lambda})'$, we generate $\boldsymbol{\Sigma}$ from,

$$p(\boldsymbol{\Sigma}, | \boldsymbol{\rho}, \boldsymbol{\Lambda}, \text{data}) = |\boldsymbol{\Sigma}|^{-1/2(T+d_0+q+1)} \exp(-1/2 \operatorname{tr} \boldsymbol{\Sigma}^{-1} (\mathbf{Q} + \mathbf{S}_0)),$$

which is the kernel of an Inverted Wishart distribution with parameters $d = d_0 + T$, $\mathbf{S} = \mathbf{Q} + \mathbf{S}_0$.

This form follows from the likelihood function (26) and the prior specified in Table 1.

At the end of each iteration, we have draws of the structural and covariance parameters. We repeat iterations 15,000 times saving each iteration. To avoid dependence on the starting value of the algorithm, we delete the first 5,000 draws. We use next 10,000 draws to compute the statistical summaries of marginal posterior densities for the structural parameters (i.e. mean, standard deviation and HPDI).

References

- [1] Ball, L., (1994), "Credible Disinflation with Staggered Price Setting," *American Economic Review*, LXXXIV, 282-289.
- [2] Blinder, A. S., Canetti, E.R.D., Lebow, D. and J.B. Rudd, (1998), *Asking About Prices: A New Approach to Understand Price Stickiness*, Russell Sage Foundation, New York.
- [3] Calvo, G. A., (1983), "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, XII, 383-398.
- [4] Campbell, J.Y. and R.J. Shiller, (1988), "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, I, 195-228.
- [5] Christiano, L.J., Eichenbaum, M. and C.L. Evans, (1997), "Sticky Price and Limited Participation Models of Money: A Comparison," *European Economic Review*, 41, 1201-1249.
- [6] Fernandez, D., (1992), *Bank Lending and the Monetary Policy Transmission Mechanism: Evidence from Japan*. Princeton, New Jersey: Princeton University.
- [7] Fuhrer, J, and G. Moore, (1995), "Inflation Persistence," *Quarterly Journal of Economics*, CX, 127-160.
- [8] Gali., J. and M. Gertler, (1999), "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 44, 195-222.
- [9] Gelfand, A.E., and D.K. Dey, (1994), "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society, Ser. B*, 56, 363-375.
- [10] Geweke J., (1999), "Computational Experiments and Reality," University of Minnesota and Federal Reserve Bank of Minneapolis, June 1999.
- [11] Khan, H. and Z. Zhu., (2002), "Estimates of the Sticky-Information Phillips Curve for the United States, Canada, and the United Kingdom," Bank of Canada Working Paper No. 2002-19.
- [12] Korenok, O., and N. Swanson, (2005), "How Sticky Is Sticky Enough? A Distributional and Impulse Response Analysis of New Keynesian DSGE Models," in process.
- [13] Linde J., (2002), "Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach," Sveriges Riksbank Working Paper Series, No. 129.
- [14] Lucas, Robert E., Jr., (1973), "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review* 63: 326-334.
- [15] Mankiw, N.G. and R. Reis, (2001), "Sticky Information: A Model of Monetary Nonneutrality and Structural Slumps," NBER Working Paper, No. 8614.
- [16] Mankiw, N.G. and R. Reis., (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, Vol. CXVII(IV),1295-1328.
- [17] Pagan, A., (1984), "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, 25, 183-209.

- [18] Phelps, E.S., (1970), "Introduction: The New Microeconomics in Employment and Inflation Theory," in E.S. Phelps et al., *Microeconomic Foundations of Employment and Inflation Theory*, New York: Norton.
- [19] Sbordone A., (2002), "Prices and Unit Labor Costs: A New Test of Price Stickiness," *Journal of Monetary Economics*, vol. 49 (2), 265-292.
- [20] Sims, C.A., (2003), "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), April 2003.
- [21] Woodford, M., (2002), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, Princeton University.
- [22] Woodford, M., (2003), "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Edited by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, Princeton University Press, 25-59.

Table 1: **Prior Distributions**

$\pi(\boldsymbol{\rho}) \sim$	$N(\mathbf{0}_{[2,1]}, \mathbf{I}_{[2,2]})$
$\pi(\lambda_{inf}) \sim$	$U([0, 1])$
$\pi(\alpha) \sim$	$U([0, 1])$
$\pi(\xi) \sim$	$N(17.2, 81)$
$\pi(\boldsymbol{\Sigma}) \sim$	$IW_2(0.01\mathbf{I}_{[2,2]}, 2)$
$\pi(\omega) \sim$	$U([0, 1])$

Notes: Here λ_{inf} and α are structural parameters of the sticky information model, ξ is a structural parameter of the sticky price model, $\boldsymbol{\rho}$ is a vector of parameters for driving process, $\boldsymbol{\Sigma}$ is variance-covariance matrix, and ω is a weighting parameter of the empirical encompassing model. N stands for the Normal distribution, U - Uniform distribution, IW - Inverted Wishart distribution.

Table 2: Full Information Likelihood Parameter Estimates

<i>A. Sticky Price Model</i>						
	<i>Prior Distribution</i>		<i>Posterior Distribution</i>		<i>HPDI</i>	
	<i>mean</i>	<i>s.d.</i>	<i>mean</i>	<i>s.d.</i>	<i>2.5%</i>	<i>97.5%</i>
ρ_1	0.3308	1.0000	0.3380	0.0699	0.2017	0.4761
ρ_2	0.2847	1.0000	0.2625	0.0654	0.1381	0.3857
ξ	17.200	9.0000	24.540	3.7996	17.614	32.198
$\sigma_{1,1}$	0.0200	0.0700	0.0003	2.5e-5	0.0003	0.0004
$\sigma_{1,2}$	0.0000	0.0200	-6.5e-6	1.8e-5	-4.1e-5	2.6e-5
$\sigma_{2,2}$	0.0100	0.0200	0.0002	1.2e-5	0.0001	0.0002

<i>B. Sticky Information Model</i>						
	<i>Prior Distribution</i>		<i>Posterior Distribution</i>		<i>HPDI</i>	
	<i>mean</i>	<i>s.d.</i>	<i>mean</i>	<i>s.d.</i>	<i>2.5%</i>	<i>97.5%</i>
ρ_1	0.0000	1.0000	0.3290	0.0705	0.1935	0.4639
ρ_2	0.0000	1.0000	0.2279	0.0680	0.0930	0.3584
λ_{inf}	0.5000	0.2887	0.2068	0.0223	0.1660	0.2520
α	0.5000	0.2887	0.8414	0.1223	0.5926	1.0000
$\sigma_{1,1}$	0.0200	0.0700	0.0003	2.8e-5	0.0003	0.0004
$\sigma_{1,2}$	0.0000	0.0200	-1.9e-5	1.8e-5	-5.4e-5	1.6e-5
$\sigma_{2,2}$	0.0100	0.0200	0.0002	1.2e-5	0.0001	0.0002

<i>C. Sticky Price Model, Alternative Solution</i>						
	<i>Prior Distribution</i>		<i>Posterior Distribution</i>		<i>HPDI</i>	
	<i>mean</i>	<i>s.d.</i>	<i>mean</i>	<i>s.d.</i>	<i>2.5%</i>	<i>97.5%</i>
ρ_1	0.3308	1.0000	0.3580	0.0745	0.2171	0.4987
ρ_2	0.2847	1.0000	0.2513	0.0744	0.1084	0.3925
ξ	17.200	9.0000	28.847	4.4486	20.548	37.814
$\sigma_{1,1}$	0.0200	0.0700	0.0003	2.2e-5	0.0002	0.0003
$\sigma_{1,2}$	0.0000	0.0200	-1.6e-5	1.6e-5	-4.8e-5	1.7e-5
$\sigma_{2,2}$	0.0100	0.0200	0.0002	1.2e-5	0.0001	0.0002

Notes: Here λ_{inf} and α are structural parameters of the sticky information model, ξ is a structural parameter of the sticky price model, ρ_1 and ρ_2 are parameters for driving process, and $\sigma_{1,1}$, $\sigma_{1,2}$ and $\sigma_{2,2}$ are parameters of variance-covariance matrix.

Table 3: Second Stage Parameters Estimates

	<i>Prior Distribution</i>		<i>Posterior Distribution</i>			
	<i>mean</i>	<i>s.d.</i>	<i>mean</i>	<i>s.d.</i>	<i>HPDI</i>	
					2.5%	97.5%
<i>Sticky price model</i>						
ξ	17.200	9.0000	24.733	2.9542	19.265	30.546
$\sigma_{1,1}$	0.0100	0.0200	0.0003	2.5e-5	0.0003	0.0004
<i>Sticky information model</i>						
λ_{inf}	0.5000	0.2887	0.1973	0.0170	0.1667	0.2313
α	0.5000	0.2887	0.8522	0.1070	0.6532	1.0000
$\sigma_{1,1}$	0.0100	0.0200	0.0003	2.7e-5	0.0003	0.0004
<i>Sticky price model, alternative solution</i>						
ξ	17.200	9.0000	28.727	3.0680	22.755	34.744
$\sigma_{1,1}$	0.0100	0.0200	0.0003	2.0e-5	0.0002	0.0003

Notes: See notes to Table 4.

Table 4: **Estimates of Marginal Likelihood**

<i>A. Full Information Likelihood Estimates</i>	
Model	$lnm(data)$
Sticky information model	812.964
Sticky price model	819.966
Sticky price model, alternative solution	819.179
Reduced form, unrestricted γ_i	805.277
Restricted VAR	802.338

<i>B. Second Stage Estimates</i>	
Model	$lnm(data)$
Sticky information model	378.975
Sticky price model	386.582
Sticky price model, alternative solution	391.988

Table 5: **Parameter Estimates for Nested Model**

	<i>Prior Distribution</i>		<i>Posterior Distribution</i>			
	<i>mean</i>	<i>s.d.</i>	<i>mean</i>	<i>s.d.</i>	<i>HPDI</i>	
					2.5%	97.5%
ρ_1	0.0000	1.0000	0.3320	0.0707	0.1968	0.4631
ρ_2	0.0000	1.0000	0.2489	0.0683	0.1137	0.3838
λ_{inf}	0.5000	0.2887	0.5221	0.2563	0.1018	0.9293
α	0.5000	0.2887	0.5264	0.2590	0.0762	0.9817
ξ	17.200	9.0000	26.702	7.5346	9.5954	40.317
ω	0.5000	0.2887	0.1804	0.1144	0.0003	0.3753
$\sigma_{1,1}$	0.0200	0.0700	0.0003	2.6e-5	0.0003	0.0004
$\sigma_{1,2}$	0.0000	0.0200	-8.3e-6	1.7e-5	-4.3e-5	2.6e-5
$\sigma_{2,2}$	0.0100	0.0200	0.0002	1.2e-5	0.0001	0.0002

Notes: See notes to Table 4. ω is a weighting parameter of the empirical encompassing model.