Adaptive Learning and Inflation Dynamics in a Flexible Price Model

Sergio Santoro

February 13, 2004

Abstract

In most of the recent macroeconomics literature, the sticky reaction of prices in response to changes in aggregate conditions has been modelled following the highly influential contribution of Calvo (1983). However, this approach has difficulties in accounting for some well-established stylized facts, like the sluggish and delayed response of inflation to demand shocks, and the positive correlation between real output and the rate of change of inflation. In this paper, we will investigate the possibility of a simple flexible prices and monopolistic competitive model to match this features, when the expectations of the firms are formed following the adaptive learning literature.

1 Introduction

In most of the recent macroeconomics literature, the sticky reaction of prices in response to changes in aggregate conditions has been modelled following the highly influential contribution of Calvo (1983). In particular, it is assumed that, in each period, firms face a constant probability to reset prices optimally. Due to its appealing analytical tractability, this approach has become the workhorse of most of monetary policy literature (see for example Clarida et al. (1999)).

However, this model has difficulties in explaining some well-established stylised facts: in particular, two robust features of data are a sluggish and delayed response of inflation to demand shocks (e.g., see Christiano et al. (2001)), and a positive correlation between real output and the rate of change of inflation (the so-called acceleration phenomenon, see Mankiw and Reis (2002)). Both of these patterns are not replicated by the Calvo-type staggered price settings; the main theoretical reason is the fact that, despite of the stickiness of price level, inflation can respond rapidly to exogenous shocks.

In this paper, we propose a different source of intrinsic stickiness in firms' pricing behavior: in particular, we assume that firms do not have an exact knowledge of the "true" economic model, but form their expectations according to their most recent estimates of the law of motion of the unknown aggregate variables. This approach is typical of the adaptive learning literature, which has received an increasing attention in recent years (see Evans and Honkapohja (2001) for an extensive monograph).

Up to now, most effort in this literature has been devoted to the issue of stability under learning, namely under which conditions a rational expectations equilibrium is a limiting point of the learning process; only recently there have been attempts to evaluate quantitatively the effect of introducing adaptive learning into a macro or finance model, and to test the ability of this framework to explain empirical facts[1]. This may be due to the caveat that learning can introduce too many degrees of freedom in the model, allowing the researcher to match any pattern of data just playing with the learning algorithm.

The aim of this paper is to build a simple flexible price model of monopolistic competition, augmented by the adaptive learning formation of agents' expectations, and to investigate (via numerical simulations) whether this is able to outperform the staggered price model in replicating the above mentioned features of data. In doing so, we will try to use a learning scheme with a basis in a payoff-maximizing choice of the agents, in order to make less severe the potential cricism toward *ad hoc* learning procedures. To assess the goodness of fit of the model, we will use some techniques to evaluate calibrated dynamic general equilibrium stochastic models presented in the survey of Canova and Ortega (2000).

It is worth noting that the approach we will develop in this paper is conceptually linked to the one of Mankiw and Reis (2002), who assume that firms are free to reset prices in each period, but that information diffuses slowly among them; in their model, in each period firms face a constant probability to update their information set.

2 Stylized Facts and Calvo Model

In recent work on monetary policy issues, the standard tool to model the firms' pricing behavior has been the so-called Calvo model: in each period, firms have a constant probability to reset their price optimally, and they do so taking into account that such a price will last for an unknown number of periods. Since this probability is independent of the last time a specific firm has reset its price, this approach leads to an analitically tractable framework which constitutes one of the building blocks of what Clarida et al. (1999) call the New Keynesian Science of Monetary Policy.

Besides its theoretical appeal, this approach has done well in replicating certain empirical patterns (like, for example, the high aurocorrelation of inflation); nevertheless, it has shown many difficulties in explaining some well-established stylized facts that are also common wisdom of policymakers. In particular, we will concentrate on two features of the Calvo model that are at odds with empirical evidence:

[1]See Timmermann (1993) and Timmermann (1996) for applications to the stock market, Marcet and Nicolini (1998) for a model that aims to explain hyperinflations in South american countries, and Sargent (1999) for and explanation of the change in the U.S. inflation pattern.

- many empirical investigations (and conventional wisdom of central bankers) show that nominal shocks have a sluggish and delayed effect on inflation, and that the impulse response of inflation is hump-shaped[2]. Instead, the Calvo model is characterized by a monotonic decreasing impulse response function[3];

- another widely documented empirical fact is a positive relationship between real output and the growth rate of inflation. This pattern, that Mankiw and Reis (2002) call acceleration phenomenon, has been shown through the use of scatterplots by many economists[4], and has been confirmed by Mankiw and Reis (2002) calculating the correlation between real output and the growth rate of inflation for U.S. data. Instead, the Calvo model predicts a slightly negative value for this statistics.

Both of these shortcomings are originated by the fact that in the baseline version of the Calvo model inflation is a purely forward looking-variable, without any intrinsic source of inertia; to overcome this problem, there have been many attempts to improve the empirical fit of the New Keynesian framework through the introduction of some source of inertia in the Phillips curve[5].

Some alternative avenues to reconcile monetary theory with data have been tried. In particular, some recent papers have abandoned the approach of assuming that firms face some kind of constraint on their possibility of resetting each period their prices; instead, they assume that prices are fully flexible, but that the information set of the firms is somehow constrained.

In this spirit, Mankiw and Reis (2002) introduced the so-called "sticky information Phillips curve": in their model, the information is supposed to spread slowly across the economy, so that each firm faces every period a constant probability to update its information set. Since it is free to reset prices every period, the firm would set prices such that its expected profit, given the latest update of the information set, is maximized. In Mankiw and Reis (2002), the authors show that this alternative assumption can outperform the Calvo model in a simple business cycle framework[6].

Another example of this new strand of literature is the model of Woodford (2001) where, following the pioneering idea of Phelps (1970) and the highly influential paper of Lucas (1972), the firms are assumed not to be able to observe correctly the level of the aggregate variables, and to take their decisions on the basis of their subjective expectations. The main difference with the Lucas' model is that the information constraint is not simply a one-period delay in the aggregate variables' observability, but, following Sims (2003), "a limited

[2]See Christiano et al. (1999) and (2001).

[3]For more details, see Section 4.3 and the Appendix.

[4]See, e.g. Abel and Bernanke (1998) and Blanchard (2000).

[5]For example, Rotemberg and Woodford (1997) have introduced a decision delay for some price setters, and Christiano et al. (2001) have added various other source of nominal rigidity together with price stickiness, like adjustment costs and wage stickiness.

[6]For an application of the sticky information Phillips curve to optimal monetary issues, see Ball et al. (2003).

capacity of the private decision-makers to *pay attention* to all of the information in their environment". In this way, the agents' decision process is converted into a signal-extraction problem: in fact, letting q_t be the nominal income, each firm i can observe in period t only the private signal $z_t(i)$ of the form:

$$z_t(i) = q_t + v_t(i)$$

where $v_t(i)$ is an idiosyncratic noise term. Woodford uses this model to study the impulse response functions of inflation and output, and finds that it can replicate data patterns better than the Calvo approach, at least for a reasonable region of the parameters' space.

One of the most common ways to model the economic behavior in a world characterized by constraints on the information sets and bounded rationality is the adaptive learning[7]. This approach typically deals with agents that does not know the "correct" model of the world, but use data to estimate it like an econometrician would do. Following this line of reasoning, the most natural way to model how the individuals do their estimations is to assume that they have a mental model of the law of motion of the relevant variables in the economy (the "perceived law of motion", or PLM), and that they estimate its parameters via OLS[8]. Given these estimates, the endogenous variables will follow what is called the "actual law of motion" (ALM). Calling ϕ_t the N-dimensional vector of parameters estimates, R_t the matrix of its second moments, z_t the regressors and p_t the endogenous variables, it can be shown that, with the appropriate initial conditions, the updating algorithm:

$$\begin{aligned} \phi_t &= \phi_{t-1} + t^{-1} R_{t-1}^{-1} z_{t-1}(p_t' - z_{t-1}' \phi_{t-1}) \\ R_t &= R_{t-1} + t^{-1}(z_{t-1} z_{t-1}' - R_{t-1}) \end{aligned} \tag{1}$$

delivers the same sequence of estimates $\{\phi_t\}_{t=0}^{\infty}$ as the standard OLS techniques applied period by period; for this reason, it is called recursive least squares algorithm (RLS). Given these estimates, the ALM (in case of a linear model) is:

$$p_t = T^{'}(\phi_{t-1}) z_{t-1} + \varepsilon_t$$

where the function $T : \Re^N \to \Re^N$ is the mapping from the estimated coefficients to the actual coefficients determined by the estimates, and is ε_t a white noise. The appealing feature of this formulation is that it can be studied with the tools of stochastic approximation[9], with the result that for large classes of models the asymptotic dynamics are governed by the stability properties of a deterministic ordinary differential equation; in particular we have that a rational expectations equilibrium is (locally) stable under adaptive learning if it is a (locally) stable

[7] For extensive monographs, see Sargent (1993) and Evans and Honkapohja (2001).

[8] Note that the PLM may or may not be of the same functional form of the rational expectations solution of the model.

[9] For an extensive monograph on stochastic approximation, see Benveniste et al. (1990).

rest point of the ordinary differential equation[10][11]:

$$\frac{d\phi}{d\tau} = T(\phi) - \phi$$

Since a rational expectation equilibrium is a rest point of the T-mapping, RLS learning has often been invoked to argue that assuming rational expectations is not a too restrictive hypothesis (at least in the limit), and a lot of effort has been devoted to study the learnability of equilibriums in widely used economic models.

There are also other adaptive algorithms employed in the literature; one of the most used alternatives is to substitute in (1) the factor t^{-1} with a constant gain (or tracking parameter) $0 < \gamma < 1$. Since in this case the estimates always react to any new shock (also asymptotically), the system never converges to a fixed value, but under some technical conditions it can settle down as a normal distribution, whose support shrinks to zero as γ approaches zero[12].

There have been some effort in trying to reconcile inflation data with monetary models applying the adaptive learning techniques. In Orphanides and Williams (2002), the authors assume an ad hoc model with a Phillips curve which includes a lag of inflation as an explanatory variable, and a demand relation that expresses output gap as a function of the real interest rate deviation from its equilibrium value, and study the design of the optimal policy in a setting of adaptive learning through a constant gain algorithm.

More closely related to this work is the paper of Williams (2003), who introduce adaptive learning (both in the RLS and the constant gain versions) in a standard New Keynesian framework with monopolistic competition and staggered prices à la Calvo. He analyzes to what extent the introduction of adaptive learning matters for business cycle statistics; he found that, quantitatively, this change is of second-order importance. However, in his model he still assumes that the agents optimize taking into account the Calvo constraint on the pricing resetting possibility.

Moreover, Sargent (1999) shows how a mispecification by the policymakers of the "true" structural relations of the economy, coupled with constant gain learning, may lead to a system that oscillates most of the time around the high inflation equilibrium (the time-consistent one, according to Barro and Gordon (1983)), occationally moving towards the low inflation time-inconsistent (or Ramsey) equilibrium, when a suitable sequence of shocks occurs.

We aim to do a first step towards the construction of a bridge between the adaptive learning literature and the other limited information approaches, in assuming no price stickiness, and instead taking the imperfect information as the main source of inertia in the model. We will therefore introduce adaptive learning in a monopolistic competitive, flexible prices setting, with an exogenous

[10] In the adaptive learning terminology, an equilibrium which is a stable solution of the differential equation reported below is defined an E-stable equilibrium.

[11] For the derivation of this result, see Marcet and Sargent (1989) and Evans and Honkapohja (2001), Chapter 6 and 8.

[12] For more on constant gain, see below Section 3.1.

process for nominal output, and compare its performance to that of an analogous model where firms behave according to Calvo model.

3 The Model

The production side is characterized by a continuum of firms that produce differenciated goods in a competitive monopolistic framework. The demand side is characterized by a representative consumer with rational expectations, who solves the following problem:

$$\max E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, N_t) \tag{2}$$

$$s.t. \int_o^1 P_{it} C_{it} di + M_t = M_{t-1} + W_t N_t + \int_0^1 \Pi_{it} di - T_t$$

where C_{it} and P_{it} denotes the demand of good i and its price, respectively, M_t is the money stock hold at the end of period t, T_t are transfers from government, N_t is labor supply, W_t is nominal wage, Π_{it} is the profit from the sell of good i[13], and C_t represents the CES aggregate of consumption:

$$C_t = \left[\int_0^1 C_{it}^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}}$$

As shown in Dixit and Stigliz (1977), maximizing the CES index of consumption subject to a certain level of overall expenditure leads to the demand schedule for good i:

$$C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\theta} C_t$$

where:

$$P_t = \left[\int_0^1 P_{it}^{1-\theta} di \right]^{\frac{1}{1-\theta}}$$

Moreover, consumer faces a cash-in-advance constraint[14]:

$$P_t C_t \leq M_t$$

which is assumed to be binding. Doing so, we can close the model with the simplest possible specification for the demand side of the economy, i.e. the quantity theory. This is different from the specification used as a standard framework of monetary policy evaluation, which derives an IS relationship from a money-in-utility setup; otherwise, it is useful to simplify the analysis in this early stage, and such a simplifying assumption has been already used in many

[13] We are assuming that firms are owned by the representative consumer.

[14] Note that, in writing this constraint, we are implicitly assuming that at time t the money market closes before the opening of commodity markets.

of the papers that propose alternatives to the Calvo model[15]. The extention of this approach to a more standard specification of the demand side of the economy is left as future work.

Assuming separability of the $U(\cdot)$ function between its arguments, it is easily shown that:

$$\frac{W_t}{P_t} = -\frac{U_N(N_t)}{U_C(C_t)} \equiv G(C_t, N_t)$$

We assume that each firm produce a differenciated good according to the strictly increasing and concave production function:

$$Y_{it} = A_t F(N_{it})$$

where A_t denotes a technology indicator, and N_{it} is firm i labour demand, and Y_{it} good i output. In equilibrium, market clearing implies:

$$Y_{it} = C_{it}, \quad Y_t = C_t$$

so that the demand schedule can be rewritten (in logs) as:

$$Y_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\theta} Y_t \tag{3}$$

To model firms' behaviour, we make the same distinction between the optimization and the inference problem underlined in Townsend (1983). In particular we assume:

- firms can freely reset prices in each period, but they can observe aggregate variables only with a one period delay; so, in each period firm i's optimization problem is given by the static expected profit maximization:

$$\max_{P_{it}} E_t^i \left[P_{it} Y_{it} - W_t N_{it}\right] \tag{4}$$

where $E_t^i x_t$ is firm i's (in general non rational) expectations of x_t, formed using time t information set. This maximization is done subject to a perceived demand schedule:

$$Y_{it} = \left(\frac{P_{it}}{E_t^i P_t}\right)^{-\theta} E_t^i Y_t \tag{5}$$

As is shown in Woodford (2002), maximizing (4) subject to (5) yields:

$$\frac{P_{it}}{E_t^i P_t} = \frac{\theta}{\theta - 1} CO^i(Y_{it}, Y_t)$$

where CO^i denotes the cost function of firm i, and x_z indicates the derivative of x with respect to z. Loglinearizing this expression around the full-information equilibrium $\frac{P_{it}}{E_t^i P_t} = 1$, $Y_t = Y_t^N$ (where Y_t^N represents natural output) yields:

$$p_{it} = E_t^i p_t + \xi E_t^i y_t$$

[15] It is common to Mankiw and Reis (2002), Ball et al. (2003), and Woodford (2001); in a learning framework, a similar assumption has been used by Adam (2003).

where $x_t \equiv \log X_t$, and $y_t = \log(Y_t/Y_t^N)$, and ξ is defined as:

$$\xi \equiv \frac{\varepsilon^{FI}_{CO^i_{Y_{it}},Y_{it}} + \varepsilon^{FI}_{CO^i_{Y_{it}},Y_t}}{1 + \theta\varepsilon^{FI}_{CO^i_{Y_{it}},Y_{it}}}$$

where $\varepsilon^{FI}_{x,z}$ denotes the elasticity of x with respect to z evaluated in the full-information equilibrium. Using (5) to substitute for $E^i_t y_t$, we get:

$$p_{it} = E^i_t p_t + \frac{\xi}{1-\theta\xi} y_{it}$$

Integrating over i, and assuming homogeneous expectations, we obtain:

$$p_t = E^*_t p_t + \frac{\xi}{1-\theta\xi} y_t \tag{6}$$

Equation (6) can be rewritten as:

$$p_t = \frac{1-\xi\theta}{1+\xi(1-\theta)} E^*_t p_t + \frac{\xi}{1+\xi(1-\theta)} q_t \tag{7}$$

where $q_t = y_t + p_t$. To close the model, we need a process for nominal output[16], and a rule for expectations formation; the first one is given by an AR(1) process for the growth rate of nominal output[17]:

$$\Delta q_t = (1-\rho)g + \rho\Delta q_{t-1} + u_t \tag{8}$$

where u_t is an i.i.d. shock. Substituting out q_t from (7) using (8)[18], we get:

$$\begin{aligned} p_t \quad = \quad & \frac{1-\xi\theta}{1+\xi(1-\theta)} E^*_t p_t + \frac{\xi(1-\rho)}{1+\xi(1-\theta)} g + \frac{\xi(1+\rho)}{1+\xi(1-\theta)} q_{t-1} \\ & - \frac{\xi\rho}{1+\xi(1-\theta)} q_{t-2} + \frac{\xi}{1+\xi(1-\theta)} u_t \end{aligned} \tag{9}$$

It can be easily shown that the minimum state variable (MSV) solution of the model given by (8)-(9) under rational expectations is[19]:

$$p_t = (1-\rho)g + (1+\rho)q_{t-1} - \rho q_{t-2} + \frac{\xi}{1+\xi(1-\theta)} u_t$$

[16]Because of the cash-in-advance constraint, money and nominal output are equivalent.

[17]See for analogous processes Woodford (2001), or Mankiw and Reis (2001). Moreover, Christiano et al. (1998) argue that an AR(1) process for growth rate of money is empirically plausible.

[18]Since agents do not observe contemporaneous nominal output, we write the law of motion for price level in terms of observable variables.

[19]I.e., $E^*_t = E_t$.

- The inference problem is modeled following the literature of adaptive learning[20]. In particular, we assume that agents do not know the exact MSV solution of the model given by (8)-(9) but, instead, form their expectations according to the perceived law of motion (PLM):

$$p_t = a_{t-1} + b_{t-1}q_{t-1} + c_{t-1}q_{t-2} + \eta_t \qquad (10)$$

where the vector $\phi_{t-1} \equiv (a_{t-1}, b_{t-1}, c_{t-1})'$ denotes the estimates of model parameters computed by agents using information available on aggregate variables at time t[21]; these estimates are updated according to the recursive algorithm:

$$\begin{aligned} \phi_t &= \phi_{t-1} + \gamma_t R_{t-1}^{-1} z_{t-1}(p_t - \phi'_{t-1} z_{t-1}) \\ R_t &= R_{t-1} + \gamma_t (z_{t-1} z'_{t-1} - R_{t-1}) \end{aligned} \qquad (11)$$

where $z_t = (1, q_t, q_{t-1})'$ and $\{\gamma_t\}$ is a sequence of nonincreasing values called "gain parameters". The precise path followed by this sequence will be described in the next subsection. Equation (10) implies that:

$$E_t^* p_t = a_{t-1} + b_{t-1}q_{t-1} + c_{t-1}q_{t-2} \qquad (12)$$

which can be plugged into (9) to obtain the actual law of motion (ALM):

$$\begin{aligned} p_t = {} & \left(\frac{1-\xi\theta}{1+\xi(1-\theta)} a_{t-1} + \frac{\xi(1-\rho)}{1+\xi(1-\theta)} g \right) + \left(\frac{1-\xi\theta}{1+\xi(1-\theta)} b_{t-1} + \right. \\ & \left. \frac{\xi(1+\rho)}{1+\xi(1-\theta)} \right) q_{t-1} + \left(\frac{1-\xi\theta}{1+\xi(1-\theta)} c_{t-1} - \frac{\xi\rho}{1+\xi(1-\theta)} \right) q_{t-2} \\ & + \frac{\xi}{1+\xi(1-\theta)} u_t \end{aligned} \qquad (13)$$

Given (13), the T-mapping is:

$$\begin{aligned} T(a) &= \frac{1-\xi\theta}{1+\xi(1-\theta)} a + \frac{\xi(1-\rho)}{1+\xi(1-\theta)} g \\ T(b) &= \frac{1-\xi\theta}{1+\xi(1-\theta)} b + \frac{\xi(1+\rho)}{1+\xi(1-\theta)} \\ T(c) &= \frac{1-\xi\theta}{1+\xi(1-\theta)} c - \frac{\xi\rho}{1+\xi(1-\theta)} \end{aligned} \qquad (14)$$

which can be easily shown to imply that the MSV solution under rational expectations is E-stable.

Equation (13), together with the process for nominal output (8) and the stochastic recursive algorithm (11), constitutes our model.

[20] See Evans and Honkapohja (2001).

[21] I.e., the sequence $\{p_i, q_i, q_{i-1}\}_{i=1}^{t-1}$.

3.1 The Gain Parameter

At this stage of the analysis, it is necessary to be explicit about the form taken by the γ_t's in (11). In the learning literature there are two main strategies used to calibrate them, corresponding to two different hypothesis of how the agents conceive the world:

- decreasing gain: assume that it is a decreasing sequence with the properties that $\sum_{t=0}^{\infty} \gamma_t = \infty$ and $\sum_{t=0}^{\infty} \gamma_t^2 < \infty$, as we would obtain setting $\gamma_t = t^{-1}$. In this case, equation (11) becomes a particular case of the RLS algorithm (1), assigning the same weight to every observation. This last remark means that a decreasing gain is a reasonable assumption if agents think that the model's parameters are constant over time, so that each observation has the same information content;
- constant gain: assume that $\gamma_t = \gamma$, where γ is a small positive constant. In this case, our algorithm does not deliver the same estimates as the OLS anymore, since past data are downweighted. The assumption behind this behavior is that agents believe structural changes to occur, even if they are able neither to model them nor to predict in which period they will take place. As a result, they will update their estimates given their belief that more recent data embed more information on the structure of the economy. As mentioned in Section 2, this specification of the gain sequence prevents the parameters' estimates to converge to any particular value, since they will be significantly influenced by any new shock.

In what follows, we will employ the constant gain specification; before proceeding, however, there are two logical problems that we have to take into account.

First of all, in the standard adaptive learning setup, the agents take their decisions treating their expectations as if they correspond to the "true" model. If this kind of behavior can be justified in a decreasing gain case[22], it seems to contrast with the basic assumption that motivates constant gain, i.e. that agents are convinced that economic structure shifts over time. In Tetlow and von zur Muehlen (2001) this issue is investigated in the context of the Sargent (1999) model; they allow policy makers to make their decisions taking into account model uncertainty in a Bayesian way[23]. Their conclusion is that this method does not yields results quantitatively different from the Sargent's more standard approach.

Another relevant question that we now should address is how to calibrate γ. In particular, this choice is potentially subject to a high degree of arbitrariness,

[22] Actually, the fact that the ALM has time varying parametrs, can arise misspecification issues in a context of decreasing gain learning, as noted in Bray and Savin (1986); for a time-varying parameters estimation approach in an adaptive learning approach, see McGough (2000).

[23] In other words, taking into account the standard errors of the estimates.

that could allow the model to replicate any empirical pattern we want: in fact, as pointed out in Marcet and Nicolini (1998), the presence of too many degrees of freedom in designing the learning algorithms has always made this class of models hardly falsifiable, hence preventing many researchers from using them to match data. The introduction in the setup of an additional exogenous parameter, that could in principle be used to make the model behave as we want, makes this criticism particularly sound, and the "wilderness of irrationality" particularly dangerous.

A possible way out of both these pitfalls is to endogenize the value of γ, making it the outcome of some optimal choice of the agents. The way we have decided to do it is using standard game-theoretic concepts. We assume that firms are interested in minimizing the mean square error (MSE) of their forecasted inflation; in this case, we can define a *misspecified equilibrium* as a value γ^* of the gain parameter which minimizes the MSE of an individual, when all the rest of the economy update its expectations using the same γ^*. More formally, we look for a fixed point of the function:

$$\gamma^* = \arg\min_{\gamma} \frac{1}{T} \sum_{t=1}^{T} \left[p_t(\widehat{\gamma}) - E_t^* p_t(\gamma)\right]^2 \equiv f(\widehat{\gamma}) \tag{15}$$

where T is the time horizon taken into account (in our case, 100 periods). This is an *equilibrium* in the sense that no agent has an incentive to deviate from this strategy, and is *misspecified* in the sense that agents behave as if they perceive the economy as a time varying parameters system, when the "true" model is characterized by constant structural parameters[24]. This approach is closely related to the concept of equilibrium in learning rules of Evans and Honkapohja (1993), and to the internal consistency requirement introduced in Marcet and Nicolini (1998).

Note that, in this way, the tracking parameter is not freely chosen anymore, but becomes a function of the other structural parameters of the economy, solving the second problem above mentioned; moreover, it implies that even if an individual agent does not think that the model is really time-varying, he will play according to the misspecified equilibrium if he thinks that the other agents will do the same[25]. Hence, this kind of approach allows us to deal also with the first problem that we discussed above[26]. If we model the choice of the gain parameter in this way, standard issues of how the agents could end up coordinating on a certain equilibrium (even with bounded rationality) arise, but they are far beyond the scope of this paper.

As an additional remark, observe that in Orphanides and Williams (2003), where a constant gain algorithm is implemented, γ is left as a free parameter, and the behavior of the model for different values of it is studied.

[24]Note that, as we point out below, it is not necessary that each agent believes the economy to be time varying, but only that in his opinion all the other agents have this belief.

[25]This is a typical coordination problem.

[26]Notice that some kind of myopic behavior from the agents' side must be assumed since, even if playing γ^* minimize the MSE in finite time, it makes convergence to the rational expectations equilibrium impossible.

Moreover, a technical remark on the asymptotic behavior of parameters' estimates under constant gain learning is to be done: as above mentioned, under certain conditions, they will converge in distribution to a normal with an E-stable equilibrium as a mean; unfortunately, in this model one of the sufficient conditions required to obtain this result is not satisfied[27]. However, very long simulations of the model (10000 periods) show that the parameters' estimates really converge to a neighborhood of the rational expectations values after a few periods (less than the fifty that we usually discard at the beginning of each simulation), around which they keep oscillating.

It is interesting to note that the problem that is behind the construction of the best-reply function $f(\bullet)$ (i.e., the minimization of the MSE with respect to the constant gain used by an individual that cannot influence the non-stationary process that is trying to track) is similar to the framework analyzed in Chapter 4, Part I of Benveniste et al. (1990). In that context, the authors study how to derive analytically the value of the tracking parameter that minimizes the expected value of the square of the distance between the sequence of the actual values of the time-varying coefficients of the process to track, and the estimated values of these coefficients. First of all, they decompose the objective function into the sum of the distance between the mean dynamics of the two sequences, and the variance of their distance. Moreover, they show that the optimal choice of γ is the result of a compromise between *tracking* and *accuracy*: in other words, higher (lower) values of the gain parameter reduce (increase) the distance between the mean dynamics of the two sequences, thus reducing (increasing) the magnitude of the MSE, and increase (reduce) the variance of the distance between the two sequences, hence increasing (reducing) the MSE.

These results cannot be directly applied to our model for the same reason we outlined above, when we talked about the asymptotic behavior of parameters' estimates. However, these tools provide useful insights on how our model will behave: in fact, we observe that an increase (decrease) of $\widehat{\gamma}$, on one hand, does not influence the mean dynamics of the time-varying coefficients of p_t (which are given by the rational expectations values, as mentioned above), while on the other hand it increases (decreases) the variance of these coefficients. Loosely applying the concepts sketched above, the best-reply of the individual firm would be to reduce γ^*; hence, we can expect $f(\bullet)$ to be a decreasing function. And this is exactly what we obtained, when we computed a numerical approximation of $f(\bullet)$[28].

4 Empirical Results

4.1 Calibration Strategy

We need to calibrate five parameters (ξ, θ, ρ, g), plus the initial conditions ϕ_0 for the RLS algorithm; the chosen values are summarized in Table 1.

[27] In particular, the law of motion of the state variables $[q_{t-1}, q_{t-2}]$ contains a unit root.

[28] See next section.

ξ	θ	ρ	g	a_0	b_0	c_0
0.15	6	0.7	0	$(1-\rho)g$	0.05	-0.05

Table 1: Baseline calibration

The parameter γ is a function of these other seven coefficients, as clarified in the previous section.

The value for ξ is suggested in Woodford (2002) as an empirically plausible value for U.S. economy, and is used also in Woodford (2001); it also lies in the range of values examined in Mankiw and Reis (2001). θ is chosen according to the standard New Keynesian literature, while ρ is chosen to match empirical facts (see below). The choice to set the trend g to zero is justified by the aim to concentrate on the model's behaviour at business cycle frequency.

We now need to be more explicit on the strategy adopted to set the value of the tracking parameter. First of all, note that neither a formal proof of the existence and the uniqueness of a fixed point of the expression (15)[29], nor an analytical expression of γ^* as a function of the parameters' vector $(\xi, \theta, \rho, g, a_0, b_0, c_0)'$ have been obtained; hence, we had to search for a numerical approximation of this equilibrium[30]. We used the following procedure: first of all, we set up a grid of 19 values (0.05, 0.1, 0.15,...,0.95), then we draw 1000 realizations of the nominal output shock $\{u_i\}_{i=1}^{150}$; then, we computed the corresponding sequences $\{p_i\}_{i=1}^{150}$ using a fixed value of $\widehat{\gamma}$. Then, we throw away the first fifty values of each sequence, to dampen the influence of initial conditions, compute the empirical MSE for each $\{p_i\}$ and for each possible γ, average across all realization, and look for the γ for which the resulting value is minimum. This procedure has been repeated for all the 19 possible values of $\widehat{\gamma}$, and we got that the only fixed point is at $\gamma^* = 0.55$.

Note that this value is much higher than those used in Orphanides and Williams (2003)[31]; in Figure 1 we have plotted the evolution over time of the parameters' estimates in one of the 10000 stochastic simulations that we performed to analyze the behavior of our model (see subsections below). As we can observe in Figure 1, we have a very rapid convergence to a neighborhood of the rational expectations equilibrium, followed by wide oscillations around it. This last feature is due to the very high value of the tracking parameter, which makes the estimates of the parameters very sensible to any new shock.

To conclude, the expectations has been initialized at values that have the same sign as the rational expectations parameters, but that deliver the desired hump-shaped impulse response for inflation; for values closer than those to the rational expectations, would yield a peak response only two periods after the nominal shock.

[29] As is instead derived, in a simpler context, in Evans and Ramey (2001).

[30] However, since the values obtained numerically for γ^* are a monotonic decreasing function of $\widehat{\gamma}$, as explained in the previous section, this makes us conjecture that the "true" equilibrium exists and is unique.

[31] The largest value of the tracking parameter that they adopt is 0.1.

4.2 Persistence of Monetary Shocks

A well-known shortcoming of the Calvo staggered price framework is that, even if prices responds sluggishly to monetary shocks, inflation does not: the highest effect of the shock is experienced in the first period, and then it monotonically decays.

In a framework similar to the one developed in the previous section, this drawback of the Calvo model can be easily seen starting from the well known New Keynesian Phillips Curve[32]:

$$\pi_t = ky_t + \beta E_t \pi_{t+1} \tag{16}$$

where β denotes the subjective discount rate, $E_t[\cdot]$ is the rational expectations operator, and k is a function of the parameters ξ, β and α, where the latter indicates the probability that a given firm does not review its price in a given period; the exact form of this function of parameters is not important for our purposes: the only relevant point is that it assumes positive values, as shown in Woodford (2001).With a nominal demand process like (8), it can be shown that inflation can be written in terms of an MA(∞) process, whose coefficients are a monotonic decreasing sequence, independently of the values of the parameters[33]; this result has an immediate consequence in terms of the impulse response function: as claimed before, the peak effect of a monetary shock is reached on impact, and the inflation returns monotonically towards the pre-shock value.

This is at odds with what is empirically observed, and with what is considered conventional wisdom; in fact, there is an extensive literature[34] that stresses the fact that the maximum effect of a monetary shock is reached between 1 and 2 years after the impact of the shock.

As shown in Figure 2, with a value of ρ of 0.7, our model is able to generate some persistence of the nominal shock; in particular, the peak effect of the shock is reached after three periods. Even if it is not as much as in data, nevertheless it is a better performance than the Calvo model. The reason is straightforward: since agents do not know the exact model, when they observe a discrepancy between the actual and the forecasted inflation, they are not aware how much of it is due to the presence of a shock, and how much to an imprecise estimate of the parameters; hence, they react with more caution than what would be optimal for a Calvo-agent, and smooth their reaction over more than one period. Formally, we can observe from equation (7):

$$\Delta p_t \equiv \pi_t = \frac{1-\xi\theta}{1+\xi(1-\theta)}\left(E_t^* p_t - E_{t-1}^* p_{t-1}\right) + \frac{\xi}{1+\xi(1-\theta)}\Delta q_t$$

which can be rewritten as:

$$\pi_t = \frac{1-\xi\theta}{1+\xi(1-\theta)}\left(E_t^* \pi_t - E_{t-1}^* \pi_{t-1}\right) + \frac{1-\xi\theta}{1+\xi(1-\theta)}\pi_{t-1} + \frac{\xi}{1+\xi(1-\theta)}\Delta q_t \tag{17}$$

[32]For the derivation, see for e.g. Woodford (2001).

[33]For the proof, see the Appendix.

[34]See, e.g. Adam (2003) and Christiano et al. (2001).

This equation makes explicit the fact that adaptive learning has introduced a backward-looking element in the inflation dynamics; this element would disappear only in the case in which the agent comes from a period of no forecasting error (i.e., in the case of $\pi_{t-1} = E^*_{t-1}\pi_{t-1}$).

Even if the model can generate more realistic dynamics, it still has to be improved to match data better. Moreover the impulse response of Figure 2 is clearly dependent on the initial conditions; so we could consider different initial conditions, like extracting them from an assumed prior distribution, or introducing a "training period"[35], i.e. a period during which the economy is run in the rational expectations equilibrium, and after which the agents use OLS to estimate the coefficients; this estimates would constitute the initial conditions for the impulse response exercise. Unfortunately the model has problems in replicating an enough hump-shaped impulse response function for inflation with initial conditions too close to the rational expectations equilibrium, since the backward-looking element above mentioned is too weak to ensure enough delay in the inflation response to the shock. However, at least two periods of delay in the peak effect of the shock can be obtained for a wide range of initial conditions.

We are also interested in checking whether the inflation inertia generated by adaptive learning (and captured by equation (17)) is enough to match empirical data on first order autocorrelation of inflation; in particular, we consider the value of 0.76 reported in Mankiw and Reis (2002), who compute first order autocorrelation of the CPI using Hodrick-Prescott filtered U.S. data. Since the model cannot be solved analytically, the population value of any statistic is not available. Hence, the only way is to perform stochastic simulations, and then to compare the value of the statistic for this simulated sample with the empirical one. In doing so, a standard problem arises: since we want to assess how "close" is our model to reality, we need a metric. Our choice, following the approach outlined in Canova and Ortega (2000), is to perform 10000 stochastic simulations of the model, reporting not only the mean of the simulated statistic of interest (i.e., the first order autocorrelation of inflation), as is usually done, but to use information on all the simulated distribution. In other words, we check whether the actual value lies between the 5^{TH} and the 95^{TH} percentiles[36]. It turns out that 0.76 almost coincides with the 95^{TH} percentile of the empirical distribution, denoting the capacity of adaptive learning to generate a realistic degree of inflation inertia, even without any other source of rigidity.

This result is more remarkable, if we take into account the extremely high value of γ we are using: in fact, we would expect that a learning scheme so sensitive to every forecast error would induce a more erratic behavior of inflation expectations, hence dampening the possibility of the model of replicating the empirical value of inflation autocorrelation. More properly, we can think about increasing the value of γ as a trade-off: in fact, a value too small of the tracking

[35] See Williams (2003).

[36] In a sense, we are taking our model as a null hypothesis; for a survey of these techniques to evaluate calibrated dynamic stochastic general equilibrium models, see Canova and Ortega (2000).

parameter would nullify the role of learning[37], while a value too large would make the expectations "overreact", introducing noise in the inflation process. To confirm this conjecture, we performed stochastic simulations of the model for different values of the constant gain, and for each of them we computed the median and the 95^{TH} percentile, plotting the results in Figure 4. A quick inspection of the figure shows that the highest values of both are reached when $\gamma = 0.05$; for every level of the tracking larger than this, the first order autocorrelation generated by the model starts declining monotonically. In Figure 5 we have plotted the evolution over time of the parameters' estimates of one of the simulations conducted with $\gamma = 0.05$; it is evident just by comparing Figure 5 and Figure 1 how the smaller value of the constant gain let the estimates vary in smoother waves than those observed for $\gamma = 0.55$, when we had a very "nervous" behavior. Moreover, the estimates of the coefficients on q_{t-1} and q_{t-2} tend to change in a trascurable way and to keep close to the rational expectations equilibrium value, when agents have had enough time to learn.

The pattern displayed in Figure 4 could help explaining why Orphanides and Williams (2002) observe that inflation persistence is *increasing* in the value of γ: actually, they assumed magnitudes of the gain parameter so small[38] that their simulations moved along the increasing side of Figure 4. Actually, we see that, setting exogenously the constant gain[39] would have allowed us to obtain a better performance in replicating the empirical patterns.

The fact that our approach was able to generate a first order autocorrelation of inflation broadly consistent with data, even with an endogenously determined value for the constant gain, strengthen the result.

5 Acceleration Phenomenon

The hump-shaped impulse response function of inflation is not the only empirical pattern which is difficult to reconcile with the Calvo model; another example is the widely documented positive and significant correlation between the level of real output and the growth rate of inflation[40]. This same pattern is reported also in Mankiw and Reis (2002), who computed this correlation for Hodrick-Prescott filtered U.S. data for CPI inflation, obtaining a value of 0.38. On the other hand, the Calvo model does not exhibit this pattern; the main reason is the interaction between a monotonically decreasing impulse response of inflation, and a positive response of real output (at least in the short run) to the nominal shock. These two features generate (after a positive shock) the contemporaneous presence for many periods of decreasing inflation and high output, thus explaining the negative correlation. To check formally this intuition, Mankiw and Reis (2002) calculated the population cross-correlation $corr(y_t, \pi_{t+2} - \pi_{t-2})$ for a Calvo-type

[37] In the limit, for $\gamma = 0$, we would have $\phi_t = \phi_{t-1}$.
[38] The largest value they assume is 0.1.
[39] In particular, giving it a value of 0.05.
[40] See, e.g. Abel and Bernanke (1998) and Blanchard (2000), who use a scatterplot to document this phenomenon.

staggered price model for a wide range of values of the key parameters, always obtaining negative numbers.

To check whether our model can instead match this feature of the data, we followed a procedure analogous to that outlined in the previous subsection, performing 10000 stochastic simulations, and then computing the simulated distribution of $corr(y_t, \pi_{t+2} - \pi_{t-2})$ (which is plotted in Figure 5); the mean of this distribution is 0.4, very "close" to the actual value of 0.38. To evaluate how "close" it is, we need a metric; we choose to check whether 0.38 lies between the 5^{TH} and the 95^{TH} percentiles of the simulated distribution. These percentiles are equal to 0.24 and 0.54, which implies that our model is broadly consistent with this empirical pattern. Moreover, we obtained that no one of the draws has delivered a negative value of $corr(y_t, \pi_{t+2} - \pi_{t-2})$[41], denoting a stark difference with respect to the Calvo model.

6 An Alternative Specification

The previous results have been obtained assuming that firms form their expectations on current prices regressing p on nominal output lagged of one and two periods (plus a constant). It seems reasonable, since individual profits depend on the difference between individual and aggregate price level, so that we can expect agents to estimate the law of motion of the relevant aggregate variable (i.e., the price level), and then forecasting its current value.

As a robustness check, we will try also a different approach. In particular, we want to check whether the previous results[42] have been driven by the fact that the price level is a nonstationary variable; hence, we will assume now that firms estimate the law of motion of the inflation rate (which is stationary), and that they forecast its current level using the most recent estimates of this law of motion; then, given the value of p_{t-1} (which is known at time t) and the identity:

$$E_t^* \pi_t \equiv E_t^* p_t - p_{t-1}$$

they obtain $E_t^* p_t$, which is used in their decision process. To derive the PLM, note that from equation (9) we get:

$$\begin{aligned}
\pi_t &= \frac{1-\xi\theta}{1+\xi(1-\theta)} E_t^* p_t + \frac{\xi(1-\rho)}{1+\xi(1-\theta)} g + \frac{\xi(1+\rho)}{1+\xi(1-\theta)} q_{t-1} \\
&\quad - \frac{\xi\rho}{1+\xi(1-\theta)} q_{t-2} + \frac{\xi}{1+\xi(1-\theta)} u_t - p_{t-1} \\
&= \frac{1-\xi\theta}{1+\xi(1-\theta)} E_t^* \pi_t + \frac{\xi(1-\rho)}{1+\xi(1-\theta)} g + \frac{\xi(1+\rho)}{1+\xi(1-\theta)} q_{t-1} \\
&\quad - \frac{\xi\rho}{1+\xi(1-\theta)} q_{t-2} - \frac{\xi}{1+\xi(1-\theta)} p_{t-1} + \frac{\xi}{1+\xi(1-\theta)} u_t
\end{aligned}$$

[41]Note that, on the contrary of what is observed for inflation persistence, the acceleration phenomenon is a robust feature of the model, since it holds for a wide range of values of ρ.

[42]In particular, the very high value of the tracking parameter.

which implies a PLM of the form:

$$\pi_t = \tilde{a}_{t-1} + \tilde{b}_{t-1} q_{t-1} + \tilde{c}_{t-1} q_{t-2} + \tilde{d}_{t-1} p_{t-1} + \eta_t$$

and the ALM:

$$\begin{aligned}\pi_t \quad = \quad & \left(\frac{1-\xi\theta}{1+\xi(1-\theta)}\tilde{a}_{t-1} + \frac{\xi(1-\rho)}{1+\xi(1-\theta)}g\right) + \left(\frac{1-\xi\theta}{1+\xi(1-\theta)}\tilde{b}_{t-1} + \right. \\ & \left.\frac{\xi(1+\rho)}{1+\xi(1-\theta)}\right) q_{t-1} + \left(\frac{1-\xi\theta}{1+\xi(1-\theta)}\tilde{c}_{t-1} - \frac{\xi\rho}{1+\xi(1-\theta)}\right) q_{t-2} \\ & \left(\frac{1-\xi\theta}{1+\xi(1-\theta)}\tilde{d}_{t-1} - \frac{\xi}{1+\xi(1-\theta)}\right) p_{t-1} + \frac{\xi}{1+\xi(1-\theta)} u_t\end{aligned} \tag{18}$$

It is easy to see that this model has a unique rational expectations equilibrium given by:

$$\pi_t = (1-\rho)g + (1+\rho)q_{t-1} - \rho q_{t-2} - p_{t-1} + \frac{\xi}{1+\xi(1-\theta)} u_t$$

which is equivalent to say:

$$p_t = (1-\rho)g + (1+\rho)q_{t-1} - \rho q_{t-2} + \frac{\xi}{1+\xi(1-\theta)} u_t$$

In other words, the two forecasting strategy are equivalent under rational expectations.

To see what changes under constant gain learning, we repeated the same excercises described in the previous section. First of all, we looked for the equilibrium value of the tracking parameter in this new context, obtaining $\gamma^* = 0.15$. A couple of remarks are now necessary:

- this value is considerably lower than the 0.55 obtained in the previous section. The reasons why it is so are still an open question. A possible explanation would be that now agents are trying to forecast a stationary variable;
- the best reply function is monotonically decreasing also under this alternative specification, making us conjecture that this is a common feature of this kind of models.

Even with these different assumptions on how firms forecast the aggregate level of prices, the adaptive learning approach does well in accounting for inflation dynamics[43]. In fact, setting the initial condition of $\tilde{d}$ in a small enough neighborhood of the rational expectations value[44], the shape of the impulse response function is analogous to the one obtained previously. Moreover, the

[43] In all the simulations, we used the same set of realizations of the shock that we used in the baseline case, in order to make the comparison consistent.

[44] In particular, for $\tilde{d} \in (-0.6, -1.5)$.

5^{TH} and the 95^{TH} percentiles of the simulated distribution of the first order autocorrelation of inflation are 0.57 and 0.79, respectively, so that they include the actual value of 0.76.

Also the correlation between the level of real output and the growth rate of inflation implied by the model is consistent with the one empirically observed: in fact, the 5^{TH} and the 95^{TH} percentiles of the simulated distribution are 0.18 and 0.46, respectively, so that 0.38 is included between them, but zero is not.

7 Conclusions and Future Research

The starting point of this paper is the difficulty of the Calvo model to replicate some well-established empirical facts. In particular, this model *per se* is not able to generate an hump-shaped impulse response function for inflation, nor a positive correlation between real output and inflation growth. To reconcile this approach with empirics, additional sources of inertia have been introduced by the New Keynesian literature.

However, a new line of research has been recently developed, whose key points are flexible prices coupled with some form of boundedly rational behavior; this paper shares this modeling strategy, and aims to investigate the properties of a simple flexible prices, monopolistic competitive setup augmented by non-rational expectations, modeled following the adaptive learning approach.

The main result is that, with reasonable parameters values, this setup can considerably improve the performance of the Calvo model, generating inflation and output dynamics that are broadly consistent with the two stylized facts above mentioned; moreover, also the inflation autocorrelation is not at odds with what is empirically observed.

As a side issue, we studied the relationship between the constant gain and inflation autocorrelation to show how, keeping this parameter free to assume any value, we could make the model match almost any empirical pattern, hence stressing the importance of endogenizing this coefficient, linking it to some optimal behavior of the agents.

As future research, we could possibly extend the model modifying the demand side under two respects: first of all, we could make the consumer's problem fully dynamic, dispensing with the cash-in-advance constraint and introducing a riskless bond in the budget constraint, in order to obtain an Euler equation as an optimization condition, from which we could derive an IS schedule; on the other hand, we could introduce adaptive learning also on the consumer's side, in a way consistent with what we assume for the firms. Unfortunately, as shown in Preston (2003), the union of an Euler-type condition and non-rational adaptive learning lead to a suboptimal behavior, since it does not take into proper account the intertemporal budget constraint. Hence, we should look for a formulation that is compatible with optimization behavior, but that can still be used as a demand schedule.

Another interesting issue would be to remove the exogeneity assumption for money, and to suppose instead that the monetary authority pursues an optimal policy; Orphanides and Williams (2002) went in this direction, but considered only the optimal rule in the restricted class of linear rules, while, as they pointed out, the "true" optimal rule is a nonlinear function of the states of the system (including the time-t estimates). Moreover, they used an exogenously determined constant gain, when it would be preferable to implement an endogenous one.

8 Appendix

In this section we will prove the statements made in the text about the impulse response function of the Calvo price setting model.

First of all, we will solve the model formed by the New Keynesian Phillips Curve (equation (16)) and the nominal demand process that we have used throughout this paper (equation (8)), finding the MA(∞) representation of the inflation process. Then, we will prove that the coefficients of this MA(∞) representation (and, therefore, the impulse response function) are a monotonic decreasing sequence.

To begin with, recall the New Keynesian Phillips Curve:

$$\pi_t = ky_t + E_t\pi_{t+1}$$

where we have set $\beta = 1$ for notational simplicity; since β is usually calibrated at 0.99, it does not seem a restrictive assumption; k is a function of structural parameters which assumes positive values. For any arbitrary sequence$\{q_t\}_{t=0}^{\infty}$ for the nominal output,the only stationary solution is[45]:

$$p_t = \lambda p_{t-1} + (1-\lambda)^2 \sum_{j=0}^{\infty} \lambda^j q_{t+j} \tag{19}$$

We now assume that q_t follows the process given by (8), which can be represented in an MA(∞) form as[46]:

$$\Delta q_t = \sum_{j=0}^{\infty} \rho^j u_{t-j}$$

so that:

$$q_t = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \rho^j u_{t-j-k}$$

To derive the process of inflation, we follow Mankiw and Reis (2002) and guess that it is stationary, so that can be represented in the MA(∞) form:

$$\pi_t = \sum_{j=0}^{\infty} \varphi_j u_{t-j}$$

[45] See the appendix of Mankiw and Reis (2002).

[46] We omit the constant term in the process for nominal output, since we calibrate it to zero throughout the paper.

and the price level is the non-stationary process:

$$p_t = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \varphi_j u_{t-j-k}$$

where the $\{\varphi_j\}$ are unknown. Plugging this guess into equation (19), and taking into account the particular process considered for nominal output, we get:

$$\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \varphi_j u_{t-j-k} = \lambda \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \varphi_j u_{t-1-j-k} + (1-\lambda)^2 \sum_{j=0}^{\infty} \lambda^j \sum_{i=0}^{\infty} \sum_{k=\max\{j-i,0\}}^{\infty} \rho^i u_{t+j-i-k}$$

Matching coefficients, we get:

$$\varphi_0 = (1-\lambda)^2 \sum_{j=0}^{\infty} \lambda^j \sum_{i=0}^{j} \rho^i = \frac{1-\lambda}{1-\rho\lambda} \tag{20}$$

for the coefficient on u_t, and for an arbitrary $m > 1$:

$$\varphi_m = (\lambda - 1) \sum_{j=0}^{m-1} \varphi_j + \left[(1-\lambda)^2/(1-\rho)\right] \left[1/(1-\lambda) - \rho^{m+1}/(1-\rho\lambda)\right] \tag{21}$$

Now that the sequence of MA coefficients $\{\varphi_j\}$ as been characterized, it is possible to show that it is monotonic decreasing. First of all, we will show that $\varphi_1 - \varphi_0 < 0$; in fact, using (21):

$$\begin{aligned} \varphi_1 - \varphi_0 &= (\lambda-2)\varphi_0 + \frac{(1-\lambda)^2}{(1-\rho)(1-\lambda)} - \frac{(1-\lambda)^2\rho^2}{(1-\rho)(1-\rho\lambda)} \\ &= (\lambda-2)\frac{1-\lambda}{1-\rho\lambda} + \frac{(1-\lambda)^2}{(1-\rho)(1-\lambda)} - \frac{(1-\lambda)^2\rho^2}{(1-\rho)(1-\rho\lambda)} \end{aligned}$$

Simple algebra shows that:

$$\begin{aligned} \varphi_1 - \varphi_0 &= \frac{(1-\lambda)\left[-(1-\rho)^2 + \lambda - 2\lambda\rho + \lambda\rho^2\right]}{(1-\rho)(1-\rho\lambda)} \\ &= \frac{(1-\lambda)\left[-(1-\rho)^2 + \lambda(1-\rho)^2\right]}{(1-\rho)(1-\rho\lambda)} \end{aligned}$$

Since $\lambda < 1$, $\lambda(1-\rho)^2 - (1-\rho)^2$ is less than zero, and so is $\varphi_1 - \varphi_0$. Now, observe that for any $m \geq 2$, from (21) we obtain:

$$\varphi_m - \varphi_{m-1} = (\lambda-1)\left(\sum_{j=0}^{m-1} \varphi_j - \sum_{j=0}^{m-2} \varphi_j\right) + \frac{(1-\lambda)^2\left(\rho^m - \rho^{m+1}\right)}{(1-\rho)(1-\rho\lambda)}$$

or, equivalently:

$$\varphi_m = \lambda\varphi_{m-1} + \frac{(1-\lambda)^2\rho^m}{(1-\rho\lambda)} \tag{22}$$

Now we will show directly that also $\varphi_2 - \varphi_1$ is less than zero; in fact, from (22) we get:

$$\varphi_2 - \varphi_1 = (\lambda - 1)\,\varphi_{1+}\frac{(1-\lambda)^2\rho^2}{(1-\rho\lambda)}$$

Using equation (21) for $m = 2$ to substitute out φ_1, and equation (20) to substitute out for φ_0, we can write:

$$\varphi_2 - \varphi_1 = \left(\rho^2 + 1 - \lambda\right)\frac{1-\lambda}{1-\rho\lambda} + \frac{(1-\lambda)^2}{(1-\rho\lambda)\,(1-\rho)}\rho^2 - \frac{1-\lambda}{1-\rho}$$

Simple algebra shows that this expression is negative if and only if $\rho^2\,(2-\lambda-\rho) + 2\rho\lambda - \lambda - \rho \equiv H(\lambda,\rho)$ is negative; but this function is always negative valued, provided that $\rho < 1$, which is a stationarity condition always assumed. In fact, we have that:

$$\frac{\partial}{\partial\lambda}H(\lambda,\rho) = -\rho^2 - 1 + 2\rho$$

which is negative whenever $\rho \neq 1$[47]; so, it is sufficient to check that $H(\cdot)$ is negative valued when λ has the minimum admittable value (i.e., zero). Note that:

$$H(0,\rho) = \rho^2\,(2-\rho) - \rho \gtrless 0 \Leftrightarrow \rho\,(2-\rho) - 1 \gtrless 0$$

But the last expression is $-\rho^2 - 1 + 2\rho$, which we have already see that is negative; thus, we conclude that $\varphi_2 - \varphi_1$ is negative whenever $\rho < 1$.

To prove that also the rest of the sequence is decreasing, we proceed by induction: we assume that $\varphi_{m-1} - \varphi_{m-2} < 0$ for an arbitrary $m \geq 3$; we want to show that $\varphi_m - \varphi_{m-1} < 0$ as well. Using equation (21) we get:

$$\varphi_m - \varphi_{m-1} = \lambda(\varphi_{m-1} - \varphi_{m-2}) + \frac{(1-\lambda)^2}{(1-\rho\lambda)}\left(\rho^m - \rho^{m-1}\right)$$

Since $\varphi_{m-1} - \varphi_{m-2} < 0$ for the induction hypothesis, and $\rho^m - \rho^{m-1} < 0$ because $0 \leq \rho < 1$, we conclude that $\varphi_m - \varphi_{m-1} < 0$.

[47] In fact, the discriminant of this quadratic expression is zero, so that it has only one root at $\rho = 1$.

Figure 1: Simulated parameters estimates for gamma=0.55

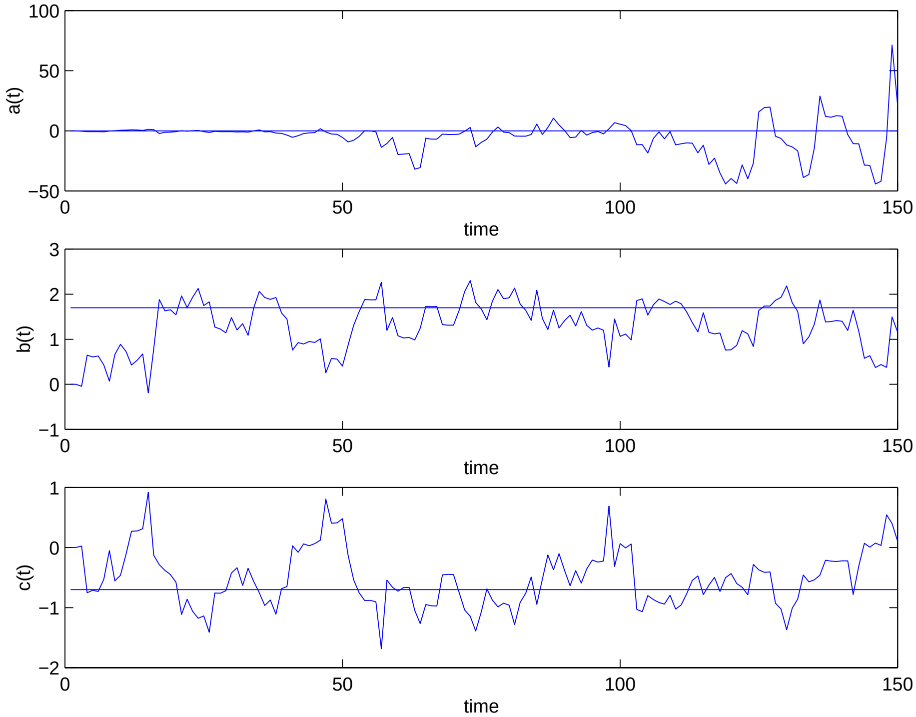


Figure 2: Impulse response function for inflation

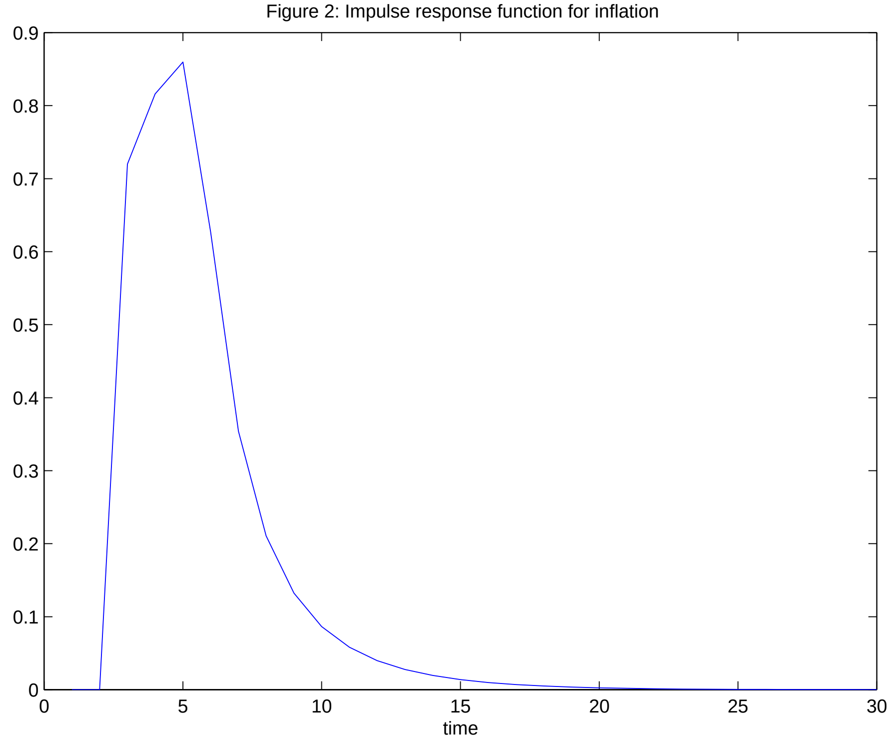
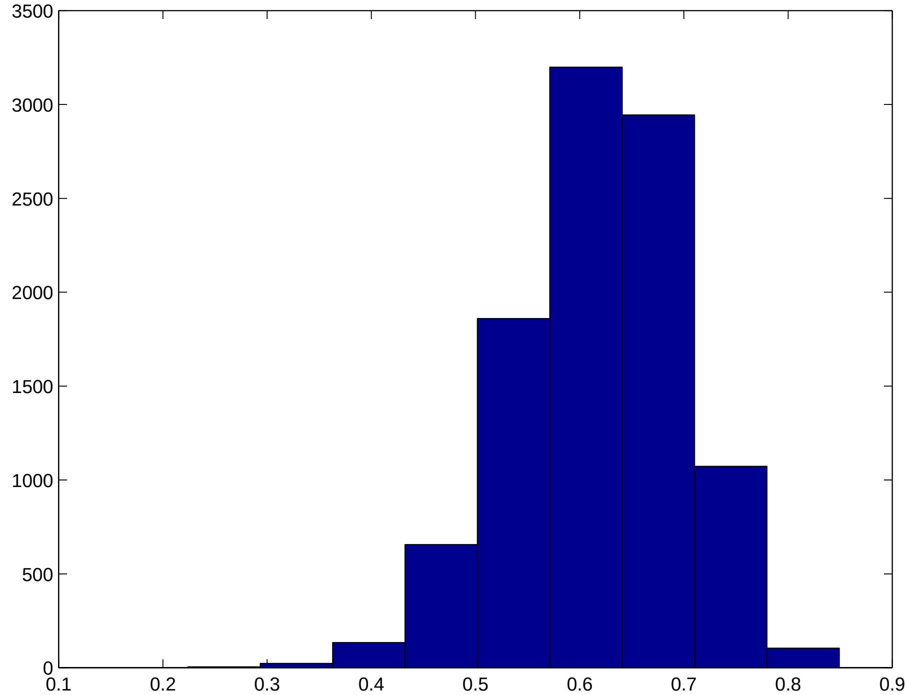
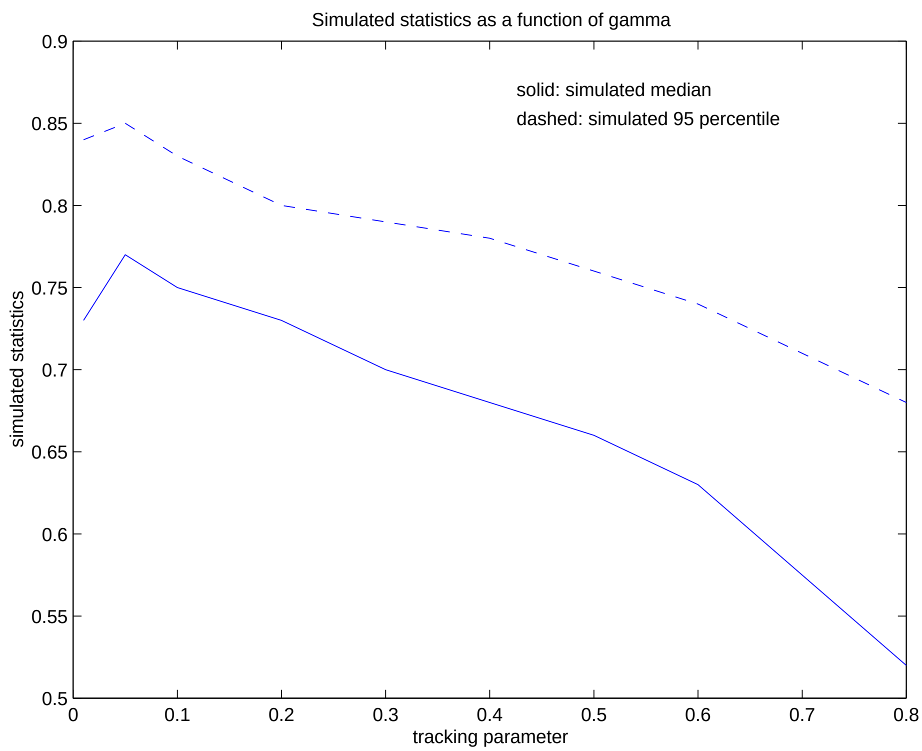


Figure 3: Simulated distribution of first order autocorrelation of inflation




Simulated statistics as a function of gamma
solid: simulated median
dashed: simulated 95 percentile
simulated statistics
tracking parameter
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

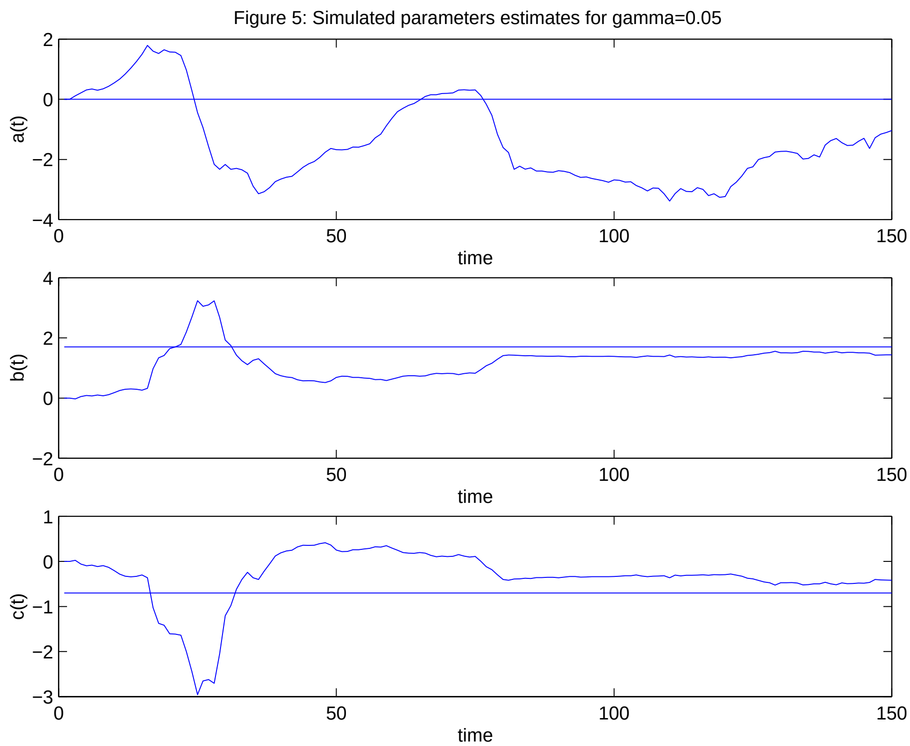
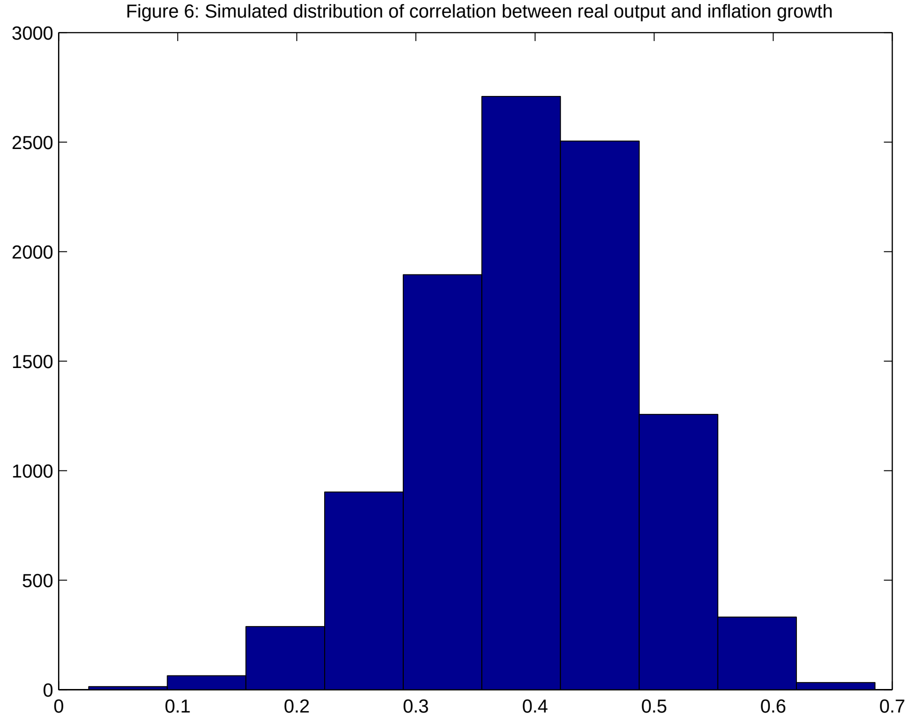


Figure 5: Simulated parameters estimates for gamma=0.05
a(t)
b(t)
c(t)
time
time
time
0
50
100
150
2
0
−2
−4
4
2
0
−2
1
0
−1
−2
−3

Figure 6: Simulated distribution of correlation between real output and inflation growth



References

[1] Abel, A. B. and B. S. Bernanke, 1998, *Macroeconomics*, Reading: Addison-Wesley.

[2] Adam, Klaus, 2003, Learning to Forecast and Cyclical Behavior of Output and Inflation, University of Frankfurt Working Paper.

[3] Ball, L., N. G. Mankiw and R. Reis, 2003, Monetary Policy for Inattentive Economies, NBER Working Paper 9491.

[4] Balvers, R. J. and T. F. Cosimano, 1990, Actively Learning About Demand and the Dynamics of Price Adjustment, *The Economic Journal* **100**: 882-898.

[5] Barro, R. J. and D. B. Gordon, 1983, A Positive Theory of Monetary Policy in a Natural Rate Model, *Journal of Political Economy* **91**: 589-610.

[6] Blanchard, O., 2000, *Macroeconomics*, Upper Saddle River: Prentice Hall.

[7] Bray, M. and N. Savin, 1986, Rational Expectations Equilibria, Learning, and Model Specification, *Econometrica* **54**: 1129-1160.

[8] Calvo, G., 1983, Staggered Prices in a Utility Maximizing Framework, *Journal of Monetary Economics* **12**: 383-398.

[9] Canova, F. and E. Ortega, 2000, Testing Calibrated General Equilibrium Models, published in R. Mariano, T. Shuermann and M. Weeks eds., *Inference using Simulation Techniques*, Oxford University Press.

[10] Christiano, L. J., M. Eichenbaum, and C. L. Evans, 1998, Modeling Money, NBER Working Paper 6371.

[11] Christiano, L. J., M. Eichenbaum, and C. L. Evans, 2001, Nominal Rigidities and the Dynamic Effects of e Shock to Monetary Policy, NBER Working Paper 8403.

[12] Clarida, R., J. Gali' and M. Gertler, 1999, The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature* **37**: 1661-1707.

[13] Dixit, A. K. and J. E. Stiglitz, 1977, Monopolistic Competition and Optimum Product Diversity, *American Economic Review* **67**: 297-308.

[14] Evans, G. W. and S. Honkapohja, 1993, Adaptive Forecasts, Hysteresis, and Endogenous Fluctuations, Federal Reserve Bank of San Francisco Economic Review.

[15] Evans, G. W. and S. Honkapohja, 2001, *Learning and Expectations in Macroeconomics*, Princeton: Princeton University Press.

[16] Evans, G. W. and S. Honkapohja, 2002, Adaptive Learning and Monetary Policy Design, University of Oregon Working Paper.

[17] Evans, G. W. and G. Ramey, 2001, Adaptive Expectations, Underparameterization and the Lucas Critique, University of Oregon Working Paper.

[18] Lucas, R. E., Jr, 1972, Expectations and the Neutrality of Money, *Journal of Economic Theory* **4**: 103-124.

[19] Heinemann, M., 2000, Convergence of Adaptive Learning and Expectational Stability: The Case of Multiple Rational Expectations Equilibria. *Macroeconomic Dynamics 4: 263-288.*

[20] Honkapohja, S. and K. Mitra, 2002, Learning Stability in Economics with Heterogeneous Agents, Unuversity of Helsinki Working Paper.

[21] Mankiw, G. N. and R. Reis, 2002, Sticky Information Versus Sticky Prices: a Proposal to Replace the New Keynesian Phillips Curve, *Quarterly Journal of Economics* **114**.

[22] Marcet, A. and J. P. Nicolini, 1998, Recurrent Hyperinflations and Learning, CEPR Working Paper 1875.

[23] Marcet, A. and T. J. Sargent, 1989, Convergence of Least-Squares Learning Mechanisms in Self-Referential Linear Stochastic Models, *Journal of Economic Theory* **48**: 337-368.

[24] McGough, B., 2000, Statistical Learning with Time-Varying Parameters, Oregon State University Working Paper.

[25] Orphanides, A. and J. C. Williams, 2002, Imperfect Knowledge, Inflation Expectations, and Monetary Policy, Federal Reserve Bank of San Francisco Working Paper.

[26] Phelps, E. S., 1970, Introduction: The New Microeconomics in Employment and Inflation Theory, published in E. S. Phelps et al. eds., 1970, *Microeconomic Foundations of Employment and Inflation Theory*, New York: Norton.

[27] Preston, B., 2003, Learning about Monetary Policy Rules when Long-Horizon Expectations Matter, Princeton University Working Paper.

[28] Rotemberg, J. and M. Woodford, 1997, Optimization-Based Econometric Framework for the Evaluation of Monetary Policy, published in B. S. Bernanke and J. Rotembergeds., *NBER Macroeconomic Annual 1997*, Cambridge: MIT Press.

[29] Sargent, T. J., 1993, *Bounded Rationality in Macroeconomics*, Oxford: Oxford University Press.

[30] Sargent, T. J., 1999, *The Conquest of American Inflation*, Princeton: Princeton University Press.

[31] Sims, C. A., 2003, Implications of Rational Inattention, *Journal of Monetary Economics* **50**: 665-690.

[32] Tetlow, R. and P. von zur Muehlen, 2001, Avoiding Nash Inflation: Bayesian and Robust Responses to Model Uncertainty, Federal Reserve Bank Economics and Finance Discussion Paper.

[33] Timmermann, A. G., 1993, How Learning in Financial Markets Generats Excess Volatility and Predictability in Stock Prices, *Quarterly Journal of Economics* **108**: 1135-1145.

[34] Timmermann, A. G., 1996, Excess Volatility and Predictability of Stock Prices in Autoregressive Dividend Models with Learning, *Review of Economic Studies* **63**: 523-557.

[35] Townsend, R. M., 1983, Forecasting the Forecast of Others, *Journal of Political Economy* **91**: 546-588.

[36] Williams, N., 2003, Adaptive Learning and Business Cycles, Princeton University Working Paper.

[37] Woodford, M., 2001, Imperfect Common Knowledge and the Effects of Monetary Policy, published in P. Aghion, R. Frydman, J. Stiglitz, M. Woodford eds., 2002, *Knowledge, Information and Expectations in Modern Macroeconomics: in Honor of Edmund S. Phelps*, Princeton: Princeton University Press.

[38] Woodford, M., 2002, *Optimizing Models with Nominal Rigidities*, unpublished, Princeton University.