Macroeconomic Dynamics, **13**, 2009, 167–188. Printed in the United States of America.
doi:10.1017/S136510050807051X

LEARNING THE INFLATION TARGET

RICARDO NUNES
Federal Reserve Board

We propose a framework in which expectations have a rational and a learning component. We describe a solution method for these frameworks and provide an application to the Volcker disinflation with the New Keynesian model. Although the model with rational expectations does not seem to account for this episode, results improve when a small and empirically plausible proportion of private agents are learning. The learning component is argued to be more robust and plausible than the rule-of-thumb expectations present in the hybrid Phillips curve.

Keywords: Adaptive Learning, Heterogeneity, Volcker Disinflation

1. INTRODUCTION

We set up a theoretical model in which some agents are rational but the others need to learn the behavior of the economy. Learning is modeled as in Marcet and Sargent (1989) and Evans and Honkapohja (2001). Economic agents are assumed to behave as econometricians because they have limited information about the underlying economic model. Like that of Erceg and Levin (2003), this paper examines the Volcker disinflation, which can be seen as a change in monetary policy regime. The assumption that the private sector learns is specially suited to the examination of a regime change, because rational expectations (henceforth, RE) unrealistically assume that the expectations of the private sector catch up immediately.

The approach in the present paper is related to the rule-of-thumb literature. For instance, Galí and Gertler (1999) consider the New Keynesian (henceforth, NK) model with rational and rule-of-thumb firms, which expect future inflation to be equal to lagged inflation. In the present model, the learning algorithm uses the same structural form of RE, which is theoretically more defensible and provides a more robust justification of backward-looking behavior than Galí and Gertler (1999). In

I am grateful to Albert Marcet, Kosuke Aoki, Gregg Forte, Jordi Galí, Massimo Guidolin, Christian Haefke, Ramon Marimon, Martin Ellison, two anonymous referees, and seminar participants at Pompeu Fabra, GREQAM, and the Xth Spring Meeting for Young Economists at HEI for helpful comments. Any remaining errors are my own. Financial support from the Fundação para a Ciência e Tecnologia and the prize money received for the award at the 2005 Xth Spring Meeting for Young Economists are gratefully acknowledged. A previous version of this paper received the best paper award at the 2005 Xth Spring Meeting of Young Economists. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. Address correspondence to: Ricardo Nunes, Board of Governors of the Federal Reserve System, 20th Street and Constitution Avenue NW, Washington, DC 20551, USA; e-mail: ricardo.p.nunes@frb.gov.

 1365-1005/09

addition, unlike that of Galí and Gertler (1999), our learning algorithm satisfies the internal consistency criterion proposed by Marcet and Nicolini (2003). We also present empirical estimations of the learning speed that support the parameters found in the internal consistency criterion.

To evaluate the proposed framework, we maintain the simplest NK model as our benchmark and we describe the effects of a disinflation. The NK model with rational expectations unrealistically predicts that unanticipated credible disinflations can be achieved instantaneously and with no output costs. Those results are clearly at odds with empirical findings. Analyzing a panel of countries, Ball (1994) shows that disinflations very frequently cause recessions and that inflation adjusts slowly to target. This paper shows that if all agents are learning, inflation moves slowly to target while the output gap is temporarily affected, but the economy takes too long to converge to its new steady state. Because the speed of convergence depends on the speed at which private agents learn, the proposed model makes the learning speed endogenous.

The literature is debating whether the Phillips curve is mainly forward-looking. Galí and Gertler (1999) found the NK Phillips curve to be mainly forward-looking. Linde (2005) has challenged this view, and Galí et al. (2005) have supported it. The present paper sheds some light on this issue. We find that the Volcker disinflation can be roughly matched if the proportion of learners is about 30%. Having even a small proportion of learners generates several differences from the case of pure RE. As long as some agents are learning, the forecasts of rational agents are significantly changed; hence the assumption of pure RE may be rather misleading. The paper also finds that although our model can explain some features of the disinflation experiment, the same model with rule-of-thumb firms is less satisfactory.

Erceg and Levin (2003) also build a learning model that successfully explains disinflation dynamics. We consider that expectations are a mixture of rational expectations and least-squares learning. In contrast, Erceg and Levin (2003) model the private sector as making use of the Kalman filter to infer the value of the unobservable inflation target.[1] Private agents in their model thus have specific alternative regimes in mind, along with the associated transition probabilities. In our model, learning agents are not aware of the specific structural changes that may occur and are ready to learn any structural change. Another difference is that under least-squares learning, economic agents do not know how specific changes in parameters modify the economic structure. Agents using the Kalman filter know the exact structure of the economy. Both our model and that of Erceg and Levin (2003) can be interpreted as a lack of credibility of the central bank. Even if the central bank announces the disinflation, agents need to learn and form expectations with the data.[2]

Other papers have used least-squares learning to explain transition dynamics. Marcet and Nicolini (2003) explain recurrent hyperinflations, and Giannitsarou (2006) accounts for a transition after a fiscal reform. We focus on the learning of private agents, and thus depart from Sargent (1999) and Primiceri (2006). For instance, Primiceri's work offers central bank learning as the explanation for the rise of inflation in the 1970s and its decline in the 1980s. The Primiceri model can account for costly disinflations because it assumes a completely backward-looking Phillips curve. In contrast, we focus on the learning side of firms and on how such a setup can induce empirically plausible disinflation patterns. Other works dealing with monetary policy and learning include those of Evans and Honkapohja (2003), Ferrero (2007), and Orphanides and Williams (2005, 2006). Milani (2005) and Santoro (2003) also explain inflation persistence by modeling the private sector as learning. Other papers introducing the issue of heterogeneous expectations in a learning framework include those of Evans et al. (1993) and Giannitsarou (2003).

The paper is structured as follows: Section 2 describes the model, Section 3 evaluates the homogeneous expectations model, Section 4 presents and evaluates the heterogeneous expectations model, and Section 5 concludes.

2. DESCRIPTION OF THE MODEL

We use the standard NK model in reduced form as derived by Galí (2007) and Woodford (2003).[3] The derivation of the model is available in the previous references; this section describes only the reduced-form equations.[4] The NK Phillips curve (henceforth NKPC) is given by

$$\pi_t = \kappa z_t + \beta \tilde{E}_t \pi_{t+1}, \tag{1}$$

where π_t is inflation, z_t is the output gap, and $\tilde{E}_t$ denotes expectations, which may not be rational. It is common to consider a stochastic term in the NKPC denoted by cost-push shock. The investment–savings curve is described as

$$z_t = \tilde{E}_t z_{t+1} - \sigma^{-1}(r_t - r_t^{\mathrm{n}} - \tilde{E}_t \pi_{t+1}), \tag{2}$$

where r_t is the interest rate set by the central bank, and r_t^{n} is the natural interest rate. Usually the natural interest rate is assumed to follow an AR(1) process,

$$r_t^{\mathrm{n}} = \rho r_{t-1}^{\mathrm{n}} + \varepsilon_t. \tag{3}$$

The model is closed with an equation for the interest rate. We will assume a Taylor (1993) rule of the type

$$r_t = \pi^* + \varphi_\pi(\pi_t - \pi^*) + \varphi_z z_t, \tag{4}$$

where the constant term in the interest rate rule is set to make the inflation target π^* consistent in equilibrium. Using the interest rate rule in equation (2) and rearranging the system, one obtains

$$y_t = a + b\tilde{E}_t y_{t+1} + \varkappa r_t^{\mathrm{n}}, \tag{5}$$

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

where $y_t = [z_t, \pi_t]'$, $a = [1/(\sigma + \varphi_z + \kappa\varphi_\pi)][\pi^*(\varphi_\pi - 1), \kappa\pi^*(\varphi_\pi - 1)]'$, $\varkappa = [1/(\sigma + \varphi_z + \kappa\varphi_\pi)][1, \kappa]'$, and

$$b = \frac{1}{\sigma + \varphi_z + \kappa\varphi_\pi}\begin{bmatrix} \sigma & 1 - \beta\varphi_\pi \\ \kappa\sigma & \kappa + \beta(\sigma + \varphi_z) \end{bmatrix}. \quad (6)$$

Taking the results of Blanchard and Kahn (1980), Bullard and Mitra (2002) find that if $\kappa(\varphi_\pi - 1) + (1 - \beta)\varphi_z > 0$, then the model described in equation (5) has a unique RE solution. Note that Bullard and Mitra (2002) consider a model where $a = 0$. If a is a constant matrix, condition (1c) in Blanchard and Kahn (1980) must be satisfied. Because this condition is trivially met, uniqueness conditions are equivalent.

The minimal state variable (MSV) solution takes the form $y_t = A + Cr_t^{\mathrm{n}}$.[5] In the learning literature, economic agents are assumed to be ignorant of the RE solution coefficients but do know the functional form of the MSV solution. At each moment, the private sector will use the available data to estimate A and C with a learning algorithm; we denote the estimates of A and C obtained at time t by A_t and C_t. Thus, at time t the private sector will think that the economy behaves as $y_t = A_t + C_t r_t^{\mathrm{n}}$, an equation known as the perceived law of motion (henceforth PLM). Given the estimates of the private sector, expectations are formed as $\tilde{E}_t y_{t+1} = A_t + C_t \rho r_t^{\mathrm{n}}$. Inserting expectations into equation (5) and rearranging the terms to fit the functional form of the MSV solution yields $y_t = a + bA_t + (bC_t\rho + \varkappa)r_t^{\mathrm{n}}$, which describes the actual law of motion (henceforth, ALM) of the economy. The mapping from the PLM to the ALM is called the T-map, and in this model it is given by

$$T(A, C) = (a + bA, bC\rho + \chi), \quad (7)$$

where we drop the time subscripts for convenience. The fixed point in the T-map is the RE solution, in which the PLM and ALM are equal. It is common to determine whether the learning equilibrium converges to the RE solution by checking the E-stability condition. Bullard and Mitra (2002) find that in this model the condition for E-stability is equal to the condition for uniqueness.

Regarding the way that learners' parameters are updated, the model assumes that private agents use a recursive least squares (RLS) formula given by

$$\phi_t = \phi_{t-1} + R_t^{-1} x_{t-1}(y_{t-1} - \phi_{t-1}' x_{t-1})' * \alpha_{\mathrm{M},t} \quad (8)$$

$$R_t = R_{t-1} + \alpha_{R,t}(x_{t-1}x_{t-1}' - R_{t-1}) \quad (9)$$

where

$$\phi_t = \begin{bmatrix} A_{z,t} & A_{\pi,t} \\ C_{z,t} & C_{\pi,t} \end{bmatrix} = \begin{bmatrix} A_t' \\ C_t' \end{bmatrix}, \quad x_{t-1} = \begin{bmatrix} 1 \\ r_{t-1}^{\mathrm{n}} \end{bmatrix}, \quad \alpha_{\mathrm{M},t} = \begin{bmatrix} \alpha_{\mathrm{out},t} & 0 \\ 0 & \alpha_{\mathrm{inf},t} \end{bmatrix}.$$

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

If $\alpha_{R,t} = \alpha_{\text{out},t} = \alpha_{\text{inf},t} = t^{-1}$, then all the observations in the data receive the same weight. A common assumption in the literature is that of a constant gain algorithm where $\alpha_{\text{R},t} = \alpha_{\text{out},t} = \alpha_{\text{inf},t} = p$, and $p \in (0, 1)$. A constant tracking parameter means that recent observations are given more weight, and the past is discounted. The constant gain can be interpreted as agents taking the possibility of regime changes into account. If a regime change occurs, agents adapt more quickly because recent data are given more weight.

The present model considers a RLS algorithm more general than those considered in Evans and Honkapohja (2001). It allows different tracking parameters for updating the parameters related to the output and inflation equations. In a regime shift, it will be optimal for learners to use higher tracking parameters; nevertheless, the regime shift may not have equal consequences for the two regressions that learners perform. We will determine the tracking parameters in an optimal way, without restricting learners to use the same tracking parameter in different regressions.

2.1. The Heterogeneous Framework

Previous work has argued that including lags of inflation in the Phillips curve improves the empirical fit of the model.[6] Galí and Gertler (1999) consider that some firms are rational, whereas others think that future inflation will be equal to last-period inflation. Roberts (1998) considers expectations to be a weighted average of rational expectations and lags of past data. In practical terms it is assumed that

$$\tilde{E}_t y_{t+1} = \psi y_{t-1} + (1 - \psi) E_t^{\text{R}} y_{t+1}, \qquad \textbf{(10)}$$

where ψ is the weight given to last-period data. The rule of thumb of using past data as a proxy for future expectations can be appealing either because it is a less costly approach to forming expectations or on the ground that expectations are not entirely rational.

The present model replaces the rule-of-thumb terms in the last equation with a learning component. This option seems more natural for several reasons. First, the learning algorithm uses the same structural form of RE. Without a regime break and a sufficiently large number of data, learners will form expectations rationally. This feature contrasts with the rule-of-thumb literature, in which agents would never form expectations rationally. Second, as we will see in a later section, we can determine the learning speed endogenously. In contrast, with rule of thumb, the learning speed is not endogenous and in general is not optimal. Hence, the model assumes that aggregate expectations are a weighted average of the expectations of learners and rationals:

$$\tilde{E}_t y_{t+1} = \psi E_t^{\text{L}} y_{t+1} + (1 - \psi) E_t^{\text{R}} y_{t+1}. \qquad \textbf{(11)}$$

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

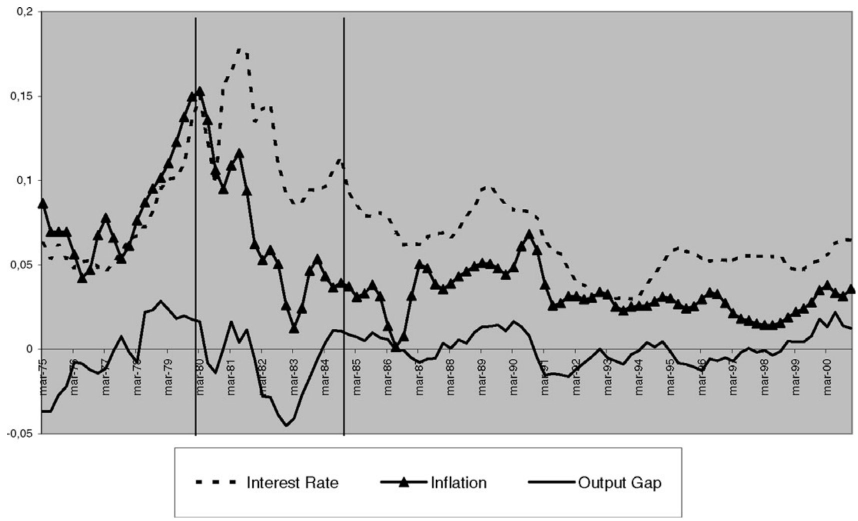

FIGURE 1. Interest rate, inflation, and output gap.

Inserting equation (11) into equation (5) yields

$$y_t = a + b\psi E_t^{\mathrm{L}} y_{t+1} + b(1-\psi) E_t^{\mathrm{R}} y_{t+1} + \varkappa r_t^{\mathrm{n}}. \qquad \textbf{(12)}$$

A fixed proportion of agents, $1-\psi$, are rational and therefore know how learners form expectations and know the proportion of learners and rationals in the economy. If $\psi = 0$, then rational agents predict that inflation can jump immediately to its new target and that no recession occurs. On the other hand, if the proportion of learners is nearly 1, then rationals should predict as learners do, and disinflations will be more costly.

3. MODEL EVALUATION

3.1. The Disinflation Episode

Figure 1 plots inflation, the output gap, and the interest rate in the United States. In the 1980s, inflation was brought down considerably, and this period is often referred to as the Volcker disinflation. A methodology for identifying the beginning and the end of a disinflation has not been established. Ball (1994), using a detrending method that differs from that in the present paper, reports that trough inflation occurred 16 quarters after peak inflation, and considers that the effects in the output gap lasted for 20 quarters.[7] During this period the economy went into recession, as can be seen by observing the output gap in figure 1. The sacrifice ratio, a common measure for disinflations, is computed as the cumulative output gap divided by the change in inflation. The computation of the sacrifice ratio is quite sensitive to the methodology used to calculate the components of the ratio.

Ball (1994) assumes that natural output would grow log-linearly from the start of the disinflation to the end. That methodology serves the author's purposes by identifying the relative costs of different episodes in a systematic way. The sacrifice ratio computed with Ball (1994) methodology and assuming that the disinflation lasts for 16 quarters and 20 quarters is 0.89% and 1.38%, respectively.[8] For the data series presented in this paper, and considering only the recession periods, the sacrifice ratio is 0.56%. In accordance with the literature, we obtain natural output by using the bandpass filter of Baxter and King (1999). The resulting sacrifice ratio is lower than has been previously assumed, but the detrending method used here is closer to what one would expect to obtain with a theoretical model.

The calibration uses the values $\kappa = 0.1$, $\sigma = 0.64$, $\rho = 0.35$, and $\beta = 1.0$, which are consistent with Nunes (2005) and Woodford (2003).[9] For the interest rate rule, we assume $\varphi_\pi = 1.5$ and $\varphi_z = 0.5$, as in the original parameterization suggested by Taylor. The standard deviations for the innovation in the natural interest rate and the cost-push shock are 0.015 and 0.01, respectively. For those parameters, the solution is unique and E-stable. At the beginning of the disinflation, agents' expectations and inflation are assumed to be 15.3%; the inflation target is then lowered to 3.7%. Those values correspond to filtered inflation 1980:Q1 and 1984:Q4, as can be observed in Figure 1.

Some empirical estimates are available for the parameter ψ. According to Galí and Gertler (1999), the weight on backward-looking behavior is estimated to be between 0.2 and 0.4. The authors conclude that the fraction of firms with backward-looking behavior is smaller than that with forward-looking behavior, but statistically different from 0. Nunes (2005) examines an equation similar to that in Galí and Gertler (1999), but where inflation expectations are an average of RE and expectations from the Survey of Professional Forecasters.[10] Nunes (2005) reports the following result[11]

$$\pi_t = \underset{(0.0065)}{0.0260} mc_t + \underset{(0.0706)}{0.3432} S_t\pi_{t+1} + (1 - 0.34329) E_t\pi_{t+1}, \tag{13}$$

where $S_t\pi_{t+1}$ is the average inflation expectation from the Survey of Professional Forecasters, and mc_t is marginal cost.[12] Branch and Evans (2006) and estimates presented later in this paper provide evidence that survey results can be accurately described by a learning algorithm. Therefore, equation (13) is the closest empirical counterpart to the model presented here.

In the next section, the model is simulated with different weights of RE.

3.2. The Limiting Cases

The RE model can be obtained by setting $\psi = 0$. Such a model cannot account for disinflation behavior. As Mankiw and Reis (2002) emphasize, the NK model under RE and an unanticipated credible disinflation generates an immediate reduction of inflation and imposes no output costs.

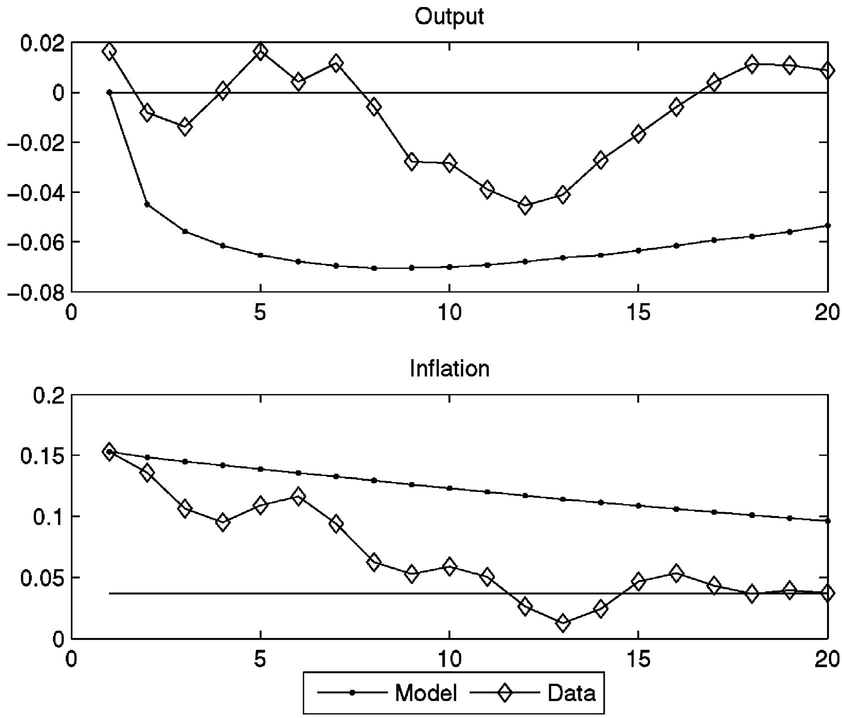

FIGURE 2. Learners' economy—output and inflation.

The other extreme case is $\psi = 1$, in which the economy is solely inhabited by learners. A simulation of that model requires values for $\alpha_{\text{inf}}, \alpha_{\text{out}}$.[13] The internal consistency criterion for this model implies that ($\alpha_{\text{inf}} = 0.5, \alpha_{\text{out}} = 0.5$). We will examine this criterion in more detail in the model with heterogeneous expectations.[14] The parameters found in our analysis are higher than the values that the literature usually assumes; the main reason is that the large bulk of the literature does not focus on regime changes.

Figure 2 plots the average inflation and output gap generated by 5,000 simulations of the model. Qualitatively the model performs well, inflation decreases sluggishly, and the economy experiences a temporary recession. In the symmetric case, i.e., raising the inflation target, inflation rises sluggishly to the target while the economy experiences a boom. The drawback of this model is that convergence is too slow.

To obtain a more precise notion of convergence, the mean and the standard deviation around the mean are reported in Table 1.[15] The last row of the table reports the first period for which mean inflation and the mean output gap differ from their steady state levels by less than 0.005.

At both $t = 16$ and $t = 20$, average inflation and average output are still far from their steady states. The economy takes 85 quarters to reach its new steady state, a

TABLE 1. Output–inflation convergence

	Output		Inflation	
	$t = 16$	$t = 20$	$t = 16$	$t = 20$
Mean	−0.0615	−0.0535	0.1060	0.0963
St. Dev.	0.0227	0.0232	0.0161	0.0167
End	84		85	

slow convergence in comparison with the Volcker disinflation. The limiting case of $\psi = 1$ can explain disinflations qualitatively but fails on a quantitative dimension.

4. SOLVING THE MODEL WITH HETEROGENEOUS EXPECTATIONS

The previous section showed that the limiting cases of pure rational expectations or pure learning do not seem to provide satisfactory transition dynamics for the Volcker disinflation. The model with heterogeneous expectations offers a nontrivial solution, which we will explain in detail.

4.1. The I.I.D. Case

The analysis starts with a simple case in which the natural interest rate in the heterogeneous expectation model is i.i.d. For expositional purposes, learners are assumed to use a unique tracking parameter. In this case the learning algorithm is simplified to

$$E_t^{\mathrm{L}} y_{t+1} = E_{t-1}^{\mathrm{L}} y_t + \alpha\left(y_{t-1} - E_{t-1}^{\mathrm{L}} y_t\right). \tag{14}$$

To solve the rational agents problem, we conjecture that the MSV solution takes the form

$$y_t = A + B E_t^{\mathrm{L}} y_{t+1}. \tag{15}$$

Hence

$$y_{t+1} = A + B E_{t+1}^{\mathrm{L}} y_{t+2} \tag{16}$$

$$y_{t+1} = A + B\left[E_t^{\mathrm{L}} y_{t+1}(1-\alpha) + \alpha y_t\right]. \tag{17}$$

Because $E_t^{\mathrm{R}} y_{t+1} = A + B[E_t^{\mathrm{L}} y_{t+1}(1-\alpha) + \alpha y_t]$, and after arranging terms, the ALM is

$$\begin{aligned} y_t &= [I - b(1-\psi)B\alpha]^{-1}\{[a + b(1-\psi)A]\} \\ &\quad + [I - b(1-\psi)B\alpha]^{-1}\left\{[b\psi + b(1-\psi)(1-\alpha)B]E_t^{\mathrm{L}} y_{t+1}\right\}. \end{aligned} \tag{18}$$

Solving for rational agents requires consideration of the following two equations:

$$A = [I - b(1-\psi)B\alpha]^{-1}[a + b(1-\psi)A] \tag{19}$$

$$B = [I - b(1-\psi)B\alpha]^{-1}[b\psi + b(1-\psi)(1-\alpha)B]. \tag{20}$$

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

The second equation is a quadratic matrix equation on B that can be solved using the generalized eigenvalues method.[16] With a solution for B, solving for A is trivial.

The MSV solution changed with the introduction of rational agents. Adaptive learners realizing that their expectations are taken into account by rationals, would revise their MSV, which in turn would cause rationals to reestimate their MSV, and so on ad infinitum. However, if learners behave as if rationals did not exist, then the expectations of learners do converge to equilibrium. In that case, convergence conditions under tracking can be computed by examining directly the learning algorithm, which is[17]

$$E_t^{\mathrm{L}} y_{t+1} = E_{t-1}^{\mathrm{L}} y_t + \alpha\left(y_{t-1} - E_{t-1}^{\mathrm{L}} y_t\right). \qquad \mathbf{(21)}$$

Using equations (18)–(20), equation (21) can be rewritten as

$$E_t^{\mathrm{L}} y_{t+1} = [I(1-\alpha) + \alpha B] E_{t-1}^{\mathrm{L}} y_t + \alpha A + \alpha [I - b(1-\psi) B\alpha]^{-1} \varkappa r_{t-1}^{\mathrm{n}}. \quad \mathbf{(22)}$$

The previous system will be stable as long as the matrix $[I(1-\alpha) + \alpha B]$ has eigenvalues with an absolute value smaller than 1. The asymptotic mean of expectations is $Ey = (I-B)^{-1}A$, which corresponds to the RE equilibrium $[0, \pi^*]'$. The distribution is asymptotic with variance given by vec $\sum = \{I - [I(1-\alpha)+\alpha B] \otimes [I(1-\alpha)+\alpha B]\}^{-1} \mathrm{vec}(\alpha[I-b(1-\psi)B\alpha]^{-1}\varkappa\sigma\{\alpha[I-b(1-\psi)B\alpha]^{-1}\varkappa\}')$. Using the normality assumption for r_t^{n}, one concludes with $E_t^{\mathrm{L}} y_{t+1} \sim N([0,\pi^*]', \sum)$.

4.2. The Autocorrelated Shocks Case—Approximate Linear Solution

In the last section, the fact that learners and rationals were predicting only a constant for output and inflation made the analysis easier. The technical difficulty with rational agents is that they must take into account the process by which learners form expectations. Once learners estimate more than an average, an explicit MSV for which a fixed point exists is not so easily obtained. When the natural interest rate is autocorrelated, as is usually assumed in the NK model, learners follow the algorithm equations (8) and (9), which is no longer linear.

A simplifying assumption will enable an explicit solution; that is, rationals do not realize that learners' parameters will change in the following period. We denote these agents as near-rationals. Near-rationals will still be aware of regime shifts; and if they were the sole inhabitants of the NK economy, disinflations would still be costless. Near-rationals will solve the equation

$$y_t = A^{\mathrm{R}} + B^{\mathrm{R}} A_t^{\mathrm{L}} + C^{\mathrm{R}} r_t^{\mathrm{n}} + D^{\mathrm{R}} C_t^{\mathrm{L}} r_t^{\mathrm{n}}, \qquad \mathbf{(23)}$$

where variables with the R superscript are variables for near-rationals and variables with the L superscript are variables for learners' MSV, which is $y_t = A_t^{\mathrm{L}} + C_t^{\mathrm{L}} r_t^{\mathrm{n}}$. Expectations for near rationals are formed as $E_t^{\mathrm{R}} y_{t+1} = A^{\mathrm{R}} + B^{\mathrm{R}} A_t^{\mathrm{L}} + C^{\mathrm{R}} \rho r_t^{\mathrm{n}} + D^{\mathrm{R}} C_t^{\mathrm{L}} \rho r_t^{\mathrm{n}}$; i.e. learners parameters A_t and C_t are not taken to evolve. So inserting

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

these expectations into the ALM yields

$$y_t = a + b(1-\psi)A^{\mathrm{R}} + b(I\psi + (1-\psi)B^{\mathrm{R}})A_t^{\mathrm{L}} + [b(1-\psi)C^{\mathrm{R}}\rho + \varkappa]r_t^{\mathrm{n}} + b[I\psi + (1-\psi)D^{\mathrm{R}}]\rho C_t^{\mathrm{L}} r_t^{\mathrm{n}}. \tag{24}$$

The solution must satisfy $A^{\mathrm{R}} = [I - b(1-\psi)]^{-1}a$, $B^{\mathrm{R}} = [I - b(1-\psi)]^{-1}b\psi$, $C^{\mathrm{R}} = [I - b(1-\psi)\rho]^{-1}\varkappa$, $D^{\mathrm{R}} = [I - b(1-\psi)\rho]^{-1}b\psi\rho$. The way that this problem was solved ensures that once learners converge to RE equilibrium, so do near-rationals.[18] Hence, the relevant question to be posed is whether learners, expectations will converge to equilibrium. The relevant T-map is

$$T(A^{\mathrm{L}}, C^{\mathrm{L}}) = \{a + b(1-\psi)A^{\mathrm{R}} + b[I\psi + (1-\psi)B^{\mathrm{R}}]A^{\mathrm{L}}, \times b(1-\psi)C^{\mathrm{R}}\rho + b[I\psi + (1-\psi)D^{\mathrm{R}}]C^{\mathrm{L}}\rho + \chi\}. \tag{25}$$

The fixed point on the T-map is

$$A^{\mathrm{L}} = \{I - b[I\psi + (1-\psi)B^{\mathrm{R}}]\}^{-1}[a + b(1-\psi)A^{\mathrm{R}}] \tag{26}$$

$$C^{\mathrm{L}} = \{I - b[I\psi + (1-\psi)D^{\mathrm{R}}]\rho\}^{-1}[\varkappa + b(1-\psi)C^{\mathrm{R}}\rho]. \tag{27}$$

The fixed point for A^{L} and C^{L} corresponds to the RE equilibrium, namely $A^{\mathrm{L}} = [0, \pi^*]'$. This result is not surprising; it means that the presence of near-rationals does not alter the long-run behavior of the economy.

E-stability is obtained if all eigenvalues in the matrices $b[I\psi + (1-\psi)B^{\mathrm{R}}] - I$ and $b[I\psi + (1-\psi)D^{\mathrm{R}}]\rho - I$ have negative real parts. It is easy to check that for the parameterization considered, E-stability is obtained.[19]

Our approximate solution is useful in three respects. First, one can analyze the E-stability conditions directly. Second, when the economy converges, learners' parameters will not change and the solution is then exact. Third, our analysis helps identify the form of a solution that we may apply to a nonlinear case.

4.3. The Autocorrelated Shocks Case—Full Nonlinear Solution

The solution for near-rationals is accurate in the steady state but may perform poorly during a transition period. This section seeks a solution for rational agents that allows a different solution during the transition period. This is a standard procedure in models where the RE solution during a transition period may differ from the steady state solution, as for instance in Marcet and Marimon (1992).

We will apply the parameterized expectations algorithm, which is described in Lorenzoni and Marcet (1999). Rational agents must be able to accurately forecast variables. The goal of the algorithm is to find a forecasting function for rational agents. We first choose as an initial guess for RE the near-rationals solution. The first step is to use the guess in 1,000 simulations of the disinflation period for 20 quarters. In the second step, we regress y_{t+1} on the regressors of the near-rationals solution, $\{1, A_t^{\mathrm{L}}, \rho r_t^{\mathrm{n}}, C_t^{\mathrm{L}}\rho r_t^{\mathrm{n}}\}$. The third step is to form a new guess for the solution. The new guess is a weighted average between the previous

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

guess and the parameters estimated in the second step. We repeat the steps until the guess equals the estimated parameters. The RE solution is a fixed point in the previously described algorithm, at which rational agents accurately forecast next-period variables.[20] With the expectations of rational agents given by the solution found previously, we can then simulate the model.

4.4. The Tracking Parameters

Internal consistency criterion. Making the vector of the tracking parameters $\alpha = (\alpha_{\text{inf}}, \alpha_{\text{out}})$ endogenous avoids introducing additional degrees of freedom into the model and prevents agents from making big mistakes. Making the vector endogenous involves the concept of internal consistency (henceforth, IC), first introduced by Marcet and Nicolini (2003). For notational purposes, let $y_t(\alpha)$ denote the values generated in the economy when learners use α as tracking parameters, and let $E_t^{\bar{\alpha}} y_{t+1}(\alpha)$ denote learners' predictions at t of y_{t+1} when learners in the economy use α as tracking parameters and the predictions are made using $\bar{\alpha}$. For a given time horizon T and a positive number close to zero ε, the vector α is consistent if

$$\left\{\frac{1}{T}\sum_{t=1}^{T}[y_{t+1}(\alpha) - E_t^{\alpha} y_{t+1}(\alpha)]^2\right\} \leq \min_{\bar{\alpha}} E\left\{\frac{1}{T}\sum_{t=1}^{T}[y_{t+1}(\alpha) - E_t^{\bar{\alpha}} y_{t+1}(\alpha)]^2\right\} + \varepsilon. \tag{28}$$

That is, α is internally consistent if all learners use it and if predictions made by these tracking parameters are good when compared with predictions made by other tracking parameters $\bar{\alpha}$. We chose the time horizon to be 16 or 20 quarters, which is the duration of the Volcker disinflation episode. We computed results for $\varepsilon = 0.00001$, which approximately corresponds to 1.5% and 1% of the MSE for output and inflation when the tracking parameters are $(\alpha_{\text{inf}} = 0.5, \alpha_{\text{out}} = 0.3)$. Robustness analysis can be found in Appendix A. Expectations on equation (28) are computed by Monte Carlo integration using, 1000 simulations.

The results presented here refer to the case in which $\psi = 0.3$. We present results for inflation in Figure 3 and for output in Figure 4. In those figures, the horizontal axes represent the tracking parameters $(\alpha_{\text{inf}}, \alpha_{\text{out}})$, and the vertical axes represent $(\bar{\alpha_{\text{inf}}}, \bar{\alpha_{\text{out}}})$. Figure 3 and 4 show the results for $T = 20$ and $\varepsilon = 0.00001$. In these figures a value of 1 means that condition (28) is met, and a value of 0 means that $(\bar{\alpha_{\text{inf}}}, \bar{\alpha_{\text{out}}})$ is inefficient given that all learners use $(\alpha_{\text{inf}}, \alpha_{\text{out}})$. If a 1 occurs in the diagonal it means that $(\alpha_{\text{inf}}, \alpha_{\text{out}})$ is internally consistent. When predicting inflation, one can see in Figure 3 that $\alpha_{\text{inf}} = 0.5$ is always optimal. In Figure 4 we see that when α_{inf} is low, $\alpha_{\text{out}} = 0.5$ predicts well. When α_{inf} is high, $\alpha_{\text{out}} = 0.1$ becomes better. For intermediate values of α_{inf}, the best predictor is $\alpha_{\text{out}} = 0.3$. The pair $(\alpha_{\text{inf}} = 0.5, \alpha_{\text{out}} = 0.3)$ is internally consistent, and the robustness analysis showed that the pair $(\alpha_{\text{inf}} = 0.5, \alpha_{\text{out}} = 0.1)$ can also pass this criterion. The first pair is the more robust, and the results are almost the same if we use the second pair.

out	inf																										sum
0,9	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
0,9	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
0,7	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,5	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,5	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,5	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
0,5	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,5	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,3	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,3	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,3	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
0,3	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,3	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,1	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,1	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,1	0,5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
0,1	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,1	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		0,1	0,1	0,1	0,1	0,1	0,3	0,3	0,3	0,3	0,3	0,5	0,5	0,5	0,5	0,5	0,7	0,7	0,7	0,7	0,7	0,9	0,9	0,9	0,9	0,9	out
		0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	inf

FIGURE 3. Internal consistency table for inflation.

out	inf																										sum
0,9	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,9	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,7	0,1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,5	0,9	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	8
0,5	0,7	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	8
0,5	0,5	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	8
0,5	0,3	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	8
0,5	0,1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0	8
0,3	0,9	0	1	1	0	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	11
0,3	0,7	0	1	1	0	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	11
0,3	0,5	0	1	1	0	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	11
0,3	0,3	0	1	1	0	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	11
0,3	0,1	0	1	1	0	0	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	11
0,1	0,9	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	10
0,1	0,7	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	10
0,1	0,5	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	10
0,1	0,3	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	10
0,1	0,1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	10
		0,1	0,1	0,1	0,1	0,1	0,3	0,3	0,3	0,3	0,3	0,5	0,5	0,5	0,5	0,5	0,7	0,7	0,7	0,7	0,7	0,9	0,9	0,9	0,9	0,9	out
		0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	0,1	0,3	0,5	0,7	0,9	inf

FIGURE 4. Internal consistency table for output.

Sargent (1982) reports four cases of hyperinflation that were stopped suddenly and argues that those cases are consistent with RE. In accordance with that study, Ball (1994) reports that when the inflation change is bigger, the sacrifice ratio is smaller. The present model can account for the previous observations. For bigger changes in inflation, the IC requirement will imply bigger tracking parameters, thus reducing the cost of the disinflation. That is, in the Volcker disinflation learners may have adapted slowly, whereas in hyperinflations, learners may have adapted very quickly.[21]

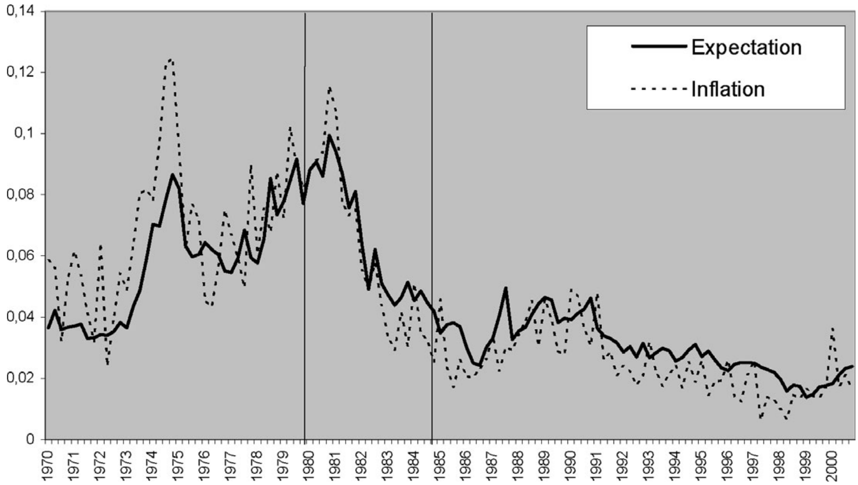

FIGURE 5. Survey of professional forecasters GDP deflator expectations.

Some Empirical Evidence. One would also like to know whether the tracking parameters that we found in the IC analysis reflect the actual learning behavior during the Volcker disinflation. The IC is a theoretical concept that may induce tracking parameters other than the ones used in reality. In equation (13), Nunes (2005) suggests that 30% of agents made forecasts matching those in the Survey of Professional Forecasters. Branch and Evans (2006) and the present paper suggest that surveys can reflect learning behavior. Figure 5 compares one-quarter-ahead expectations from the Survey of Professional Forecasters to one-quarter-ahead GDP inflation. For most of the disinflation period, expectations systematically overpredict inflation. Moreover, Figure 5 also shows that, in periods of rising inflation, expectations underpredict inflation; hence the overall picture is consistent with learning behavior. Surveys can be used to obtain an indicative measure of the tracking parameter. When inflation is regressed on inflation expectations, the results are[22]

$$\underset{}{\pi_t} = \underset{(0.003)}{-0.005} + \underset{(0.08)}{1.113} S_{t-1}\pi_t. \tag{29}$$

Because the hypothesis that the constant is zero and the coefficient on inflation expectations is 1 is not rejected, unbiasedness over the full sample cannot be rejected. The significance level of the test is 0.23. The tracking parameter is obtained by estimating equation $S_t\pi_{t+1} = (1-\alpha)S_{t-1}\pi_t + \alpha\pi_t$, which yields[23]

$$S_t\pi_{t+1} = \underset{(0.000)}{-0.000} + (1 - \underset{(0.043)}{0.175})S_{t-1}\pi_t + 0.175\pi_t. \tag{30}$$

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

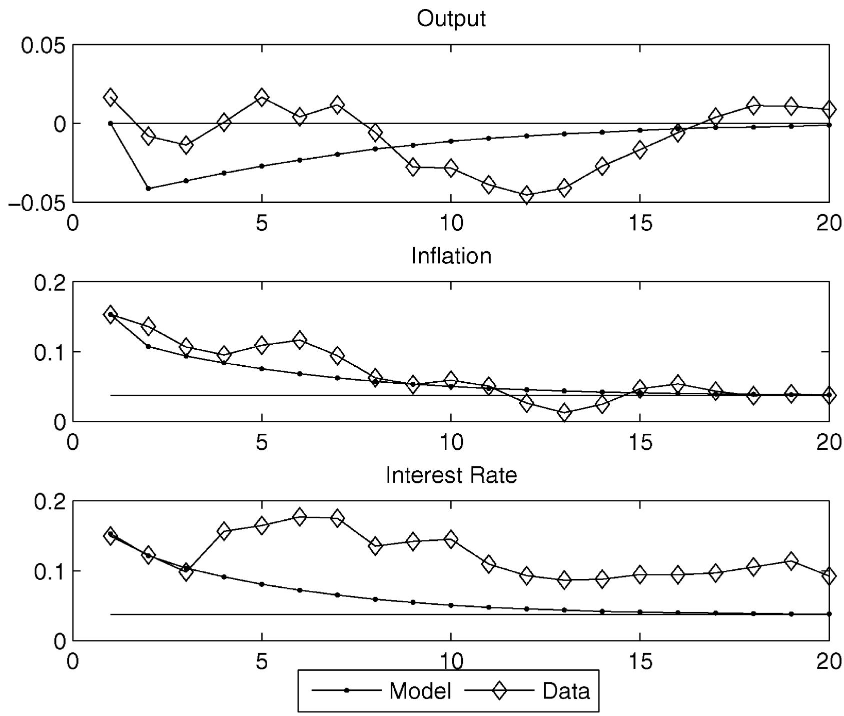

FIGURE 6. Learners' and rationals' economy.

The previous regression suggests that private agents used an average tracking parameter of 0.175 over the full sample. However, in periods of structural change, the tracking parameter is probably higher. Because the period of interest is the Volcker disinflation, we repeat the analysis for the subsample of 1980:Q1 to 1984:Q4. This corresponds to 20 periods, as considered in the IC analysis:

$$S_t\pi_{t+1} = \underset{(0.001)}{0.001} + (1 - \underset{(0.077)}{0.442})S_{t-1}\pi_t + 0.442\pi_t. \quad \textbf{(31)}$$

Note that the tracking parameter estimate, 0.442, is very close to the value obtained with the IC criterion.[24] It is not our claim that one can model survey expectations in such a simple way as we did here. Nevertheless, the fact that the estimated and the internally consistent tracking parameter are similar provides reassurance that our model yields a good description of the Volcker disinflation.

4.5. Evaluating the Heterogeneous Model

Figure 6 plots the average paths of inflation, the output gap, and the interest rate, assuming $\alpha_{\text{out}} = 0.3$, $\alpha_{\text{inf}} = 0.5$, and $\psi = 0.3$. Assuming that the disinflation lasts

https://doi.org/10.1017/S136510050807051X Published online by Cambridge University Press

TABLE 2. Convergence to new steady state

	Output		Inflation	
	$t = 16$	$t = 20$	$t = 16$	$t = 20$
Mean	−0.0035	−0.0011	0.0401	0.0382
St. Dev.	0.0192	0.0194	0.0111	0.0107
End	15		15	

for 16 quarters, the sacrifice ratio is 0.56%, a value similar to the computations in this paper but lower than the estimates of Ball (1994). The economy where learners and rationals coexist can better capture some features of the data.

The model captures the reduction in the interest rate. But the interest rate in the model is below that in the data. A potential explanation is that during the Volcker disinflation, RE agents thought the disinflation could be abandoned. Although the central bank implemented a disinflation, medium- and long-run interest rates stayed high. The present paper advances the simplest model as the benchmark, but both the yield curve and the possibility of abandoning the disinflation policy might have played a role in the real data.

In line with the data, the model predicts a temporary recession. The recession in the model is immediate, whereas in reality it developed more slowly. The basic fact that the disinflation was costly is captured by the model. The timing of the recession could be improved by introducing adjustment costs or a time-to-build feature.

More importantly, in the model, inflation gradually moves to target and matches the data quite well. Overall, the model matches the key stylized facts of disinflations. Inflation moves down slowly, the economy experiences a temporary recession, and in the long run, interest rates are reduced. Also, convergence to the new steady state is not too slow. As shown in Table 2, the economy is very near its new steady state after 15 quarters.

One can also compare the welfare loss of the disinflation instead of the sacrifice ratio. We consider that welfare is measured by quadratic deviations of output and inflation from their targets, and the weight on output deviations is 0.048.[25] Obviously, the welfare loss is 0 when there are no learners. If only learners exist, the welfare loss is −0.1424; for the model with both types of agents, the loss is −0.0309.

4.6. Further Exercises

The NK model is linearized around a zero steady state value of inflation. Whether this local approximation is a good one for the problem at hand is not entirely clear. We did not follow the approach of pursuing a global approximation because learning is better understood in linear models. We checked whether the results changed if we kept the magnitude of the disinflation, but assumed a zero inflation steady state at the end of the disinflation. When we simulated the model with a disinflation from 11.6% to 0%, the results did not change. In accordance with the previous observation, in Ball (1994) the initial level of inflation has no clear effect on the sacrifice ratio.

We also considered another version of the Phillips curve, in which the firms that do compute optimal prices adjust prices by the inflation target. Setting $\beta = 0.99$, the results with this Phillips curve are very similar to the results presented previously. In this case the Phillips curve is described by

$$\pi_t = (1 - \beta)\pi^* + \kappa z_t + \beta \tilde{E}_t \pi_{t+1}. \tag{32}$$

Using the same calibration as in the learning model, we also performed the disinflation experiment with rule-of-thumb firms, as in equation (10). Output and inflation were back to target in four and five quarters, respectively, and the sacrifice ratio was 0.09%, an extremely low value when compared with the data. The rule-of-thumb firms give less importance to past data than learning firms. The speed at which rule-of-thumb firms discount past data is not supported by the IC criterion. Moreover, the empirical analysis of survey expectations also did not support the rule-of-thumb specification.

5. CONCLUSIONS

This paper proposed a framework in which expectations are a weighted average of RE and learning. We discussed some difficulties that arise in this setup and how to tackle them. The model is similar in spirit to the rule-of-thumb literature that supports, for instance, the hybrid Phillips curve. The paper argued that including learning makes for a more solid model than does the rule-of-thumb approach. The learning algorithm uses the same functional form of RE, and one can determine the optimal learning speed. In general, the rule of thumb will not discount past data at an optimal speed, whereas learning does.

To evaluate the model, we provided an application to the Volcker disinflation. We used the standard NK model as a benchmark. This model under RE cannot account for the empirical patterns of unanticipated credible disinflations. When we assumed that all agents were learning, the qualitative features of the disinflation were matched. Nevertheless, convergence to the new steady state seemed to be too slow. When expectations were set as an average of RE and learning, then the disinflation was roughly matched both qualitatively and quantitatively.

This work has some implications for the debate on the hybrid Phillips curve. Although we managed to match the transitions of the Volcker disinflation, the model with rule-of-thumb firms implies a transition that is too fast. The main issue is that rule-of-thumb firms discount past data at a very high and nonoptimal pace. The literature has also debated whether the hybrid Phillips curve is mainly forward-looking. Although the present experiment is not a definitive test, our results suggest that the Volcker disinflation can be roughly matched when the weight on learning is around 30%. Even though the weight put on learning seems to be small, the model behaves in a way very different from the pure RE case. Future research could well tackle the question of whether other episodes and types of agents can be successfully modeled with the sort of expectations described here.

NOTES

1. Other related applications of Bayesian learning are in Andolfatto and Gomme (2003) and Schorfheide (2005).
2. Another work on imperfect credibility is Ball (1995).
3. The pricing assumption that this model uses was first suggested by Calvo (1983).
4. Preston (2005, 2006) points out that the model derived under learning would be different. We do not tackle this issue here and introduce learning from the reduced form equations, as is done, for instance, in Bullard and Eusepi (2005).
5. Further details on MSV solutions are in McCallum (1983).
6. For instance, Clarida et al. (1999), Fuhrer (1997), Fuhrer and Moore (1995), Roberts (1997, 1998).
7. The data appendix contains the description of the data series used throughout the paper.
8. We are assuming that output grows log-linearly during 16 quarters or 20 quarters and the change in inflation is 11.6%.
9. The value of $\kappa = 0.026$ in equation (13) has to be multiplied by four to consider annualized inflation. A similar adjustment was performed for σ.
10. Adam and Padula (2003), Carroll (2003), Kozicki and Tinsley (2002), and Roberts (1997, 1998), have pointed to the importance of survey expectations in determining the dynamics of inflation.
11. The coefficients on survey and rational expectations have been constrained to sum to 1. The results are very similar for the unrestricted estimation. The sample period is 1968:Q4 to 2004:Q2. The set of instruments contains four lags of inflation and two lags of marginal cost, wage inflation, output gap, and expected inflation; the results are robust to the inclusion of other instruments. The equation is estimated with GMM, and the J-test is 0.6181. In the first-stage regression, the F-statistic is 45 with a p-value of 0.00.
12. The estimation follows Galí and Gertler (1999) and Galí et al. (2005) in estimating the NKPC with marginal cost. In contrast, Rudd and Whelan (2005) argue that the NKPC with marginal cost provides a poor model of price inflation.
13. For the estimation of the variance–covariance matrix, the private sector is assumed to use a low tracking parameter of 0.05. Such a low value is optimal because even with the regime change, the steady state values of this matrix are not altered. Assuming other values for the parameter does not change the results.
14. Further details are in the Appendix on the internal consistency criterion.
15. The standard deviation at a given date is given by

$$\sqrt[2]{\frac{\sum_{n=1}^{\mathrm{n}} (x_{t,n} - \bar{x}_t)^2}{N}}.$$

The realization is denoted by n, and the mean at date (t) is given by $\bar{x}_t$. The mean $\bar{x}_t$ is reported in the first row of Table 1.

16. A discussion is in Uhlig (1999).
17. This approach closely follows Evans and Honkapohja (2001, Section 3.3).
18. Molnar (2004) used the method derived here and assumed that the proportion of near rationals depends on their forecasting performance. Confirming our results, the author concluded that agents with more rationality make convergence faster.

19. Evans and Honkapohja (1998) showed that E-stability implies local convergence of the learning algorithm in a class of models that contain the NK framework. The difference between the present model and the NK framework are the matrices that constitute the T-map. The results of Evans and Honkapohja (1998) can be applied to the model presented in this section. Also note that when the economy converges, as is assumed by the E-stability concept, near-rationals become completely rational and do not commit mistakes. Convergence conditions under recursive least squares and tracking are not always the same; simulation-based results suggest that if agents use a tracking algorithm, the economy also converges to equilibrium.

20. We also tested for different functional forms of the solution for rational agents; the results were robust. In particular, we checked how the mistakes of near-rationals could be improved by the regressors $1/t$, $1/t^2$. This solution undoubtedly imposes continuity between the transition period solution and the near-rationals solution.

21. In addition, the experiences reported in Sargent (1982) involved drastic policy changes in the exchange rate regime and fiscal stance that cannot be modeled here; such prominent policy changes may also have led learners to make a rapid adaptation. Also, the Calvo pricing assumption may be inappropriate in hyperinflationary environments.

22. We are considering a Newey West correction of 12 lags.

23. To check for robustness, we also estimated $S_t\pi_{t+1} = (1-\alpha)S_{t-2}\pi_{t-1} + \alpha\pi_{t-1}$. The results do not change.

24. The R^2 of this regression is 0.49 and the Durbin–Watson statistic is 2.3.

25. The weight on the output gap is the one assumed in Woodford (2003).

REFERENCES

Adam, K. and M. Padula (2003) Inflation Dynamics and Subjective Expectations in the United States. Working paper, European Central Bank.

Andolfatto, D. and P. Gomme (2003) Monetary policy regimes and beliefs. *International Economic Review* 44 (1), 1–30.

Ball, L. (1994) What determines the sacrifice ratio? In N.G. Mankiw (ed.), *Monetary Policy*, pp. 155–182. Chicago: University of Chicago Press.

Ball, L. (1995) Disinflation and imperfect credibility. *Journal of Monetary Economics* 35, 5–24.

Baxter, M. and R. King (1999) Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics* 81 (4), 575–593.

Blanchard, O.J. and C.M. Kahn (1980) The solution of linear difference models under rational expectations. *Econometrica* 48, 1305–1316.

Branch, W. and C. Evans (2006) A simple recursive forecasting model. *Economics Letters* 91, 158–166.

Bullard, J. and S. Eusepi (2005) Did the great inflation occur despite policymaker commitment to a Taylor rule? *Review of Economic Dynamics* 8, 324–359.

Bullard, J. and K. Mitra (2002) Learning about monetary policy rules. *Journal of Monetary Economics* 49, 1105–1129.

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

Carroll, C. (2003) Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118 (1), 269–298.

Clarida, R., J. Galí and M. Gertler (1999) The science of monetary policy: A new Keynesian perspective. *Journal of Economic Literature* 37 (4), 1661–1707.

Erceg, C.J. and A.T. Levin (2003) Imperfect credibility and inflation persistence. *Journal of Monetary Economics* 50, 915–944.

Evans, G. and S. Honkapohja (1998) Economic dynamics with learning: New stability results. *Review of Economic Studies* 65, 23–44.

Evans, G. and S. Honkapohja (2001) *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Evans, G. and S. Honkapohja (2003) Expectations and the stability problem for optimal monetary policies. *Review of Economic Studies* 70, 807–824.

Evans, G., S. Hokapohja, and T.J. Sargent (1993) On the preservation of deterministic cycles when some agents perceive them to be random fluctuations. *Journal of Economic Dynamics and Control* 17, 705–721.

Ferrero, G. (2007) Monetary policy, learning and the speed of convergence. *Journal of Economic Dynamics and Control* 31 (9), 3006–3041.

Fuhrer, J. (1997) The (un)importance of forward looking behavior in price specifications. *Journal of Money, Credit, and Banking* 29 (3), 338–350.

Fuhrer, J. and G. Moore (1995) Inflation persistence. *Quarterly Journal of Economics* 110 (1), 127–159.

Galí, J. (2007) *Monetary Policy, Inflation and the Business Cycle*. Princeton, NJ: Princeton University Press, forthcoming.

Galí, J. and M. Gertler (1999) Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics* 44, 195–222.

Galí, J., M. Gertler and J.D. López-Salido (2005) Robustness of the estimates of the hybrid new Keynesian Phillips curve. *Journal of Monetary Economics* 52, 1107–1118.

Giannitsarou, C. (2003) Heterogeneous learning. *Review of Economic Dynamics* 6, 885–906.

Giannitsarou, C. (2006) Supply-side reforms and learning dynamics. *Journal of Monetary Economics* 53, 231–306.

Kozicki, S. and P.A. Tinsley (2002) Alternative Sources of the Lag Dynamics of Inflation. Working Paper, Federal Reserve Bank of Kansas City.

Linde, J. (2005) Estimating new-Keynesian Phillips curves: A full information maximum likelihood approach. *Journal of Monetary Economics* 52, 1135–1149.

Lorenzoni, G. and A. Marcet (1999) Parameterized expectations approach. Some practical issues. In R. Marimon and A. Scott (eds.), *Computational Methods for the Study of Dynamic Economies*. pp. 143–171. Oxford, UK: Oxford University Press.

Mankiw, N.G. and R. Reis (2002) Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics* 117 (4), 1295–1328.

Marcet, A. and R. Marimon (1992) Communication, commitment and growth. *Journal of Economic Theory* 58 (2), 219–249.

Marcet, A. and J. Nicolini (2003) Recurrent hyperinflations and learning. *American Economic Review* 93 (5), 1476–1498.

Marcet, A. and T. Sargent (1989) Convergence of least-squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory* 48, 337–368.

McCallum, B. (1983) On nonuniqueness in rational expectations models: An attempt at perspective. *Journal of Monetary Economics* 11, 139–168.

Milani, F. (2005) Adaptive Learning and Inflation Persistence. Manuscript, University of California Irvine.

Molnar, K. (2004) Learning with Expert Advice. Mimeo, Universitat Pompeu Fabra.

Nunes, R.C. (2005) Inflation Dynamics: The Role of Expectations. Working paper, SSRN.

Orphanides, A. and J. Williams (2005) Expectations, learning and monetary policy. *Journal of Economic Dynamics and Control* 29 (11), 1807.

Orphanides, A. and J. Williams (2006) Monetary policy with imperfect knowledge. *Journal of the European Economic Association* 4, 366–375.

Preston, B. (2005) Learning about monetary policy rules when long-horizon expectations matter. *International Journal of Central Banking* 1 (2), 81–126.

Preston, B. (2006) Adaptive learning, forecast-based instrument rules and monetary policy. *Journal of Monetary Economics* 53 (3), 507–535.

Primiceri, G. (2006) Why inflation rose and fell: Policymakers' beliefs and U.S. Postwar stabilization policy. *Quarterly Journal of Economics* 121, 867–901.

Roberts, J. (1997) Is inflation sticky? *Journal of Monetary Economics* 39, 173–196.

Roberts, J. (1998) Inflation Expectations and the Transmission of Monetary Policy. Finance and Economics Discussion Paper No. 98-43, Board of Governors of the Federal Reserve System, Washington, DC.

Rudd, J. and K. Whelan (2005) Does labor's share drive inflation? *Journal of Money, Credit, and Banking* 37, 297–312.

Santoro, S. (2003) Adaptive Learning and Inflation Dynamics in a Flexible Price Model. Manuscript, Universitat Pompeu Fabra.

Sargent, T. (1982) The ends of four big inflations. In R. Hall (ed.), *Inflation: Causes and Effects*, pp. 41–97. Chicago: University of Chicago Press.

Sargent, T.J. (1999) *The Conquest of American Inflation*. Princeton, NJ: Princeton University Press.

Schorfheide, F. (2005) Learning and monetary policy shifts. *Review of Economics Dynamics* 8, 392–419.

Taylor, J. (1993) Discretion versus policy rules in practice. *Carnegie–Rochester Conference Series on Public Policy* 39, 195–214.

Uhlig, H. (1999) A toolkit for analyzing nonlinear dynamic stochastic models easily. In R. Marimon and A. Scott (eds.), *Computational Methods for the Study of Dynamic Economies*, pp. 30–61. Oxford, UK: Oxford University Press.

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

APPENDIX A: INTERNAL CONSISTENCY

The magnitude of shocks influences the prediction performance of tracking parameters. Insofar as there is a regime change, it is optimal to give more importance to recent observations. On the other hand, shocks create noise in the economy, and if one gives more importance to recent observations, then predictions will be distorted by recent shocks.

We first assumed that there are an autocorrelated natural interest rate and a nonautocorrelated cost-push shock with standard deviations set to 0.015 and 0.01 respectively. We also did sensitivity analysis by considering the presence of other shocks and different standard deviations. The results were robust. In one alternative specification, we considered two nonautocorrelated shocks that influence equation (12) directly.

We estimated the data standard deviations of the actual output gap and of inflation from 1980 to 1984 to be 0.02 and 0.04, respectively. We set the magnitude of shocks so that for the internally consistent tracking parameter, the model yields plausible variances when compared with the data for the Volcker disinflation. Note, however, that this implies variances that are too high for the period after the disinflation.

In the model with $\psi = 1$, the internally consistent parameters correspond to $\alpha_{\text{out}} = 0.5$ and $\alpha_{\text{out}} = 0.7$ and to α_{inf} ranging from 0.5 to 0.9. For any of these parameters the simulated disinflation is slower than the Volcker disinflation.

For $\psi = 0.3$, the internally consistent parameters were $\alpha_{\text{inf}} = 0.5$, $\alpha_{\text{out}} = 0.3$, as can be seen in Figures 3 and 4. We conducted sensitivity analysis assuming an alternative value of $T = 16$ and $\varepsilon = 0.00002$; the results were robust. The parameters $\alpha_{\text{inf}} = 0.5$, $\alpha_{\text{out}} = 0.1$ are also internally consistent for $\varepsilon = 0.00002$, but the simulated series are very similar under the two sets of parameters.

APPENDIX B: DATA

Inflation was computed using the seasonally adjusted monthly consumer price index for all urban consumers and all items, as computed by the Bureau of Labor Statistics. Quarterly inflation is computed as the sum for the months in the quarter divided by the sum for the months in the previous quarter. The reported series were filtered using the bandpass filter, eliminating components with periodicity of less than 4 quarters.

We used the GDP series at constant prices and seasonally adjusted. A first series was computed by eliminating the components with a periodicity of less than 32 quarters. A second series was computed by eliminating components of less than 4 quarters. The output gap is computed as the log of the second series divided by the first series.

The actual federal funds interest rate source is published by the Federal Reserve System. The monthly annualized rates were transformed to quarterly annualized rates using a geometric average.

The Survey of Professional Forecasters reports the mean values of the predicted GDP deflator for the current quarter and the following quarter. Expected inflation is computed as the annualized change from the current quarter prediction to the next quarter prediction.