

# Adaptive Expectations, Underparameterization and the Lucas Critique\*

George W. Evans  
University of Oregon

Garey Ramey  
University of California, San Diego

November 29, 2001

## Abstract

A striking implication of the replacement of adaptive expectations by Rational Expectations was the “Lucas Critique,” which showed that expectation parameters, and endogenous variable dynamics, depend on policy parameters. We consider this issue from the vantage point of bounded rationality, where for transparency we model bounded rationality by means of simple adaptive expectations. We show that for a range of processes, monetary policy remains subject to the Lucas critique. However, there are also regimes in which the expectation parameter is locally invariant and the Lucas critique does not apply.

*JEL classification:* E52, E31, D83, D84.

*Keywords:* Lucas critique, expectations, underparameterization, bounded rationality, learning.

---

\*Corresponding author: Professor George W. Evans, Department of Economics, 1285 University of Oregon, Eugene, OR 97403-1285.  
Phone: (541)-346-4662. Fax: (541)-346-1243. email: [gevans@oregon.uoregon.edu](mailto:gevans@oregon.uoregon.edu).

# 1 Introduction

The adaptive expectations hypothesis was introduced by Cagan (1956) and Friedman (1957) as a plausible and empirically meaningful approach to modeling expectations of future variables in a world of uncertainty. Their apparent empirical success led to widespread utilization of the adaptive expectations hypothesis before it was ultimately swept away by the rational expectation revolution, initiated by Muth (1961) and advanced by Lucas (1976) and Sargent and Wallace (1975). Rational expectations has the great advantage of providing *optimal* expectations; under the standard of optimality, adaptive expectations suffers by comparison and should be rejected.

One of the most salient implications of Rational Expectations is the critique of traditional policy making presented in Lucas (1976). The traditional theory of economic policy is characterized as treating the time series process followed by the economy as fixed and invariant with respect to exogenous changes in policy. Under rational expectations, however, the forecast or expectation rule will be affected by policy changes and, if the economy is in turn affected by expectations, these will alter the time series process followed by the economy. Lucas provided examples of this phenomenon based on prominent macroeconomic models.

Our objective here is to reconsider the Lucas critique in the context of adaptive expectations. The starting point of our argument is Muth (1960). In that paper Muth showed that adaptive expectations, with an appropriate adaptation parameter, are fully rational if the variable being forecasted follows an exogenous IMA(1,1) stochastic process, i.e. if the first difference of the variable is a first-order moving average process. Rational expectations, however, assumes that the true process generating the data is known, an assumption that many feel to be implausibly strong. Recently Evans and Honkapohja (1993,2001) and Sargent (1999) have argued that adaptive expectations may be a reasonable, if not fully rational, forecast method when the true process is unknown.

We consider a simple expectations augmented Phillips curve model in which the monetary policy rule follows an unknown regime switching process. In a world in which the true data generating process is complex, economic agents can be expected to use simple underparameterized representations of the process to make their forecasts. By appropriately tuning the free parameters of the forecast rule, they can obtain the best forecast rule within this class. That is, an appropriate bounded rationality assumption is that

agents, in the terminology of Sargent (1999, Ch. 6) have “optimal misspecified beliefs.” We choose Cagan-Friedman adaptive expectations for our underparameterized class of expectations because of their simplicity. Using this model, we are able to demonstrate vividly the scope for the Lucas critique to operate in a setting of bounded rationality. These findings suggest that the Lucas critique may play an important role under more generalized classes of underparameterized expectation functions.

Our central results are as follows. Even in the special case in which expectations do not affect realized inflation, so that inflation is exogenous, changes in the economic policy process may induce changes in the forecast rule and therefore in the stochastic process followed by other economic variables. That is, the Lucas critique has a range of validity even when expectations are not rational, but formed adaptively. The reason is that the optimal choice of the adaptive expectations parameter may depend on the stochastic process followed by policy. However, this is not always the case. For some parameter settings the optimal choice of the adaptive expectation parameter is to set it for maximal filtering. When this is the case, small changes in policy parameters have no effect on the expectation function and the Lucas critique does not apply.

When expectations have a nonzero effect on inflation, some new features emerge. We show how the size of the region in which the Lucas critique applies can be expanded as a consequence of feedback between agents’ choices of learning rules. Further, the Nash equilibrium choice of  $\gamma$  may differ from the socially optimal setting. This arises because the feedback induces a forecast externality.<sup>1</sup>

Section 2 presents the model, analysis is carried out in Sections 3 and 4, and Section 5 concludes.

## 2 Model

### 2.1 Economic Structure

We consider a simple macroeconomic model, inspired by Lucas (1973) and Fischer (1977), in which aggregate output is affected by unanticipated price

---

<sup>1</sup>Throughout this paper we follow Lucas (1976) and treat policy as a stochastic process set once and for all by the government. The recent book by Sargent (1999), which discusses some of the issues treated here, emphasizes an adaptive approach to economic policy.

level changes. Let aggregate supply be specified as follows:

$$q_t = \alpha(p_t - p_t^e), \quad (1)$$

where  $q_t$  and  $p_t$  denote aggregate output and the aggregate price level in period  $t$ , and  $p_t^e$  indicates private agents' expectation of  $p_t$  formed in period  $t-1$ . Equation (1) is, of course, the simplest form of the standard expectations augmented Phillips curve. Aggregate demand is given by

$$q_t = m_t - p_t + \eta_t, \quad (2)$$

where  $m_t$  denotes the nominal money stock in period  $t$ , and  $\eta_t$  represents an exogenous white noise process with variance  $Var(\eta)$ .

Finally, the monetary authority determines the money stock according to the following policy rule:

$$m_t = p_{t-1} + \mu_t, \quad (3)$$

where  $\mu_t$  is a *regime switching process* defined as follows:

$$\mu_t = \begin{cases} \mu_{t-1} & \text{with probability } 1 - \varepsilon \\ \nu_t & \text{with probability } \varepsilon \end{cases},$$

where  $0 \leq \varepsilon \leq 1$  and  $\nu_t$  is an independently and identically distributed process with an arbitrary distribution having mean  $\bar{\nu}$  and variance  $Var(\nu)$ . Thus, apart from the random shocks, the money supply fully accommodates any changes to the previous period's price level. This form of the policy rule is chosen in order to present most simply the central ideas of the paper.

The model may be reexpressed in terms of inflation rates. Let the inflation rate be denoted by  $\pi_t = p_t - p_{t-1}$ , and let  $\pi_t^e$  denote inflation expectations formed in period  $t-1$ , i.e.,  $\pi_t^e = p_t^e - p_{t-1}$ . Combining (1), (2) and (3), we have

$$\pi_t = \frac{\alpha\pi_t^e + \mu_t + \eta_t}{1 + \alpha}. \quad (4)$$

We remark that the structure (4) arises in other economic models. An important example is the Muth (1961) "cobweb model" of an isolated market. Demand for output  $q_t$  depends negatively on price  $p_t$ , with inverse demand curve given by  $p_t = x_t - bq_t$  for some exogenous process  $x_t$ . Supply depends positively on expected price,  $q_t = cp_t^e$ , where  $p_t^e$  denotes the price at  $t$  expected by agents in  $t-1$ . If the exogenous component of demand follows the above regime switching process  $x_t = \mu_t + \eta_t$  then we obtain the reduced form

$p_t = -bcp_t^e + \mu_t + \eta_t$ . This has the same form as (4) except that now there is a negative feedback from expected prices. The analysis of this paper applies equally well to this set-up. A special case of interest is if  $b = 0$  so that demand follows an exogenous stochastic process, as would be appropriate in an open economy or for a monopolist with infinitely elastic demand. This leads to the zero feedback case analyzed in Section 3.

## 2.2 Expectation Formation

We turn now to the problem of how agents form inflation forecasts. Assume for convenience that at time  $t - 1$  the whole history of data  $\{\pi_i\}_{i=-\infty}^{t-1}$  is available. Let  $\beta_{t-1} = \pi_t^e$  denote the forecast of  $\pi_t$  made at time  $t-1$ . If agents attempt to minimize the Mean Square Error of one-step ahead forecasts, then the problem is to choose  $\beta_{t-1}$  to minimize

$$MSE = E(\pi_t - \beta_{t-1})^2.$$

The standard rational expectations forecasting procedure assumes full knowledge of the true process generating  $\pi_t$ , and the form of the optimal forecasting rule would be closely linked to that of the true process. We instead consider the problem of forecasting  $\pi_t$  when the forecaster *does not* know the true data-generating process.

In contrast to the rational expectations procedure, we assume that the forecaster might misspecify the true process by using a forecast rule that is too simple to make fully-rational forecasts possible. To make our analysis transparent, the class of forecast rules we consider is a mild extension of the venerable adaptive expectations class introduced by Friedman and Cagan. In particular, we restrict attention to forecasts derived from the following one-parameter class:

$$\beta_t = \gamma \sum_{i=0}^{\infty} (1 - \gamma)^i \pi_{t-i}, \text{ where } 0 \leq \gamma \leq 1. \quad (5)$$

In the case in which  $\gamma = 0$ , we interpret this sum as

$$\beta_t = \lim_{T \rightarrow \infty} T^{-1} \sum_{i=0}^T \pi_{t-i}. \quad (6)$$

Existence of the limit is shown below. For  $0 < \gamma < 1$  this class of forecasts is called exponential smoothing, and for  $\gamma = 0$  it reduces to the sample mean.

The forecasts (5) and (6) can be represented recursively by

$$\beta_t = \beta_{t-1} + \gamma_t(\pi_t - \beta_{t-1}),$$

where  $\{\gamma_t\}_{t=1}^{\infty}$  is a nonstochastic gain sequence. For  $\gamma_t = \gamma$ , where  $0 < \gamma \leq 1$ , we have

$$\beta_t = \beta_{t-1} + \gamma(\pi_t - \beta_{t-1}). \quad (7)$$

This is a *fixed gain rule*, corresponding to (5), and gives the classic Cagan-Friedman *adaptive expectations* formula. For  $\gamma_t = 1/t$  we obtain

$$\beta_t = \beta_{t-1} + t^{-1}(\pi_t - \beta_{t-1}), \quad (8)$$

which corresponds to (6), i.e., the sample average, where for convenience we now treat the system as starting at  $t = 0$ . This is known as a *decreasing gain rule*. Note also that this can be thought of as least squares learning when  $\pi_t$  is modeled as following an *i.i.d.* process with a fixed but unknown mean.

Intuitively, the optimal choice of the gain parameter  $\gamma$  depends on the relative importance of tracking versus filtering the observed data. Tracking indicates the responsiveness of the forecast rule to a time-varying conditional mean, while filtering refers to the ability of the forecast rule to eliminate data-induced noise in the forecasts. A high value of  $\gamma$  increases tracking, but sacrifices filtering. The choice  $\gamma = 0$  represents maximal filtering with no tracking.

### 2.3 Equilibrium

We consider Nash equilibria of the model, in which all agents select a common value of  $\gamma$  that minimizes MSE. In particular,  $\gamma$  is a Nash equilibrium if it minimizes *MSE* for each agent, given the inflation process that arises when all other agents in the economy choose  $\gamma$ .

The equilibrium path of inflation expectations is determined as follows. If  $\gamma > 0$ , then inflation expectations satisfy  $\pi_t^e = \beta_{t-1}$ , where  $\beta_t$  is formed using the exponential smoothing class (5). For the choice  $\gamma = 0$ , combine (4), (8) and  $\pi_t^e = \beta_{t-1}$  to obtain

$$\beta_t = \beta_{t-1} + t^{-1}\theta(\mu_t + \eta_t - \beta_{t-1}), \quad (9)$$

where  $\theta = (1 + \alpha)^{-1}$ . This set-up fits the standard Robbins-Monroe formulation for stochastic recursive algorithms. Applying these stochastic approx-

imation results,<sup>2</sup> it can be shown that, for all initial  $\beta_0$ , with probability one  $\beta_t$  converges under (9) to the unique equilibrium of the associated ordinary differential equation  $\frac{d\beta}{d\tau} = \theta(E(\mu_t + \eta_t) - \beta)$ . Since  $E(\mu_t + \eta_t) = \bar{\nu}$ , we have  $\beta_t \rightarrow \bar{\nu}$  under (9) with probability one. In view of (6), it follows that  $\beta_t = \bar{\nu}$  with probability one.

### 3 Zero Feedback Case: Theoretical Results

We begin by considering the case of  $\alpha = 0$ , in which realized inflation is unaffected by inflation expectations. This eliminates feedback between agents' expectation formation decisions, allowing closed-form solutions for the Nash equilibrium to be obtained. The case of  $\alpha > 0$  is analyzed in the following section.

Given  $\alpha = 0$ , the inflation process is

$$\pi_t = \mu_t + \eta_t. \quad (10)$$

Since  $Cov(\eta_t, \beta_{t-1}) = 0$  and  $E\pi_t = E\beta_{t-1}$ , we see that  $MSE$  is given by

$$MSE = Var(\pi_t) - 2Cov(\mu_t, \beta_{t-1}) + Var(\beta_t). \quad (11)$$

In order to compute  $MSE$  for the case  $0 < \gamma \leq 1$ , we need the asymptotic moments of the stationary process implied by (7) and (10). These are computed in the following lemma.<sup>3</sup>

**Lemma 1.** *For the process defined by (7) and (10), we have*

$$Cov(\mu_t, \beta_{t-1}) = \frac{\gamma(1-\varepsilon)}{1-(1-\gamma)(1-\varepsilon)}Var(\nu), \quad (12)$$

$$Var(\beta_t) = \frac{\gamma}{2-\gamma} \left( Var(\eta) + \frac{1+(1-\gamma)(1-\varepsilon)}{1-(1-\gamma)(1-\varepsilon)}Var(\nu) \right). \quad (13)$$

---

<sup>2</sup>See Ljung (1977) or Ch. 6 of Evans and Honkapohja (2001). Note that under least squares learning,  $\gamma_t = 1/t$  satisfies the assumptions (i)  $\gamma_t > 0$  is a nonincreasing sequence, with (ii)  $\sum_{t=1}^{\infty} \gamma_t = \infty$  and (iii)  $\sum_{t=1}^{\infty} \gamma_t^2 < \infty$ .

<sup>3</sup>Proofs are found in the appendix.

Using (11), (12) and (13), the  $MSE$  when  $0 < \gamma \leq 1$  is given by

$$MSE = Var(\mu_t + \eta_t) + Var(\nu) \left\{ \frac{\gamma}{2 - \gamma} \frac{Var(\eta)}{Var(\nu)} - \frac{\gamma}{2 - \gamma} \frac{2 - 3\varepsilon - \gamma(1 - \varepsilon)}{\varepsilon + \gamma(1 - \varepsilon)} \right\}. \quad (14)$$

This leads to the following necessary first-order condition for the optimal choice  $\gamma^*$ :

$$\frac{2}{(2 - \gamma^*)^2} \frac{Var(\eta)}{Var(\nu)} - 2\varepsilon\Psi(\gamma^*) = 0, \quad (15)$$

where

$$\Psi(\gamma) = \frac{2 - 3\varepsilon - 2\gamma(1 - \varepsilon)}{(2 - \gamma)^2(\varepsilon + \gamma(1 - \varepsilon))^2}. \quad (16)$$

Necessary and sufficient conditions for nonzero  $\gamma^*$  are given in the following lemma.

**Lemma 2.**  $\gamma^* \in (0, 1)$  if and only if the following condition holds:

$$\varepsilon < \frac{2}{\frac{Var(\eta)}{Var(\nu)} + 3}. \quad (17)$$

In this case,  $\gamma^*$  is given implicitly by:

$$\frac{\varepsilon(2 - 3\varepsilon - 2\gamma^*(1 - \varepsilon))}{(\varepsilon + \gamma^*(1 - \varepsilon))^2} = \frac{Var(\eta)}{Var(\nu)}. \quad (18)$$

Now consider the  $MSE$  if  $\gamma = 0$ . Since  $\theta = 1$ , we have, with probability one, that

$$\beta_t = E\pi_t = E(\mu_t + \eta_t) = \bar{\nu}.$$

Hence when  $\gamma = 0$ , the  $MSE$  is given by  $MSE = Var(\mu_t + \eta_t)$ , i.e. by the limiting  $MSE$  under (5) as  $\gamma \rightarrow 0$ . We have arrived at the following result:<sup>4</sup>

**Theorem 1.** *The  $MSE$  is minimized by  $\gamma^* = 0$  when (17) fails to hold, and by the value  $\gamma^* \in (0, 1)$  given by (18) if (17) is satisfied.*

---

<sup>4</sup>We have excluded from consideration the choice  $\gamma < 0$ . However, we note that a value  $\gamma < 0$  cannot be optimal, since from (7) and (10) it follows that  $\beta_t$  would then follow an explosive process in which  $Var(\beta_t) \rightarrow \infty$  as  $t \rightarrow \infty$ .



Observe from (17) that when  $\varepsilon$  takes on a high value, associated with frequent switches in policy, agents select  $\gamma^* = 0$  in equilibrium. In this case, no attempt is made to learn about policy. Because policy changes frequently, policy shifts are tantamount to noise, and agents opt to filter noise out of their expectations by placing zero weight on past inflation observations. When the value of  $\varepsilon$  is low, in contrast, switches are less frequent, and agents find it beneficial to track the time-varying conditional mean of policy by setting  $\gamma^* > 0$ . Thus, characteristics of the policy rule feed back on the agents' choice of learning rule, based on the relative benefits of filtering versus tracking.

Note further that the upper bound of allowable  $\varepsilon$  in (17) becomes smaller as  $Var(\eta)$  grows relative to  $Var(\nu)$ . Higher  $Var(\eta)$  reflects an increase in background noise, which increases the attractiveness of filtering.

The theorem demonstrates that for a range of  $\varepsilon$ , a decline in  $\varepsilon$  leads agents to select higher values of  $\gamma^*$ . This relationship does not hold for all  $\varepsilon$  values, however. From (18) the following corollary is immediate.

**Corollary 1.**  $\gamma^* \rightarrow 0$  as  $\varepsilon \rightarrow 0$ .

Intuitively, as policy switches become very infrequent, the return to tracking them becomes small, and agents prefer to choose small values of  $\gamma^*$  in order to filter out noise. Full filtering emerges in the limit. It follows that the relationship between the persistence of policy and optimal tracking is nonmonotonic: as  $\varepsilon$  rises from zero, tracking first rises, then falls.

We close this section by considering the implications for the Lucas Critique in the context of adaptive expectations. We have seen that when agents underparameterize the forecast rule by using adaptive expectations with an optimally chosen adaption parameter  $\gamma^*$ , there are two regimes depending on the underlying exogenous stochastic process. Following Lucas (1976) we can consider the effect of exogenous changes in economic policy leading to changes in the  $\pi_t$  process.

**Corollary 2.** *Suppose  $\varepsilon < 2/(\frac{Var(\eta)}{Var(\nu)} + 3)$ . Then changes in  $Var(\eta)$ ,  $Var(\nu)$  or  $\varepsilon$  lead to changes in the Nash equilibrium forecast rule and the Lucas critique applies. However, if  $\varepsilon > 2/(\frac{Var(\eta)}{Var(\nu)} + 3)$  then for sufficiently small changes in  $Var(\eta)$ ,  $Var(\nu)$  or  $\varepsilon$ , the equilibrium rule remains  $\gamma^* = 0$  so that the forecast rule is not subject to the Lucas critique.*

Corollary 2 shows that, depending on the other parameters, changes in parameters of the policy rule may or may not induce changes in the forecast

rule. This result demonstrates that the Lucas critique may have a range of applicability within a setting of bounded rationality. More broadly, our findings reveal that the Lucas critique is not fundamentally tied to rational expectations. Once agents are allowed to vary the parameters of their learning rules in a purposeful way, the Lucas critique becomes salient even if rational expectations is impossible to acquire. The importance of the critique is then related to the degree of sensitivity of agents' parameter choices to the economic environment.

## 4 General Case

### 4.1 Feedbacks, Tracking and Inflation Persistence

We turn now to the general case, where  $\alpha > 0$  is considered. When  $\alpha$  is positive, agents' choice of learning rule affects the inflation process, generating a feedback to the choice of learning rule. The complications introduced by this feedback make it difficult to obtain full analytical results, and thus we limit our discussion to a local existence proof and to numerical examples for a range of parameters.

To define the Nash equilibrium we must now distinguish between the gain parameter  $\hat{\gamma}$  chosen by the representative agent and the value  $\gamma$  chosen by all other agents. For  $\gamma, \hat{\gamma} \in (0, 1]$  the system is thus defined by

$$\begin{aligned}\pi_t &= (1 - \theta)\beta_{t-1} + \theta(\mu_t + \nu_t), \text{ where } \theta = 1/(1 + \alpha) \\ \beta_t &= \gamma\pi_t + (1 - \gamma)\beta_{t-1} \\ \hat{\beta}_t &= \hat{\gamma}\pi_t + (1 - \hat{\gamma})\hat{\beta}_{t-1},\end{aligned}$$

where  $\hat{\beta}_t$  is the forecast at  $t$  of the representative agent and  $\beta_t$  is the forecast of all other agents. For the case  $\gamma = 0$  or  $\hat{\gamma} = 0$ ,  $\beta_t$  or  $\hat{\beta}_t$  is instead given by the mean of  $\pi_t$  as in (6). The representative agent attempts to minimize  $MSE(\gamma, \hat{\gamma}) = E(\pi_t - \hat{\beta}_{t-1})^2$  and a Nash equilibrium  $\gamma^*$  is defined by

$$\gamma^* = \arg \min_{\hat{\gamma} \in [0, 1]} MSE(\gamma^*, \hat{\gamma}).$$

As shown in the appendix, the theoretical results of the previous section carry over to the case of  $\alpha > 0$  sufficiently small. We have:

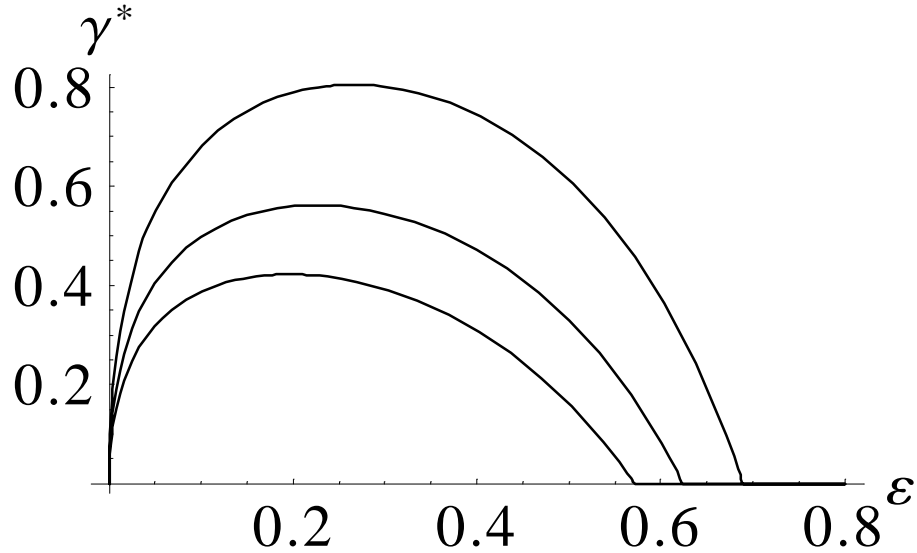


Figure 1:  $\gamma^*(\varepsilon)$  for  $\alpha = 0, 1/4, 2/3$

**Theorem 2.** *If  $\varepsilon < 2/(\frac{Var(\eta)}{Var(\nu)} + 3)$  then for all  $\alpha > 0$  sufficiently small there exists a Nash equilibrium with  $\gamma^* \in (0, 1)$ . If  $\varepsilon < 2/(\frac{Var(\eta)}{Var(\nu)} + 3)$  then for all  $\alpha > 0$  sufficiently small there exists a Nash equilibrium with  $\gamma^* = 0$ .*

We remark that Theorem 2 also applies to the Muth “cobweb” model, described at the end of Section 2.1, with the condition “ $\alpha > 0$  sufficiently small” replaced by “ $b > 0$  sufficiently small.” Figure 1 depicts the relationship between  $\gamma^*$  and  $\varepsilon$  for three values of  $\alpha$ , where  $Var(\eta) = 0.1$  and  $Var(\nu) = 0.2$  are specified. The upper curve shows the values of  $\gamma^*$  for  $\alpha = 2/3$ , the middle curve corresponds to  $\alpha = 1/4$  and the lower curve gives the values of  $\gamma^*$  for the zero feedback case  $\alpha = 0$ . The nonmonotonicity derived in the preceding section may be noted, as well as the existence of both  $\gamma^* > 0$  and  $\gamma^* = 0$  regimes in each case. Furthermore, higher values of  $\alpha$  lead to greater tracking for each  $\varepsilon$ . Feedbacks between learning rules thus serve to increase the attractiveness of tracking policy switches. Observe that as  $\alpha$  rises, the range of values for which the Lucas critique fails to apply, in the sense of Corollary 2, becomes smaller.

**Table 1: Autocorrelations of Simulated Inflation data.**

	$\alpha = 0$	$\alpha = 1/4$	$\alpha = 2/3$
$\varepsilon = 0.1$	0.60	0.70	0.84
$\varepsilon = 0.3$	0.47	0.57	0.75
$\varepsilon = 0.5$	0.35	0.42	0.59

Increased tracking has the added effect of introducing greater persistence into the inflation process. This is shown in Table 1, which reports first order autocorrelation coefficients for simulated inflation data using the Nash equilibrium learning parameter.<sup>5</sup> The  $\alpha = 0$  column indicates the autocorrelation introduced by the policy switching process itself, in the absence of any learning effects. These coefficients would be the same for the other  $\alpha$  values if agents selected  $\gamma = 0$ . However, positive choices of  $\gamma^*$  imply added inflation persistence. For  $\varepsilon = 0.1$ , increasing  $\alpha$  to 0.25 leads the autocorrelation coefficient to rise by 10 percentage points, based on the rise in tracking by agents. This effect leads to a further rise of 14 percentage points when  $\alpha$  is increased to 0.67. Results are similar for the other  $\varepsilon$  levels.

## 4.2 Learning Externalities

We close our examination by comparing the Nash equilibrium choices  $\gamma^*$  to the values that would be selected by a social planner who sought to minimize  $MSE$ . It can be verified that in the case with feedback  $MSE/\theta^2$  is given by the right-hand-side of (14) with  $\gamma$  replaced by  $\theta\gamma$ . Let  $\gamma^o$  denote the  $MSE$  minimizing level. It follows that  $\gamma^o$  is determined by the results of Section 3 with  $\gamma^*$  replaced by  $\theta\gamma^o$ .<sup>6</sup> In general the socially optimal choice of  $\gamma$  will not coincide with the Nash equilibrium value.

Figure 2 illustrates the Nash equilibrium and socially optimal values of  $\gamma$  as functions of  $\alpha$ , where  $\varepsilon = 0.3$ ,  $Var(\eta) = 0.1$  and  $Var(\nu) = 0.2$ . Clearly, Nash and optimal values are identical (at approximately  $\gamma^* = \gamma^o = 0.39$ )

<sup>5</sup>For the simulations, we specify that  $\eta$  and  $\nu$  are normally distributed, with  $\bar{\nu} = 1$ . We conduct a 10,200 period simulation and throw out the first 200 periods when computing the autocorrelation coefficients.

<sup>6</sup>Provided that  $\alpha > 0$  is not too large. For sufficiently large values of  $\alpha$  there is the possibility of an additional regime with  $\gamma^o = 1$ .

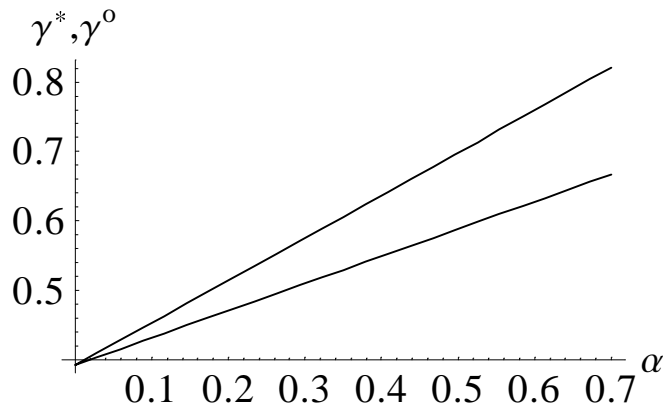


Figure 2:  $\gamma^*(\alpha)$ , upper, and  $\gamma^o(\alpha)$ , lower, for  $\varepsilon = 0.3$

for  $\alpha = 0$ , since there are no feedbacks across agents' choices. As  $\alpha$  rises, the private incentive to raise tracking increases faster than does the social incentive, and as a result the Nash equilibrium choices exceed the optimal levels. Here agents suboptimally overtrack the policy switches.

## 5 Conclusion

We have considered a simple monetary model in which agents are faced with a stochastic policy process which occasionally undergoes shifts in the mean. While the agents may have a rough idea that the system is time-varying, they are still unable to model explicitly the regime-switching process; as a consequence, they rely on simple forecast rules that underparameterize the true process. In particular we have considered the implications of agents using exponential smoothing forecasts with an optimally chosen smoothing parameter  $1 - \gamma$ . For  $0 < \gamma \leq 1$  this is the traditional adaptive expectations formula and for  $\gamma = 0$  it reduces to the sample mean. Even if the time variations of the mean are small, provided regime switches occur infrequently it will benefit agents to use an adaptive expectations forecast with  $0 < \gamma < 1$  rather than setting  $\gamma = 0$  as would be optimal in the absence of regime switches. However if regime switches occur sufficiently quickly, then the choice  $\gamma = 0$  is optimal because the importance of filtering then dominates.

We obtain simple analytical results for a version of the model in which

inflation follows an exogenous stochastic process. Even in this setting we can illustrate the basic implications for the Lucas critique of the use of underparameterized forecasting models with optimally chosen parameters. Provided the exogenous and policy parameter settings satisfy an appropriate condition, adaptive expectations with a choice of  $0 < \gamma < 1$  will be optimal and in this region the Lucas critique operates with (optimally chosen) adaptive expectations for the same reason that it applies with rational expectations. However, if this condition fails, then full filtering with  $\gamma = 0$  is optimal within our class of forecast rules and the Lucas critique does not apply for small changes in parameter setting.

Additional features arise when inflation is affected endogenously by the learning process. The feedback from expectations in this setting expands the parameter region within which the Lucas critique applies. The self-referential aspect also induces a forecast externality which, in the case  $0 < \gamma < 1$ , moves the Nash equilibrium value of  $\gamma$  away from the socially optimal value.

In the language of control theory, adaptive expectations is a simple example of a constant gain algorithm employed to track an unknown time-varying system.<sup>7</sup> Our results explore the implications of adopting this kind of algorithm in underparameterized environments. A key insight of rational expectations and the Lucas critique, that the expectations parameters depend on the parameters of the underlying stochastic process, carries over to underparameterized models. However underparameterization can also result in regions of the parameter space within which the Lucas critique does not locally apply. Exploring the generality of these results, in more elaborate frameworks, would be of considerable interest and is left to future research.

---

<sup>7</sup>See for example the discussion in Ljung and Soderstrom (1983).

## APPENDIX

*Proof of Lemma 1.* Let the regime switching process be rewritten as

$$\mu_t = (1 - \delta_t)\mu_{t-1} + \delta_t\nu_t \quad (19)$$

where  $\delta_t$  is an *i.i.d.* exogenous Bernoulli random variable taking value 1 with probability  $\varepsilon$  and 0 with probability  $1 - \varepsilon$ . Since  $E\mu_t = E\nu_t = \bar{\nu}$ , it follows that  $\mu_t - E\mu_t = (1 - \delta_t)(\mu_{t-1} - E\mu_t) + \delta_t(\nu_t - \bar{\nu})$ , so that without loss of generality we can temporarily set  $E\mu_t = E\nu_t = 0$  in order to compute the variance of  $\mu_t$ . Using  $E\delta_t = E\delta_t^2 = \varepsilon$  and  $E(1 - \delta_t) = E(1 - \delta_t)^2 = 1 - \varepsilon$ , we have

$$\begin{aligned} \text{Var}(\mu_t) &= E\mu_t^2 \\ &= EE[(1 - \delta_t)^2\mu_{t-1}^2 + \delta_t^2\nu_t^2 + 2\delta_t(1 - \delta_t)\nu_t\mu_{t-1}] \mid \mu_{t-1}] \\ &= E[(1 - \varepsilon)\mu_{t-1}^2 + \varepsilon\text{Var}(\nu)] \\ &= (1 - \varepsilon)\text{Var}(\mu_t) + \varepsilon\text{Var}(\nu). \end{aligned}$$

Hence

$$\text{Var}(\mu_t) = \text{Var}(\nu).$$

We compute  $\text{Cov}(\mu_t, \beta_{t-1})$  and  $\text{Var}(\beta_t)$  using an extension of the Yule-Walker equations to allow for the regime switching exogenous process. For use in the proof of Theorem 2, below, we also allow for the case with feedback in which  $\theta \neq 1$ . Combining (4), (7) and  $\pi_t^e = \beta_{t-1}$  gives

$$\beta_t = (1 - \gamma\theta)\beta_{t-1} + \gamma\theta(\mu_t + \eta_t), \quad (20)$$

where  $\theta = 1/(1 + \alpha)$ . In the zero feedback case  $\theta = 1$ . Again, note that  $E\beta_t = \bar{\nu}$  and  $(\beta_t - \bar{\nu}) = (1 - \gamma)(\beta_{t-1} - \bar{\nu}) + \gamma((\mu_t - \bar{\nu}) + \eta_t)$ , so that without loss of generality we can temporarily set  $\bar{\nu} = 0$  to compute variances and covariances. Multiplying the above equation successively by  $\beta_t, \beta_{t-1}, \mu_t$  and  $\eta_t$ , taking unconditional expectations and imposing stationarity we obtain

$$\begin{aligned} E\beta_t^2 &= (1 - \gamma\theta)E\beta_t\beta_{t-1} + \gamma\theta E\mu_t\beta_t + \gamma\theta E\eta_t\beta_t, \\ E\beta_t\beta_{t-1} &= (1 - \gamma\theta)E\beta_t^2 + \gamma\theta E\mu_t\beta_{t-1}, \\ E\mu_t\beta_t &= (1 - \gamma\theta)E\mu_t\beta_{t-1} + \gamma\theta E\mu_t^2, \\ E\eta_t\beta_t &= \gamma\theta E\eta_t^2. \end{aligned}$$

Furthermore, multiplying the regime switching equation (19) by  $\beta_{t-1}$  and taking expectations we also have

$$E\mu_t\beta_{t-1} = (1 - \varepsilon)E\mu_t\beta_t.$$

Setting  $\theta = 1$  and solving these five equations simultaneously, we obtain (12) and (13), where  $E\beta_t^2$  and  $E\mu_t\beta_{t-1}$  are replaced by  $Var(\beta_t)$  and  $Cov(\mu_t, \beta_{t-1})$ , respectively, to allow for nonzero means, and  $Var(\mu_t) = Var(\nu)$  is invoked. *Q.E.D.*

*Proof of Lemma 2 and Theorem 1.* Let us establish that the second-order condition for a MSE minimum is satisfied globally. Clearly, the first term in (15) is strictly increasing in  $\gamma$ . As for the second term, note that

$$\begin{aligned} \frac{\partial}{\partial \gamma} \Psi(\gamma) &= -\frac{2}{(2 - \gamma)^3(\varepsilon + \gamma(1 - \varepsilon))^3} \\ &\quad \times [2\varepsilon(1 - \varepsilon) + (2 - 3\varepsilon)^2 - 3\gamma(1 - \varepsilon)(2 - 3\varepsilon) + 6\gamma^2(1 - \varepsilon)^2]. \end{aligned}$$

The term in brackets is strictly positive, as can be established by minimizing the term with respect to  $\gamma$ . Thus  $-\varepsilon\Psi(\gamma)$  is increasing in  $\gamma$  as well and (15) determines the global minimum. Since the left-hand-side of (15) is positive at  $\gamma = 1$  and since (18) is derived from (15), we have  $\gamma^* \in (0, 1)$  if and only if (17) holds, while otherwise the constraint  $\gamma \geq 0$  becomes binding. *Q.E.D.*

*Proof of Theorem 2.* Recall that  $\theta = (1 + \alpha)^{-1}$ . We have

$$\begin{aligned} MSE &= E((1 - \theta)\beta_{t-1} + \theta(\mu_t + \eta_t) - \hat{\beta}_{t-1})^2 \\ &= Var((1 - \theta)\beta_{t-1} + \theta(\mu_t + \eta_t)) + \\ &\quad Var(\hat{\beta}_{t-1}) - 2(1 - \theta)Cov(\beta_{t-1}, \hat{\beta}_{t-1}) - 2\theta Cov(\mu_t, \hat{\beta}_{t-1}). \end{aligned}$$

We first show that  $MSE$  is a smooth function of  $\gamma$  and  $\hat{\gamma}$  for all  $\gamma, \hat{\gamma} \in [0, 1]$  and for all  $\theta$  in some neighborhood of  $\theta = 1$ . From (20),(19) and  $\hat{\beta}_t = \hat{\gamma}\pi_t + (1 - \hat{\gamma})\hat{\beta}_{t-1}$  it is straightforward to obtain the following equations for second moments (without loss of generality we set means equal to 0):

$$\begin{aligned} E\hat{\beta}_t^2 &= (1 - \hat{\gamma})E\hat{\beta}_t\hat{\beta}_{t-1} + \hat{\gamma}(1 - \theta)E\beta_{t-1}\hat{\beta}_t + \hat{\gamma}\theta E\mu_t\hat{\beta}_t + \hat{\gamma}\theta E\eta_t\hat{\beta}_t \\ E\hat{\beta}_t\hat{\beta}_{t-1} &= (1 - \hat{\gamma})E\hat{\beta}_t^2 + \hat{\gamma}(1 - \theta)E\beta_t\hat{\beta}_t + \hat{\gamma}\theta E\mu_t\hat{\beta}_{t-1} \\ E\mu_t\hat{\beta}_t &= (1 - \hat{\gamma})E\mu_t\hat{\beta}_{t-1} + \hat{\gamma}(1 - \theta)E\mu_t\beta_{t-1} + \hat{\gamma}\theta E\mu_t^2 \end{aligned}$$



$$\begin{aligned}
E\eta_t\hat{\beta}_t &= \hat{\gamma}\theta E\eta_t^2 \\
E\mu_t\hat{\beta}_{t-1} &= (1-\varepsilon)E\mu_t\hat{\beta}_t \\
E\beta_{t-1}\hat{\beta}_t &= (1-\hat{\gamma})E\beta_t\hat{\beta}_t + \hat{\gamma}(1-\theta)E\beta_t^2 + \hat{\gamma}\theta E\mu_t\beta_{t-1} \\
E\beta_t\hat{\beta}_t &= (1-\gamma\theta)E\beta_{t-1}\hat{\beta}_t + \gamma\theta E\mu_t\hat{\beta}_t + \gamma\theta E\eta_t\hat{\beta}_t.
\end{aligned}$$

In addition we have the five equations in  $E\beta_t^2$ ,  $E\beta_t\beta_{t-1}$ ,  $E\mu_t\beta_t$ ,  $E\eta_t\beta_t$  and  $E\mu_t\beta_{t-1}$  derived in the Proof of Theorem 1. This gives us 12 linear equations in 12 second moments. Thus  $E\mu_t\hat{\beta}_{t-1}$ ,  $E\beta_t\hat{\beta}_t$ , and  $E\hat{\beta}_t^2$  can be written as the first three elements of a 12x1 moment vector  $M$  which are jointly determined by the equation  $R(\theta, \gamma, \hat{\gamma})M = s(\gamma, \hat{\gamma})$ , where  $R$  is 12x12 and  $s$  is 12x1. Furthermore,  $R$  and  $s$  are smooth functions and it can be verified that  $R$  is invertible in a neighborhood of  $\theta = 1$  for all  $\gamma, \hat{\gamma} \in [0, 1]$  so that  $M = M(\theta, \gamma, \hat{\gamma})$  is well defined and smooth in a neighborhood of  $\theta = 1$  for all  $\gamma, \hat{\gamma} \in [0, 1]$ . The  $MSE$  is given by  $MSE = u(\theta, \gamma) + w(\theta)'M(\theta, \gamma, \hat{\gamma}) \equiv P(\theta, \gamma, \hat{\gamma})$ , where  $w(\theta)' = (-2\theta, -2(1-\theta), 1, 0, \dots, 0)$ . It can be seen that  $P(1, \gamma, \hat{\gamma})$  is independent of  $\gamma$  and is given by the expression for  $MSE$  given in Section 3.

Clearly  $P$  is smooth for all  $\gamma, \hat{\gamma} \in [0, 1]$  and for all  $\theta$  in some neighborhood of  $\theta = 1$  and from Section 3  $P_{33}(1, \gamma, \hat{\gamma}) > 0$  for all  $\gamma, \hat{\gamma} \in [0, 1]$ . Let  $H(\theta, \gamma) \equiv P_3(\theta, \gamma, \gamma)$ . Then  $\gamma^*$  is a Nash equilibrium if  $H(\theta, \gamma^*) = 0$  and  $P_{33}(\theta, \gamma^*, \hat{\gamma}) > 0$  for all  $\hat{\gamma} \in [0, 1]$ . Let  $\bar{\gamma}$  denote the optimal value of  $\gamma$  when  $\theta = 1$ . Since  $H(1, \bar{\gamma}) = 0$  it follows from the implicit function theorem that there exists  $\gamma^*(\theta)$  near  $\bar{\gamma}$  satisfying  $H(\theta, \gamma^*(\theta)) = 0$  provided  $H_\gamma(1, \bar{\gamma}) \neq 0$ . This condition is satisfied since  $H_\gamma(\theta, \gamma) = P_{32}(\theta, \gamma, \gamma) + P_{33}(\theta, \gamma, \gamma)$ , since  $P(1, \gamma, \bar{\gamma})$  independent of  $\gamma$  implies  $P_{32}(1, \gamma, \gamma) = 0$  for all  $\gamma$ , and since  $P_{33}(1, \bar{\gamma}, \bar{\gamma}) > 0$ . It remains to show that  $P_{33}(\theta, \gamma^*, \hat{\gamma}) > 0$  for all  $\hat{\gamma} \in [0, 1]$ . Let  $L = \min P_{33}(1, \gamma, \hat{\gamma}) > 0$ . Since  $P_{33}$  is continuous for  $\gamma, \hat{\gamma} \in [0, 1]$  and for  $\theta$  sufficiently close to 1, it follows that, for some neighborhood of  $\theta = 1$ ,  $P_{33}(\theta, \gamma^*, \hat{\gamma}) > L/2 > 0$  for all  $\hat{\gamma} \in [0, 1]$ .

The proof is completed by noting that continuity of  $P(\theta, \gamma, \hat{\gamma})$  implies that if (17) holds, then for  $\theta$  sufficiently close to 1,  $MSE$  must be lower for the individual agent in the Nash equilibrium than for the choice  $\hat{\gamma} = 0$ . *Q.E.D.*

## References

- [1] Cagan P. (1956), "The monetary dynamics of inflation." In *Studies in the Quantity Theory of Money*, ed. M. Friedman, University of Chicago

Press, Chicago.

- [2] Evans, G. and S. Honkapohja (1993), "Adaptive forecasts, hysteresis and endogenous fluctuations," Federal Reserve Bank of San Francisco *Economic Review*, 1993(1), 3-13.
- [3] Evans, G. and S. Honkapohja (2001), *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton.
- [4] Fischer, S. (1977), "Long-term contracts, rational expectations and the optimal money supply rule," *Journal of Political Economy*, 85, 191-205.
- [5] Friedman, M. (1957), *Theory of the Consumption Function*. Princeton University Press, Princeton.
- [6] Ljung, L. (1977), "Analysis of Recursive Stochastic Algorithms," *IEEE Transactions on Automatic Control*, 22, 551-575.
- [7] Ljung, L. and T. Söderström (1983), *Theory and Practice of Recursive Identification*. MIT Press, Cambridge, MA.
- [8] Lucas, Jr., R.E. (1973), "Some international evidence on output-inflation trade-offs," *American Economic Review*, 63, 326-334.
- [9] Lucas, Jr., R.E. (1976), "Econometric policy evaluation: a critique." In *The Phillips Curve and the Labor Market*, eds. K. Brunner and A. Meltzer, Vol. 1 of the Carnegie-Rochester Conferences in Public Policy, a supplemental series to the *Journal of Monetary Economics*. Amsterdam: North-Holland.
- [10] Muth, J.F. (1961), "Optimal properties of exponentially weighted forecasts," *Journal of the American Statistical Association*, 55, 299-306.
- [11] Muth, J.F. (1961), "Rational expectations and the theory of price movements," *Econometrica*, 29, 315-335.
- [12] Sargent, T.J. (1999), *The Conquest of Inflation*. Princeton University Press, Princeton.
- [13] Sargent, T.J. and N. Wallace (1975), "'Rational expectations,' the optimal monetary instrument and the optimal money supply rule," *Journal of Political Economy*, 83, 241-254.