Precautionary Learning and Inflationary Biases

Dave, Chetan and Feigenbaum, James

University of Texas at Dallas

21 October 2007

Precautionary Learning and Inflationary Biases*

Chetan Dave† James Feigenbaum‡

University of Texas at Dallas University of Pittsburgh

October 21, 2007

Abstract

Recursive least squares learning is a central concept employed in selecting amongst competing outcomes of dynamic stochastic economic models. In employing least squares estimators, such learning relies on the assumption of a symmetric loss function defined over estimation errors. Within a statistical decision making context, this loss function can be understood as a second order approximation to a von-Neumann Morgenstern utility function. This paper considers instead the implications for adaptive learning of a third order approximation. The resulting asymmetry leads the estimator to put more weight on avoiding mistakes in one direction as opposed to the other. As a precaution against making a more costly mistake, a statistician biases his estimates in the less costly direction by an amount proportional to the variance of the estimate. We investigate how this precautionary bias will affect learning dynamics in a model of inflationary biases. In particular we find that it is possible to maintain a lower long run inflation rate than could be obtained in a time consistent rational expectations equilibrium.

*We would like to thank seminar participants at the 2005 Southern Economics Association meetings and at the Federal Reserve Bank of Dallas for comments. All remaining errors are ours.

†Address: The University of Texas at Dallas; School of Economic, Political and Policy Sciences, GR 3.104; 800 W. Campbell Rd.; Richardson, TX., 75080. Email: cdave@utdallas.edu.

‡Address: The University of Pittsburgh; Department of Economics; 4900 W. W. Posvar Hall; Pittsburgh, PA., 15260. Email: jfeigen@pitt.edu.

1 Introduction

Dynamic stochastic macroeconomic models can produce multiple outcomes depending on the equilibrium concept employed. For example, following Kydland and Prescott (1977) and Barro and Gordon (1983), models in which a central bank sets monetary policy over time when facing a public with rational expectations can deliver one of two outcomes. The Nash equilibrium concept delivers a time-consistent high inflation outcome, while the Ramsey equilibrium concept delivers a time-inconsistent low inflation outcome. Assuming that the central bank learns the latent parameters of the structural Phillips curve by estimating least squares regressions over the entire time series of data on a 'perceived' (and possibly misspecified) Phillips curve in order to set monetary policy, it is possible to state a stability condition whereby the Nash outcome is selected as the one that is 'learnable', and therefore expected to arise in reality.[1] On the other hand, if more weight is given to more recent data, Sargent (1999) and Cho et al (2002) have shown that an economy may occasionally 'escape' for brief periods to the Ramsey outcome.

In this paper we ask whether information processing, modeled via statistical decision making, can enhance the learning dynamics obtained in the constant gain case. In particular, we model the information processing decision separately from the optimal policy-making decision. This dichotomy allows us to incorporate alternate assumptions on information processing. Consequently, any contributions to the dynamics of learning that alternative assumptions provide can be explicitly identified. In order to motivate this approach it is instructive to view the adaptive learning process through a statistical decision making lens, as follows.

Adaptive learning begins from the assumption of least squares estimation. Least squares estimation assumes, following Zellner (1971) and Berger (1985), that a statistical decision maker minimizes a squared error loss function, faces an underlying data generating process that is Gaussian in nature and employs Bayes' rule to update parameter estimates as new

[1] See Evans and Honkapohja (1999, 2001).

data arrives. Econometricians typically assume that the loss function is a primitive of this statistical decision problem. However, as Berger (1985) demonstrates, the squared error loss function can be derived from a second-order Taylor series approximation to a von Neumann-Morgenstern utility function defined over 'rewards' to a statistical decision maker (SDM). Within the context of a central bank that learns about the economy, a reward could be in the form of a reputation for interpreting varied socioeconomic data 'correctly'. Several recent articles have examined the effects of allowing a central bank to have an asymmetric loss function for policy-making (e.g. Ruge-Murcia (2003)). However, following Cukierman (2002) who suggests examining sources of the degree of transparency of a central banks models and objectives, we examine here the effects of asymmetry in how a central bank might learn about underlying hazy fundamentals, while keeping the policy-making objective function symmetric.

Suppose that the SDM views overestimation of the unknown parameters as more costly than underestimation. A third-order approximation to the underlying utility function is necessary for this assumption to affect the SDM's actions. This paper first demonstrates that such a SDM would have an asymmetric loss function, derived by taking a third-order approximation to the underlying utility function. The asymmetry will then cause the SDM to behave in a 'precautionary' manner. Indeed, the mathematics underlying this interpretation of the asymmetry is essentially the same as in Leland's (1968) study of precautionary saving.

Given an asymmetric loss function, this paper specifies the recursive form of the optimal estimator chosen by a precautionary SDM. Finally, within the context of the Kydland and Prescott (1977)-Barro and Gordon (1983) model, this paper discusses the dynamics of such 'precautionary learning' in terms of the stability of the Nash and Ramsey outcomes. Cho et al (2002) have shown escape dynamics induced by the assumption of a constant gain in least squares learning suggest that an economy may occasionally and briefly deviate to an outcome that has not been defined to be stable. The simulation results in this paper suggest that an economy may possess more complicated dynamics that *enhance* the likelihood of such escape. In particular, we find that the frequency of escape can be higher, or the economy can fluctuate around a lower level of inflation than the Nash outcome, depending upon the parameterized degree of asymmetry. For some parameter values, the economy can even settle around the Ramsey outcome.

The paper is structured as follows. Section 2 briefly discusses existing statistical decision theory, shows that a precautionary SDM employs an asymmetric loss function and provides the recursive form of the resulting alternative estimator. In Section 3 a form of the Kydland and Prescott (1977)-Barro and Gordon (1983) environment is specified. Section 4 provides simulation results within this context and discusses which outcome may be selected under the assumption of precautionary learning. Section 5 concludes with our finding and its' underlying intuition.

2 Precautionary Motives in Statistical Decisions

2.1 The Standard Case

Following Berger (1985), a SDM faces the following problem when deciding upon the loss function upon which to base estimation. Given a set of states of nature or parameters $\widetilde{\theta} \in \Theta$, the decision maker takes an action $\widetilde{a} \in \widetilde{A}$ in order to maximize utility (g) which is a function of the estimation error $(\widetilde{\theta} - \widetilde{a})$. Assuming that g is strictly concave and thrice continuously differentiable yields the following Taylor series approximation of $g(\widetilde{\theta}-\widetilde{a})$ around 0:

$$g(\widetilde{\theta} - \widetilde{a}) \simeq g(0) + (\widetilde{\theta} - \widetilde{a})g_1(0) + \frac{(\widetilde{\theta} - \widetilde{a})^2 g_{11}(0)}{2} \tag{1}$$

where g_1 denotes the first derivative and so on. Next, define the following expectations:

$$K_0 \equiv -E[g(0)], \quad K_1 \equiv -E[g_1(0)], \quad K_2 \equiv -E\left[\frac{g_{11}(0)}{2}\right] > 0. \tag{2}$$

Given this environment, Berger (1985) defines the loss function for estimation as,

$$L(\widetilde{\theta}, \widetilde{a}) \simeq -E[g(\widetilde{\theta} - \widetilde{a})], \tag{3}$$

which, given re-definition of the action space as $\widetilde{A}^* = \{\widetilde{a} - c, \quad c = \frac{K_1}{2K_2} | \, \widetilde{a} \in \widetilde{A}\}$ results in

$$L(\widetilde{\theta}, \widetilde{a}) = (\widetilde{\theta} - \widetilde{a})^2. \tag{4}$$

Given this symmetric loss function, optimality of the least squares estimator is assured.

2.2 The Precautionary Case

If the SDM has a precautionary motive then a third-order approximation of the utility function is required to account for this motive. We assume that the SDM seeks to maximize utility $E[g(\lambda(\widetilde{a} - \widetilde{\theta})]$ where λ is a scale parameter that measures how sensitive the decision maker is to deviations of $\widetilde{a}$ from $\widetilde{\theta}$. Define the loss function to be

$$L(\widetilde{\theta}, \widetilde{a}) = -E[g(\lambda(\widetilde{a} - \widetilde{\theta}))]. \tag{5}$$

Under the assumption that g is differentiable to the fourth degree with respect to $\widetilde{a} - \widetilde{\theta}$, a Taylor series approximation is

$$\begin{aligned} L(\widetilde{\theta}, \widetilde{a}) \;\; \approx & \;\; -E[g(0)] - \lambda(\widetilde{a} - \widetilde{\theta})E[g_1(0)] \\ & -\frac{1}{2}\lambda^2(\widetilde{a} - \widetilde{\theta})^2 E[g_{11}(0)] - \frac{1}{6}\lambda^3(\widetilde{a} - \widetilde{\theta})^3 E[g_{111}(0)]. \end{aligned} \tag{6}$$

We wish the optimal objective for the statistical decision-maker to be to choose $\widetilde{a} = \widetilde{\theta}$, which implies the conditions $E[g_1(0)] = 0$ and $E[g_{11}(0)] < 0$ so that $\widetilde{a} = \widetilde{\theta}$ is the local minimum of the loss function. It can further be assumed that $E[g(0)] = 0$ since a constant can be added to the utility function without changing the underlying structure. Next, define

$$K_2 = -E[g_{11}(0)] \tag{7}$$

$$K_3 = -E[g_{111}(0)], \tag{8}$$

then the loss function simplifies to

$$L(\widetilde{\theta}, \widetilde{a}) \approx \frac{1}{2} K_2 \lambda^2 (\widetilde{a} - \widetilde{\theta})^2 + \frac{1}{6} K_3 \lambda^3 (\widetilde{a} - \widetilde{\theta})^3. \tag{9}$$

Further, assuming as in the standard case, that the statistical decision-maker chooses $\widetilde{a}$ to minimize $E_{\widetilde{\theta}}[L(\widetilde{\theta}, \widetilde{a})]$ (given beliefs about the distribution of $\widetilde{\theta}$), the first order condition reduces to

$$K_2(\widetilde{a} - \mu_{\widetilde{\theta}}) + \frac{\lambda K_3}{2}[(\widetilde{a} - \mu_{\widetilde{\theta}})^2 + \sigma^2_{\widetilde{\theta}}] = 0 \tag{10}$$

where

$$\mu_{\widetilde{\theta}} = E_{\widetilde{\theta}}[\widetilde{\theta}] \tag{11}$$

$$\sigma^2_{\widetilde{\theta}} = E_{\widetilde{\theta}}[(\widetilde{\theta} - \mu_{\widetilde{\theta}})^2]. \tag{12}$$

Assuming that the optimal $\widetilde{a}$ follows

$$\widetilde{a} = \widetilde{a}_0 + \lambda \widetilde{a}_1 + \lambda^2 \widetilde{a}_2 + ..., \tag{13}$$

then to the zero-th order in λ

$$\widetilde{a}_0 = \mu_{\widetilde{\theta}} \tag{14}$$

and to first order in λ

$$K_2(\widetilde{a}_0 + \lambda \widetilde{a}_1 - \mu_{\widetilde{\theta}}) + \frac{\lambda K_3}{2}[(\lambda \widetilde{a}_1)^2 + \sigma^2_{\widetilde{\theta}}] = 0. \tag{15}$$

Dropping the second and higher order terms in λ one can compute

$$\widetilde{a} \approx \mu_{\widetilde{\theta}} - \frac{\lambda}{2}\frac{K_3}{K_2}\sigma^2_{\widetilde{\theta}}. \tag{16}$$

In particular, with respect to estimation problems, we assume that the loss function takes the following LINEX form,

$$L(\widetilde{a}, \widetilde{\theta}) = \exp(b(\widetilde{\theta} - \widetilde{a})) - b(\widetilde{\theta} - \widetilde{a}) - 1 \approx \frac{b^2}{2}(\widetilde{\theta} - \widetilde{a})^2 + \frac{b^3}{6}(\widetilde{\theta} - \widetilde{a})^3 \tag{17}$$

Thus,

$$K_2\lambda^2 = b^2 \tag{18}$$

$$K_3\lambda^3 = -b^3 \tag{19}$$

and,

$$\widetilde{a} = \mu_{\widetilde{\theta}} + \frac{b}{2}\sigma^2_{\widetilde{\theta}}. \tag{20}$$

This discussion is applied to the adaptive least squares learning process described below.

3 The Economic Environment

Cho et al (2002) describe a model in which the monetary authority uses least squares learning to determine its policy. The monetary authority's beliefs are described by a vector of regression coefficients $\widetilde{\gamma}$. It chooses a decision rule $h(\widetilde{\gamma})$ that causes the stochastic process for the economy to be $\widetilde{\xi}(\widetilde{\gamma})$. Given $\widetilde{\xi}(\widetilde{\gamma})$ the best fitting regression will be $\Gamma = T(\widetilde{\gamma})$. A self-confirming equilibrium is a fixed point of T.

Here we consider what happens if we separate the monetary authority into two entities: a statistician and a policy maker. The statistician's beliefs are described by a vector of regression coefficients $\widetilde{\gamma}$. Based on those regression coefficients, the statistician estimates

$\widehat{\phi}(\widetilde{\gamma})$ where ϕ is the object of interest to the policy maker, which in turn summarizes the beliefs of the policy maker about the economy. The policy maker then follows a decision rule $H(\phi)$ that causes the stochastic process for the economy to be $\widetilde{\xi}(H)$. Finally, given $\widetilde{\xi}$, the best fitting regression by the statistician in the next period will be $\Gamma(\widetilde{\xi})$. If we define $T(\widetilde{\gamma}) = \Gamma(\widetilde{\xi}(H(\widehat{\phi}(\widetilde{\gamma}))))$ then a self confirming equilibrium is a fixed point of T.

Suppose that the statisticians goal is to minimize a loss function $L(\widehat{\phi} - \phi)$ and a is a scalar that measures the degree of precaution in reporting estimates to the policy maker, where $a = 0$ implies no such precautionary motive. Then the estimated function of interest is $\widehat{\phi}(\gamma|a)$ and will be a function of a. In the event that $a = 0$ the statistician and policy maker coincide and there is no dichotomy.

3.1 The Basic Model

As an example of how the dichotomy between a policy maker and a loss averse statistician might interact, we consider the following model of inflation and unemployment adapted from Kydland and Prescott (1977) and Cho et al (2002). Let U_t represent unemployment and π_t represent inflation. The monetary authority chooses its target inflation rate x_t to minimize the loss function

$$L_p(x_t) = E[\pi_t(x_t)^2 + \alpha U_t(x_t)^2], \tag{21}$$

where U_t is governed by

$$U_t = u - \theta(\pi_t - \widehat{x}_t) + \sigma_1 W_{1t} \tag{22}$$

and

$$\pi_t = x_t + \sigma_2 W_{2t}. \tag{23}$$

where $(\sigma_1, \sigma_2) > 0$, and $(W_{1t}, W_{2t})^T$ are *i.i.d.* normally distributed shocks with zero means and identity covariance matrices. In the Phillips curve above, $\widehat{x}_t$ is the private sector's expectation of inflation at t, $u > 0$ is the natural rate of unemployment and $\theta > 0$ is the slope.

Let us suppose that the monetary authority takes $\widehat{x}_t$ as given. Then, given the above loss minimization problem,

$$x_t = \frac{\alpha\theta}{1+\alpha\theta^2}\left(u + \theta\widehat{x}_t\right). \tag{24}$$

is the optimal inflation target. Under rational expectations, $x_t = \widehat{x}_t$, so the optimal policy is $x_t^* = \alpha\theta u$. This corresponds to the time-consistent Nash equilibrium of Kydland and Prescott (1977). Thus, in equilibrium, $E[U_t] = u$ and $E[\pi_t] = \alpha\theta u$ and

$$\begin{aligned} L_p(x_t^*) &= \alpha^2\theta^2u^2 + \sigma_2^2 + \alpha u^2 + \alpha\left[\theta^2\sigma_2^2 + \sigma_1^2\right] \\ &= (1+\alpha\theta^2)(\alpha u^2 + \sigma_2^2) + \alpha\sigma_1^2. \end{aligned} \tag{25}$$

On the other hand, if the monetary authority did not try to exploit the private sector's expectations and just assumed the private sector would know what it was doing, then the enlightened policy maker's loss function is

$$L_p^e(x_t) = x_t^2 + \sigma_2^2 + \alpha u^2 + \alpha\left[\theta^2\sigma_2^2 + \sigma_1^2\right], \tag{26}$$

where $\widehat{x}_t$ is set equal to x_t. Then clearly, the optimal policy is $x_t = 0$, so $E[U_t] = u$ and $E[\pi_t] = 0$. Then

$$\begin{aligned} L_p^e(0) &= \sigma_2^2 + \alpha u^2 + \alpha\left[\theta^2\sigma_2^2 + \sigma_1^2\right] \\ &= \alpha u^2 + (1+\alpha\theta^2)\sigma_2^2 + \alpha\sigma_1^2 < L_p(x_t^*). \end{aligned} \tag{27}$$

This corresponds to the time-inconsistent Ramsey policy of Kydland and Prescott (1977). This is the optimal policy, but it is not supportable over time because the monetary authority has an incentive to inflate to get a lower unemployment rate, so there is an inflationary bias.

3.2 Least Squares Learning

Adaptive least squares learning begins by first assuming that the monetary authority has a perceived law of motion (PLM)

$$U_t = u_0 - \omega\pi_t + \eta_t, \tag{28}$$

while the actual law of motion (ALM) is given by

$$U_t = u - \theta(\pi_t - x_t) + \sigma_1 W_{1t}, \tag{29}$$

where we have assumed rational expectations on the part of the public and, as before,

$$\pi_t = x_t + \sigma_2 W_{2t}. \tag{30}$$

Suppose now that the policy maker determines policy by minimizing the loss function

$$L_p(x_t) = E_p[\pi_t(x_t)^2 + \alpha U_t(x_t)^2], \tag{31}$$

where the expectation is determined by the PLM. Then

$$\begin{aligned} L_p(x_t) &= E_p\left[(x_t + \sigma_2 W_{2t})^2 + \alpha\left(u_0 - \omega(x_t + \sigma_2 W_{2t}) + \eta_t\right)^2\right] \\ &= x_t^2 + \alpha(u_0 - \omega x_t)^2 + E_p[\sigma_2^2 W_{2t}^2 + \alpha(\eta_t - \omega\sigma_2 W_{2t})^2], \end{aligned} \tag{32}$$

which has the first-order condition

$$2x_t - 2\omega\alpha(u_0 - \omega x_t) = 0. \tag{33}$$

Thus, the monetary authority will follow the policy

$$x_t^{LS} = \frac{\alpha\omega u_0}{1+\alpha\omega^2}. \tag{34}$$

Plugging the monetary authority's policy into the ALM, we get the system of equations

$$U_t = u - \theta\left(\pi_t - \frac{\alpha\omega u_0}{1+\alpha\omega^2}\right) + \sigma_1 W_{1t} \tag{35}$$

$$\pi_t = \frac{\alpha\omega u_0}{1+\alpha\omega^2} + \sigma_2 W_{2t} \tag{36}$$

In a self-confirming equilibrium, the policy will be chosen so the ALM (35) and the PLM (28) are the same. Thus,

$$u_0 = u + \frac{\alpha\theta\omega u_0}{1+\alpha\omega^2} \tag{37}$$

$$\omega = \theta. \tag{38}$$

This system has the solution

$$u_0 = u + \frac{\alpha\theta^2}{1+\alpha\theta^2} u_0 \tag{39}$$

$$u = \frac{u_0}{1+\alpha\theta^2} \tag{40}$$

$$u_0 = (1+\alpha\theta^2)u. \tag{41}$$

Note that in the self-confirming equilibrium,

$$x^{LS} = \alpha\theta u, \tag{42}$$

which is the same high-inflation but time-consistent policy pursued by the fully rational monetary authority.

The idea is that if the ALM can be learned via least squares, the monetary authority will converge to a high-inflation policy. The question of this paper is whether introducing precautionary motives into the learning process can affect the dynamics of reaching the high-inflation policy.

3.3 Precautionary Learning

Suppose the policy maker is assumed to myopically set policy so as to minimize the loss function specified above and the private sector has rational expectations so $\widehat{x}_t = x_t$. As before, the monetary authority models the economy with a linear Phillips curve such that U_t is a linear projection on π_t and the information available at t.

Let us assume that the statistician estimates the PLM

$$U_t = u_0 - \omega \pi_t + \eta_t, \tag{43}$$

where ω is a scalar, u_0 is a constant, and η_t is a mean-zero, serially uncorrelated noise term. Let us suppose the statistician is particularly interested in knowing ω and also suppose that the statistician is more concerned about overestimating ω rather than underestimating it. As a result, the statistician chooses its estimate $\widehat{\omega}$ so as to minimize the LINEX loss function

$$L_s^t(\widehat{\omega}_t) = E_t\left[e^{a(\omega - \widehat{\omega}_t)} - a(\omega - \widehat{\omega}_t) - 1\right]. \tag{44}$$

If the statistician believes at t that $\omega \sim N(\mu_\omega^t, (\sigma_\omega^t)^2)$, then

$$L_s^t(\widehat{\omega}_t) = \exp\left(a(\mu_\omega^t - \widehat{\omega}) + \frac{a^2(\sigma_\omega^t)^2}{2}\right) - a(\mu_\omega^t - \widehat{\omega}_t) - 1, \tag{45}$$

so

$$\frac{dL_s}{d\widehat{\omega}_t} = -a \exp\left(a(\mu_\omega^t - \widehat{\omega}_t) + \frac{a^2(\sigma_\omega^t)^2}{2}\right) + a = 0, \tag{46}$$

and the optimal estimate will satisfy

$$a(\mu_\omega^t - \widehat{\omega}_t) + \frac{a^2(\sigma_\omega^t)^2}{2} = 0 \tag{47}$$

or

$$\widehat{\omega}_t = \mu_\omega^t + \frac{a}{2}(\sigma_\omega^t)^2. \tag{48}$$

Thus, if $a > 0$, it is more costly for the statistician to underestimate ω than to overestimate ω, so he makes a higher, more conservative estimate than he would if he had a quadratic loss function (corresponding to $a = 0$). The opposite is true if $a < 0$.

The statistician will use the Kalman filter to update his estimate of ω in the perceived Phillips curve to $\widehat{\omega}_t$. Then the policy maker will choose x_t so as to minimize

$$L_p(x_t) = E_p[(x_t + \sigma_2 W_{2t})^2 + \alpha\left(-\widehat{\omega}_t\left(x_t + \sigma_2 W_{2t}\right) + u_0 + \eta_t\right)^2]. \tag{49}$$

This simplifies to

$$L_p(x_t) = x_t^2 + \alpha\left[(-\widehat{\omega}_t x_t + u_0)^2 + E_p\left[\sigma_2 W_{2t}^2 + (-\widehat{\omega}_t \sigma_2 W_{2t} + \eta_t)^2\right]\right] \tag{50}$$

since the policy maker believes that both W_{2t} and η_t are mean zero. The variance-covariance matrix of W_{2t} and η_t does not affect the policy makers decision since the last term is independent of x_t. The policy maker will choose x_t to satisfy the first-order condition

$$2x_t - 2\alpha\widehat{\omega}_t\left(-\widehat{\omega}_t x_t + u_0\right) = 0. \tag{51}$$

Thus, the optimal choice will be

$$x_t^* = \frac{\alpha\widehat{\omega}_t}{1 + \alpha\widehat{\omega}_t^2} u_0. \tag{52}$$

Given the statistician's underlying precautionary motives and his beliefs, this is

$$x_t^* = \frac{\alpha\left[\mu_\omega^t + \frac{a}{2}(\sigma_\omega^t)^2\right]}{1 + \alpha\left[\mu_\omega^t + \frac{a}{2}(\sigma_\omega^t)^2\right]^2} u_0. \tag{53}$$

Using Evans and Honkapohja (2001)'s Ricatti equations, which do not involve the estimate of the error variance (but also require the inversion of a matrix), we have

$$\begin{aligned} \widehat{\xi}_{t+1|t}^{LS} &= \widehat{\xi}_{t|t-1}^{LS} + t^{-1} R_t^{-1} x_t \left(y_t - x_t^T \widehat{\xi}_{t|t-1}^{LS} \right) \quad (54) \\ R_t &= R_{t-1} + t^{-1} \left(x_t x_t^T - R_{t-1} \right), \quad (55) \end{aligned}$$

where in the present context $y_t = u_t$, $x_t = (1, \pi_t)^T$, ξ represents the recursive parameter estimate (e.g. the slope of the Phillips curve). For constant gain learning, we replace t^{-1} by γ^{-1}, where γ is a constant. Effectively, the learner will behave as though he is considers only a moving window of the last γ observations.

The estimate of the variance-covariance matrix will then be

$$P_{t+1|t}^{LS} = t^{-1} R_t^{-1} \widehat{\sigma}_{t+1|t}^2, \quad (56)$$

where

$$\begin{aligned} \widehat{\sigma}_{t+1|t}^2 &= \widehat{\sigma}_{t|t-1}^2 \\ &+ \frac{1}{t-r} \left\{ \left(y_t - x_t^T \widehat{\xi}_{t|t-1} \right)^2 \left[1 - t^{-1} x_t^T R_t^{-1} x_t \right] \right. \\ &\left. + 2 \frac{t-1}{t} (\widehat{\xi}_{t|t-1}^T R_{t-1} - S_{t-1}) R_t^{-1} x_t (y_t - x_t^T \widehat{\xi}_{t|t-1}) - \widehat{\sigma}_{t|t-1}^2 \right\}. \quad (57) \end{aligned}$$

and

$$S_t = \frac{1}{t} \sum_{i=1}^{t} y_i x_i^T \quad (58)$$

is a $1 \times r$ row vector that satisfies the Ricatti equation

$$S_t = S_{t-1} + t^{-1} (y_t x_t^T - S_{t-1}).$$

Generalizing, we then get the precautionary learning equation

$$(\widehat{\xi}_{t+1|t})_i = \left(\widehat{\xi}_{t+1|t}^{LS} \right)_i + \frac{a_i}{2} (P_{t+1|t}^{LS})_{ii}, \quad (59)$$

where $i = 1, \ldots, r$ and ξ consists as before of the regression parameters estimated by the SDM and provided to the policy-maker. We now provide simulation results of the model to evaluate whether the low inflation outcome is reachable.[2] In the simulation results we let γ^{-1} be the constant Kalman gain of the Ricatti equations as in Cho et al (2002) and vary that as well.

4 Simulation Results

Our first objective in simulating the model economy is to verify whether or not we obtain escape dynamics given a lack of precaution (in order to replicate the results, for instance, in Cho et al (2002)). Next, allowing for a degree of precaution (a) on the estimated slope of the Phillips curve (θ), we investigate whether the Ramsey outcome is attainable under the following three scenarios. First, holding the degree of precaution and the Kalman gain (γ) constant, we vary the value of the shocks hitting the simulated model economy (σ_1 and σ_2). Second, we vary the Kalman gain as in Cho et al (2002) and finally we vary the degree of precaution on the slope of the Phillips curve. The simulations are conducted under certain fixed parameters as indicated in Table 1 and were conducted for 10000 periods. Given the fixed parameter values, Nash inflation is 10% and Ramsey inflation is 0%. Next, Table 2 provides the values for the parameters that are varied and indicates the time series for the targeted rate of inflation (x_t) that arise given the varied model parameters. These time series are then plotted in Figures 1-4.

[2] We are not aware of any stochastic approximation techniques that would allow us to derive differential equations that would approximate the stochastic difference equations that result from the recursive form of the asymmetric least squares estimator suggested here.

Table 1: Fixed Parameters

Parameter	**Value**
α	1
u	5
θ	2

Table 2: Varied Parameters

Parameter	**Value**							
	Figure 1				Figure 2			
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
γ^{-1}	20	20	100	100	20	20	100	100
a	0	0	0	0	3	3	3	3
σ_1	1	1.5	1	1.5	1	1.5	1	1.5
σ_2	1	1.5	1	1.5	1	1.5	1	1.5
	Figure 3				Figure 4			
	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}
γ^{-1}	20	20	100	100	20	20	100	100
a	6	6	6	6	9	9	9	9
σ_1	1	1.5	1	1.5	1	1.5	1	1.5
σ_2	1	1.5	1	1.5	1	1.5	1	1.5

Figure 1 below provides the plots for targeted inflation when there is no precaution in learning. The plot demonstrates that as expected the time consistent high inflation Nash outcome is achieved. However, also as expected, escapes from Nash to Ramsey levels do occur for particular parameter configurations.

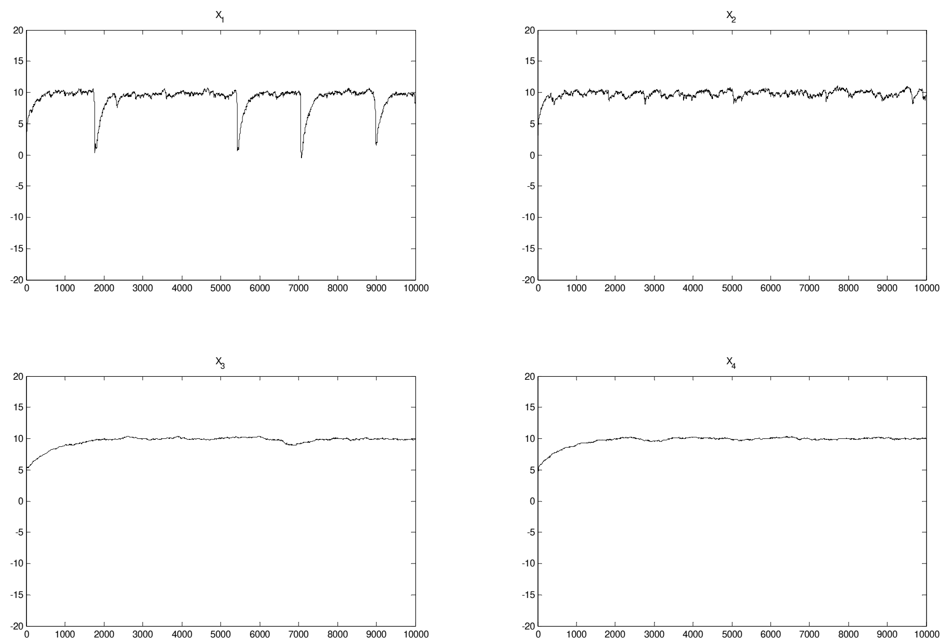

Figure 1: No Precaution

Figure 2 below provides plots for targeted inflation under a 'low' level of precaution ($a = 3$). The plots here indicate a much higher frequency of escape. Indeed for a higher gain ($\gamma^{-1} = 20$) the plots indicate quite frequent departures from the Nash outcome. When the gain is much smaller ($\gamma^{-1} = 100$) there actually seems to be overshooting behavior whence targeted inflation actually rises above the Nash outcome only to fall to the Ramsey outcome and then begin a relatively slow rise to the Nash level. That is, we see some dynamics in between the two outcomes.

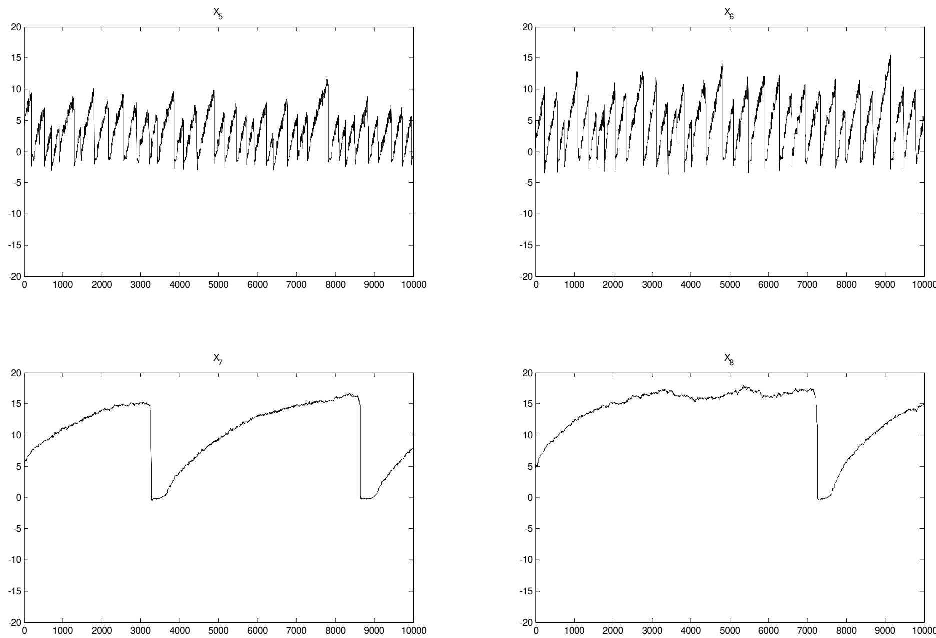

Figure 2: Low Precaution

Next, Figure 3 below provides plots for targeted inflation under a 'medium' level of precaution ($a = 6$). The plots here indicate stability at the Ramsey outcome regardless of the value for the variance of the presumed shocks to the unemployment or inflation equations provided that $\gamma^{-1} = 20$. For $\gamma^{-1} = 100$ there seems to be some overshooting above the Nash outcome level with a fall to the Ramsey outcome and a relatively slow rise to the Nash outcome. Once again, we see some dynamics at least in simulation given that the SDM employs an asymmetric least squares estimator reflecting a degree of precaution in the interpretation of data.

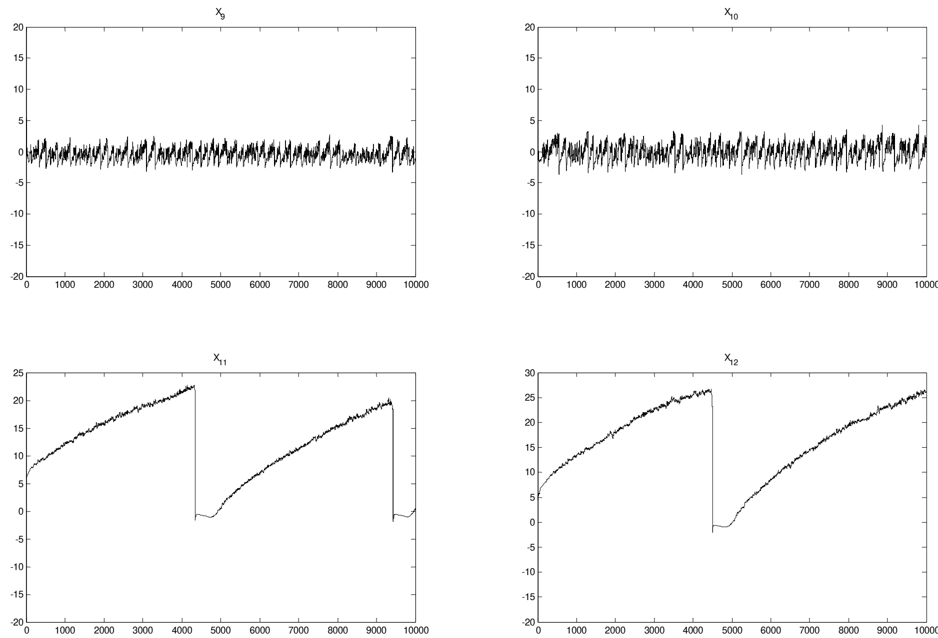

Figure 3: Medium Precaution

Finally, Figure 4 below provides plots for targeted inflation under a ‘high’ level of precaution ($a = 9$).

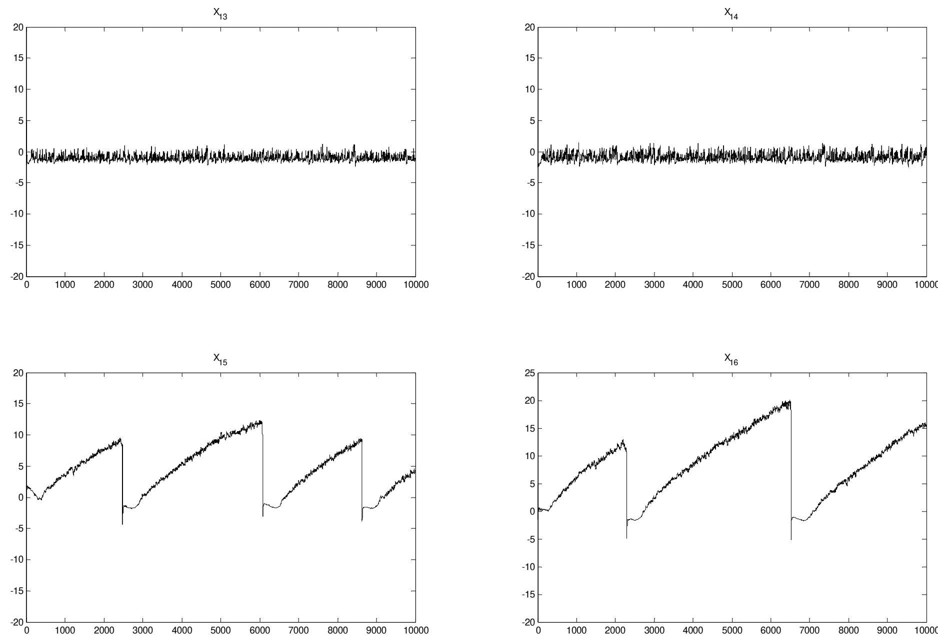

Figure 4: High Precaution

The plots indicate that for $\gamma^{-1} = 20$, regardless of the variance of the shocks hitting the unemployment and inflation equations, the targeted level of inflation is actually hovering around a negative value. Furthermore, for $\gamma^{-1} = 100$ we again seem some dynamics with only occasional overshooting.

Given these simulations for varying degrees of precaution, one might wish to ascertain what is the optimal level of precaution for the statistician to exhibit. Table 3 provides the average value of the loss function for each of the 16 simulations.

Table 3: Simulated Loss

Parameter	**Value**							
	Figure 1				Figure 2			
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
γ^{-1}	20	20	100	100	20	20	100	100
a	0	0	0	0	3	3	3	3
σ_1	1	1.5	1	1.5	1	1.5	1	1.5
σ_2	1	1.5	1	1.5	1	1.5	1	1.5
Loss	*117.40*	*134.12*	*124.32*	*133.17*	*49.54*	*73.25*	*161.32*	*235.55*
	Figure 3				Figure 4			
	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}
γ^{-1}	20	20	100	100	20	20	100	100
a	6	6	6	6	9	9	9	9
σ_1	1	1.5	1	1.5	1	1.5	1	1.5
σ_2	1	1.5	1	1.5	1	1.5	1	1.5
Loss	*31.93*	*40.25*	*220.63*	*359.19*	*32.72*	*39.77*	*65.01*	*137.89*

What is clear from Table 3 is that introduction of the precautionary parameter leads to lower values of the loss function being optimized by the policy-maker, relative to the no precaution case, when $\gamma^{-1} = 20$. For $\gamma^{-1} = 100$ this is not necessarily the case. Next, the lowest level of the loss occurs for the targeted inflation series X_9 which corresponds to the $a = 6$ 'medium' level of the precautionary parameter. As per the first panel in Figure 3 this is the case in which inflation fluctuates closely around the Ramsey outcome. In summary, it is entirely feasible that for certain values of the precautionary parameter the loss experienced by the monetary authority may be lower and the authority might still consistently target inflation around the Ramsey outcome.

5 Conclusion

The adaptive learning approach is a rich environment in which stability, perturbation, and other issues can be analyzed with respect to models with multiple possible outcomes. In this paper the approach has been to analyze learning dynamics in the event that a policy maker is split into the usual economic decision maker and into a statistical decision maker who exhibits a degree of precaution in forming estimates of the Phillips curve. In simulation we can show that we do obtain learning dynamics that are rich enough to warrant further analytical investigation. Unfortunately, we are not aware of any method by which stochastic approximation can be conducted for the constant gain case under precaution, so we limit our analysis to presenting the simulation results. These results suggest a possible tension between the pull of time-consistency and the degree to which a statistical decision maker is cautious about interpreting information (reflected in the estimator employed). We show that in simulation it is possible that the targeted rate of inflation tends to the Ramsey outcome and may stay there for some time.

References

[1] Barro, R. J. and Gordon, D. R. (1983), "A Positive Theory of Monetary Policy in a Natural Rate Model", **Journal of Political Economy**, v. 91.

[2] Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis (2nd. ed.)*, **Springer**.

[3] Cho, I-K., Sargent, T. J. and Williams, N. (2002), "Escaping Nash Inflation", **Review of Economic Studies**, v. 69.

[4] Cukierman, A. (2002), "Are Contemporary Central Banks Transparent about Economic Models and Objectives and What Difference Does it Make?", **Federal Reserve Bank of St. Louis Review**, v. 84.

[5] Evans, G. and Honkapohja, S. (1999), "Learning Dynamics", *Handbook of Macroeconomics* (v. 1A), **Elsevier**.

[6] Evans, G. and Honkapohja, S. (2001), *Learning and Expectations in Macroeconomics*, **Princeton University Press**.

[7] Kydland, F. and Prescott, E. C. (1977), "Rules Rather Than Discretion: The Inconsistency of Optimal Plans", **Journal of Political Economy**, v. 85.

[8] Leland, H. E. (1968), "Saving and Uncertainty: The Precautionary Demand for Saving", **The Quarterly Journal of Economics**, v. 82.

[9] Ruge-Murcia, F. J. (2003), "Inflation Targeting under Asymmetric Preferences", **Journal of Money, Credit and Banking**, v. 35.

[10] Sargent, T. J. (1999), *The Conquest of American Inflation*, **Princeton University Press**.

[11] Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, **Wiley**.

[12] Zellner, A. (1976), "Bayesian Estimation and Prediction Using Asymmetric Loss Functions", **Journal of the American Statistical Association**, v. 81.