

# The Stability of Macroeconomic Systems with Bayesian Learners

**James Bullard\***

Federal Reserve Bank of St. Louis

**Jacek Suda<sup>†</sup>**

Banque de France

This version: 8 October 2009

## Abstract

We study abstract macroeconomic systems in which expectations play an important role. Consistent with the recent literature on recursive learning and expectations, we replace the agents in the economy with econometricians. Unlike the recursive learning literature, however, the econometricians in the analysis here are Bayesian learners. We are interested in the extent to which expectational stability remains the key concept in the Bayesian environment. We isolate conditions under which versions of expectational stability conditions govern the stability of these systems just as in the standard case of recursive learning. We conclude that Bayesian learning schemes, while they are more sophisticated, do not alter the essential expectational stability findings in the literature.

*Keywords:* Expectational stability, recursive learning, learnability of rational expectations equilibrium.

*JEL codes:* D84, E00, D83.

---

\*Email: bullard@stls.frb.org. Any views expressed are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Banque de France.

<sup>†</sup>Email: jacek.suda@banque-france.fr.

# 1 Introduction

## 1.1 Overview

A large and expanding literature has developed over the last two decades concerning the issue of learning in macroeconomic systems. These systems have a recursive feature, whereby expectations affect states, and states feed back into the expectations formation process being used by the agents. The focus of the literature has been on whether processes in this class are locally convergent to rational expectations equilibria. Evans and Honkapohja (2001), in particular, have stressed that the expectational stability condition governs the stability of real-time learning systems defined in this way.

This line of research has so far emphasized recursive updating, including least squares learning as a special case. There has been little study of Bayesian updating in the context of expectational stability. What might one expect from an extension to Bayesian updating? There seem to be at least two lines of thought in this area. One is that Bayesian estimation is a close relative of least squares, and therefore that all expectational stability results should obtain with suitable adjustments, but without conceptual difficulties. A second, opposite view is that Bayesian agents are essentially endowed with rational expectations—indeed Bayesian learning is sometimes called “rational learning” in the literature—and therefore one should not expect to find a concept of “expectational instability” in the Bayesian case. A goal of this paper is to understand which of these views is closer to reality in abstract macroeconomic systems.

It is also important to understand how Bayesian updating might repair certain apparent inconsistencies in the recursive learning literature. Cogley and Sargent (2008), for example, have noted that there are “two minds” embedded in the anticipated utility approach to learning that has become popular. According to Cogley and Sargent (2008, p. 186),

“[The anticipated utility approach recommended by Kreps (1998)] is of two minds .... Parameters are treated as random

variables when agents learn but as constants when they formulate decisions. Looking backward, agents can see how their beliefs have evolved in the past, but looking forward they act as if future beliefs will remain unchanged forever. Agents are eager to learn at the beginning of each period, but their decisions reflect a pretence that this is the last time they will update their beliefs, a pretence that is falsified at the beginning of every subsequent period.”

In this paper, we take a first step toward studying this issue in the context of expectational stability. The Bayesian econometricians in our model will recognize that their beliefs will continue to evolve in the future. The Bayesian perspective means treating estimates as random variables, and is one way to take parameter uncertainty into account.

## **1.2 What we do**

We consider a standard version of an abstract macroeconomic model, the generalized linear model of Evans and Honkapohja (2001). Instead of assuming standard recursive learning, we think of the private sector agents as being Bayesian econometricians. In particular, the agents will then treat estimated parameters as random variables. In certain circumstances, the system will behave as if the agents are classical recursive learners, but in general, the system will behave somewhat differently from the one where agents are classical econometricians. We highlight these differences and similarities. The primary question we wish to address is whether we can describe local convergence properties of systems with Bayesian learners in the same expectational stability terms as systems with standard recursive learning.

## **1.3 Main findings**

We find expectational stability conditions for systems with Bayesian learners. We are able to isolate cases where these conditions are identical to the

conditions for non-Bayesian systems. In these cases, in terms of expectational stability, the Bayesian systems yield no difference in results *vis-a-vis* the systems with standard recursive learning.

The actual stochastic dynamical systems produced by the classical recursive learning versus the Bayesian learning assumptions are not identical, however, except under special circumstances. This means that the dynamics of the two systems will differ during the transition to the rational expectations equilibrium, even if the local asymptotic stability properties do not differ. We document via examples how the dynamics of Bayesian systems can differ from the dynamics of non-Bayesian systems with identical shock sequences. We show situations in which the differences can be material and situations where the differences are likely to be negligible.

We interpret these findings as follows. When we replace the rational expectations agents in a model with recursive least squares learners, as has been standard in this literature, we are assuming a certain degree of bounded rationality. This has been discussed extensively in the literature. However, since the systems can converge, locally, to rational expectations equilibrium, the bounded rationality eventually dissipates, which is perhaps a comforting way to think about how rational expectations equilibrium is actually achieved. Still, one might worry that if the agents were a little more rational at the time that they adopt their learning algorithm, the local stability properties of the rational expectations equilibrium might be altered. Here, “a little more rational” means that the agents use Bayesian methods while learning instead of classical recursive algorithms, and so they take into account that they will be learning in the future. It is conceivable that equilibria which were unstable under standard recursive learning might now be stable under Bayesian learning, for instance. The results in this paper suggest that this is not the case. The expectational stability conditions for the systems with Bayesian learners are not any different, at least in the cases analyzed here, from those which are commonly studied in the literature. This suggests that the stability analysis following the tradition of Marcet and Sargent (1989) and Evans and Honkapohja (2001) may have very broad appeal, and that

the assumption of standard recursive learning may be less restrictive than commonly believed.

## 1.4 Recent related literature

Bray and Savin (1986) studied learning in a cobweb model and noted that a recursive least squares specification for the learning rule implied that agents assumed fixed coefficients in an environment where coefficients were actually time-varying.<sup>1</sup> They thought of this as a misspecification, a form of bounded rationality. They asked whether convergence to rational expectations might occur at a pace that was rapid enough to cause agents to not notice the misspecification using standard statistical tests. They illustrated some cases where this was true, and others where it was not. Bray and Savin (1986) used what we would call fixed coefficient Bayesian updating; this was their source of bounded rationality. We allow agents to see their estimated coefficients as random variables. Also, the cobweb model used in the classic Bray and Savin paper does not encompass the two-step ahead expectations which will play an important role in the results reported below.

McGough (2003) studies Bray and Savin's cobweb model but allows the agents to use a Kalman filter to update parameter estimates. This allows the agents to take into account the fact that estimates are time-varying.<sup>2</sup> He finds conditions under which such a system is expectationally stable. McGough also studies a Muth model with Kalman filter updating.

Cogley and Sargent (2008) study a partial equilibrium model with a representative Bayesian decision-maker. Like Bray and Savin, they are concerned that while the agent is learning using standard recursive algorithms, fixed coefficients are assumed in the learning rule, whereas actual coefficients change along the path to the rational expectations equilibrium.<sup>3</sup> To address this, they allow the household to behave as a Bayesian decision-maker. They il-

---

<sup>1</sup>This is the same concern raised by Cogley and Sargent (2008).

<sup>2</sup>Bullard (1992) also uses the Kalman filter to allow agents to take time-varying parameters into account.

<sup>3</sup>See the quote above.

illustrate differences in decisions when households are modeled as Bayesian versus rational expectations or standard recursive learners. They argue that the standard recursive learning approximation to the Bayesian household is actually very good in the problem they study. This theme will be echoed in the results reported below, as the systems under recursive learning will not behave too differently from the systems under Bayesian learning. Cogley and Sargent did not study the question of expectational stability. We, on the other hand, do not have households making economic decisions, but instead study the reduced form model of Evans and Honkapohja (2001).

Guidolin and Timmerman (2007) study an asset pricing model with Bayesian learning. They study the nature of the asset price dynamics in this setting, comparing Bayesian systems to those with rational expectations and standard recursive least squares, similar to Cogley and Sargent (2008).

Evans, Honkapohja, and Williams (2006) study stochastic gradient learning. They show that under certain conditions the stochastic gradient algorithm can approximate the Bayesian estimator. They display expectational stability conditions for their generalized stochastic gradient algorithm, and these conditions have clear similarities to those under standard recursive least squares.

In this paper, we think of systems in which private sector expectations are important, so that learning refers to private sector learning. However, some of the learning literature emphasizes policymaker learning with a rational expectations private sector. For instance, Sargent and Williams (2005) study the effect of priors on escape dynamics in a model where the government is learning. Wieland (2000) adapts the framework of Nyarko and Kiefer (1989) to study optimal control by a monetary authority when the authority is a Bayesian learner. We do not have any policy in this paper and so we cannot address these topics.

## 1.5 Organization

We present a version of the generalized linear model of Evans and Honkapohja in the next section. We analyze this model when the agents are Bayesian

learners. We find expectational stability conditions and show that they are the same as in the case of recursive learning. However, differences can arise along transition paths to the rational expectations equilibrium. We then turn to simulations to illustrate some of the issues involved.

## 2 Environment

Evans and Honkapohja (2001) study a general linear model which can be viewed as representative of a linear approximation to a rational expectations equilibrium. This provides a common framework which will allow us to compare results clearly. We study a somewhat less general, scalar version of their model given by

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E_{t-1}^* y_t + \beta_1 E_{t-1}^* y_{t+1} + v_t, \quad (1)$$

where  $v_t \sim \mathcal{N}(0, \nu^2)$ . Here  $y_t$  is the state of the economic system,  $\alpha$ ,  $\delta$ ,  $\beta_0$ , and  $\beta_1$  are scalar parameters, and  $E_{t-1}^*$  is a subjective expectations operator, as expectations may not initially be rational. We have chosen this particular version of Evans and Honkapohja (2001), equation (1), carefully. One might be tempted to set, say,  $\delta = 0$  and  $\beta_1 = 0$ , for instance. But as we show below, both of these will have to be nonzero in order to effectively see the differences between standard recursive learning and the Bayesian learners we wish to understand.

The minimal state variable (MSV) solution is given by

$$y_t = \bar{a} + \bar{b} y_{t-1} + v_t, \quad (2)$$

where  $\bar{a}$  and  $\bar{b}$  solve

$$\alpha + (\beta_0 + \beta_1)\bar{a} + \beta_1\bar{a}\bar{b} = \bar{a}, \quad (3)$$

and

$$\delta + \beta_0\bar{b} + \beta_1\bar{b}^2 = \bar{b}. \quad (4)$$

We stress that there may be two solutions  $\bar{b}$  which solve these equations. We assign a traditional perceived law of motion (PLM), which is consistent in

form with the MSV solution (2),

$$y_t = a + by_{t-1} + v_t. \quad (5)$$

The agents use the PLM to form expectations, which then can be substituted into equation (1) to produce an actual law of motion (ALM) for the system.

In the standard analysis, agents are assumed to use recursive least squares to update their parameter estimates. Using the PLM in (5) agents are assumed to forecast according to

$$\begin{aligned} E_{t-1}^* y_t &= \hat{a}_{t-1} + \hat{b}_{t-1} y_{t-1}, \\ E_{t-1}^* y_{t+1} &= E(\hat{a}_t + \hat{b}_t y_t + v_{t+1} | Y_{t-1}) \\ &= \hat{a}_{t-1}(1 + \hat{b}_{t-1}) + \hat{b}_{t-1}^2 y_{t-1}, \end{aligned} \quad (6)$$

with  $\hat{a}_t$  and  $\hat{b}_t$  denoting the least squares estimates through time  $t$ . Substituting these equations into equation (1) we obtain the actual law of motion under recursive least squares learning

$$y_t = [\alpha + (\beta_0 + \beta_1)\hat{a}_{t-1} + \beta_1\hat{a}_{t-1}\hat{b}_{t-1}] + [\delta + \beta_0\hat{b}_{t-1} + \beta_1\hat{b}_{t-1}^2]y_{t-1} + v_t. \quad (7)$$

The expectational stability of the system will depend on mapping from the perceived to the actual law of motion. We now wish to find the counterpart of the actual law of motion, equation (7), in the case of Bayesian learning in order to compare the two.

## 3 Real time Bayesian learning

### 3.1 Priors and posteriors

We wish to assume that the private sector agents in this economy use a Bayesian approach to updating the coefficients in their perceived law of motion, that is, the scalar coefficients  $a$  and  $b$ . They have priors which are given by

$$\phi'_0 = (a_0, b_0) \sim \mathcal{N}(\mu_0, \Sigma_0), \quad (8)$$

where  $\mu'_0 = (\mu_{a,0}, \mu_{b,0})$ , and

$$\Sigma_0 = \begin{bmatrix} \sigma_{a,0}^2 & \sigma_{ab,0} \\ \sigma_{ab,0} & \sigma_{b,0}^2 \end{bmatrix}, \quad (9)$$

where  $\sigma_{xy}$  indicates the covariance of  $x$  and  $y$ . The conditional distribution of the state  $y_t$  is

$$y_t | Y_{t-1}, \phi_{t-1} \sim \mathcal{N}(a_{t-1} + b_{t-1}y_{t-1}, \nu^2), \quad (10)$$

where  $Y_{t-1}$  is the history of  $y_t$ . The distribution of  $Y_t$  conditional on  $\phi$  is

$$\begin{aligned} f(Y_t | \phi) &= f(y_t | \phi, Y_{t-1}) f(Y_{t-1} | \phi) \\ &= f(y_t | \phi, Y_{t-1}) f(y_{t-1} | \phi, Y_{t-2}) \dots f(y_2 | \phi, y_1) f(y_1 | \phi). \end{aligned} \quad (11)$$

Using these expressions we can represent a posterior distribution of  $\phi_t$ , i.e.  $\phi | Y_t$ , as

$$\begin{aligned} f(\phi | Y_t) &\propto f(Y_t | \phi) f(\phi) \\ &\propto f(y_t | \phi, Y_{t-1}) f(y_{t-1} | \phi, Y_{t-2}) \dots f(y_2 | \phi, y_1) f(y_1 | \phi) f(\phi). \end{aligned} \quad (12)$$

Assuming  $f(y_1 | \phi)$  is known (for instance,  $f(y_1 | \phi) = 1$ ), we can obtain a Normal-Normal update given by

$$\begin{aligned} f(\phi | Y_t) &\propto \mathcal{N}(\phi' z_t) \dots \mathcal{N}(\phi' z_1) \mathcal{N}(\phi_0) \\ f(\phi | Y_t) &= \mathcal{N}(\mu_t, \Sigma_t), \end{aligned} \quad (13)$$

where  $z_t = (1, y_{t-1})'$ , and where

$$\mu_t = \Sigma_t (\Sigma_0^{-1} \phi_0 + \nu^{-2} (Z_t' Y_t)), \quad (14)$$

and

$$\Sigma_t = (\Sigma_0^{-1} + \nu^{-2} (Z_t' Z_t))^{-1}, \quad (15)$$

where  $Z_t$  is the history of  $z_t$ .

### 3.2 Recursive forms

Both  $\mu_t$  and  $\Sigma_t$  can be written in a recursive form. For  $\Sigma_t$ , we can write

$$\begin{aligned}
\Sigma_t^{-1} &= \Sigma_0^{-1} + \nu^{-2}(Z_t'Z_t) \\
&= \Sigma_0^{-1} + \nu^{-2} \sum_{i=1}^t z_i z_i' \\
&= \Sigma_0^{-1} + \nu^{-2} \sum_{i=1}^{t-1} z_i z_i' + \nu^{-2} z_t z_t' \\
&= \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t'.
\end{aligned} \tag{16}$$

For  $\mu_t$ , we use period-by-period updating, taking yesterday's estimate as today's prior:

$$\begin{aligned}
\mu_t &= \Sigma_t(\Sigma_{t-1}^{-1}\mu_{t-1} + \nu^{-2}z_t y_t), \\
&= \Sigma_t \Sigma_{t-1}^{-1} \mu_{t-1} + \Sigma_t \nu^{-2} z_t y_t, \\
\mu_t - \mu_{t-1} &= (\Sigma_t \Sigma_{t-1}^{-1} - I)\mu_{t-1} + \Sigma_t \nu^{-2} z_t y_t, \\
\mu_t &= \mu_{t-1} + \Sigma_t ((\Sigma_{t-1}^{-1} - \Sigma_t^{-1})\mu_{t-1} + \nu^{-2} z_t y_t),
\end{aligned} \tag{17}$$

where  $I$  is a conformable identity matrix. Substituting the expression  $\Sigma_t^{-1} = \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t'$ , we obtain

$$\begin{aligned}
\mu_t &= \mu_{t-1} + \Sigma_t (\nu^{-2} z_t y_t - \nu^{-2} z_t z_t' \mu_{t-1}) \\
&= \mu_{t-1} + \Sigma_t \nu^{-2} z_t (y_t - z_t' \mu_{t-1}).
\end{aligned} \tag{18}$$

### 3.3 The actual law of motion

To consider the evolution of the system we have to determine the ALM under Bayesian learning. We begin with the PLM under learning

$$\begin{aligned}
y_t &= a_{t-1} + b_{t-1} y_{t-1} + v_t \\
&= \phi_{t-1}' z_t + v_t,
\end{aligned} \tag{19}$$

where  $a_t = a|Y_t$ . We now take expectations based on the PLM in order to substitute these into (1) to obtain the ALM. The necessary expectation terms

are given by

$$\begin{aligned} E_{t-1}^* y_t &= E(a_{t-1} + b_{t-1} y_{t-1} + v_t | Y_{t-1}) \\ &= \mu'_{t-1} z_t, \end{aligned} \tag{20}$$

$$\begin{aligned} E_{t-1}^* y_{t+1} &= E(a_t + b_t y_t + v_{t+1} | Y_{t-1}) \\ &= E(\phi'_t z_{t+1} | Y_{t-1}). \end{aligned} \tag{21}$$

We stress that one of the hallmarks of the Bayesian approach is that both  $y_t$  and  $b_t$  are random variables. We next have to compute  $E(\phi'_t z_{t+1} | Y_{t-1})$ . We can write joint distribution of  $\phi$  and  $y$  as

$$f(\phi_t, y_t | Y_{t-1}) = \underbrace{f(\phi_t | Y_t)}_{\text{Posterior beliefs}} \cdot \underbrace{f(y_t | Y_{t-1})}_{\text{Posterior prediction}} \tag{22}$$

$$= \mathcal{N}_\phi(\mu_t, \Sigma_t) \mathcal{N}_y(\mu'_{t-1} z_t, \nu^2 + z'_t \Sigma_{t-1} z_t). \tag{23}$$

To see the second term of (23), we write the distribution of  $y_{t+1}$  conditional on  $Y_t$  as

$$\begin{aligned} f(y_{t+1} | Y_t) &= \int f(y_{t+1} | Y_t, \phi_t) f(\phi_t | Y_t) d\phi_t \\ &= \int \mathcal{N}_y(z'_{t+1} \phi_t, \nu^2) \mathcal{N}_\phi(\mu_t, \Sigma_t) d\phi_t \\ &= \mathcal{N}_y(\mu'_t z_{t+1}, \nu^2 + z'_{t+1} \Sigma_t z_{t+1}), \end{aligned} \tag{24}$$

so that  $f(y_t | Y_{t-1})$  is as given in (23).

The density function can be written as

$$f(\phi_t) = f \begin{pmatrix} a_t \\ b_t \end{pmatrix} = \mathcal{N} \left( \begin{bmatrix} \mu_{a,t} \\ \mu_{b,t} \end{bmatrix}, \begin{bmatrix} \sigma_{a,t}^2 & \sigma_{ab,t} \\ \sigma_{ab,t} & \sigma_{b,t}^2 \end{bmatrix} \right). \tag{25}$$

Also, using (18),

$$\begin{aligned}
\mu_t &= \begin{pmatrix} \mu_{a,t} \\ \mu_{b,t} \end{pmatrix} = \mu_{t-1} + \Sigma_t \nu^{-2} z_t (y_t - z_t' \mu_t) \\
&= \begin{pmatrix} \mu_{a,t-1} \\ \mu_{b,t-1} \end{pmatrix} + \begin{pmatrix} \sigma_{a,t}^2 & \sigma_{ab,t} \\ \sigma_{ab,t} & \sigma_{b,t}^2 \end{pmatrix} \nu^{-2} \begin{pmatrix} 1 \\ y_{t-1} \end{pmatrix} (y_t - a_{t-1} - b_{t-1} y_{t-1}) \\
\mu_{a,t} &= \mu_{a,t-1} + \underbrace{\nu^{-2} (\sigma_{a,t}^2 + \sigma_{ab,t} y_{t-1})}_{\Sigma_{a,t}} (y_t - a_{t-1} - b_{t-1} y_{t-1}) \\
\mu_{b,t} &= \mu_{b,t-1} + \underbrace{\nu^{-2} (\sigma_{ab,t} + \sigma_{b,t}^2 y_{t-1})}_{\Sigma_{b,t}} (y_t - a_{t-1} - b_{t-1} y_{t-1}). \tag{26}
\end{aligned}$$

We can write

$$\begin{aligned}
f(b_t, y_t | Y_{t-1}) &= f(b_t | y_t, Y_{t-1}) f(y_t | Y_{t-1}) \\
&= \mathcal{N}_b(\mu_{b,t}, \sigma_{b,t}^2) \mathcal{N}_y(\mu'_{t-1} z_t, \nu^2 + z_t' \Sigma_{t-1} z_t) \tag{27}
\end{aligned}$$

We are interested in an expression for  $E(\phi_t' z_{t+1} | Y_{t-1})$ . As we have a joint distribution of both random variables we can compute the expectations directly:

$$\begin{aligned}
E(\phi_t' z_{t+1} | Y_{t-1}) &= E \left( (a_t, b_t) \begin{pmatrix} 1 \\ y_t \end{pmatrix} | Y_{t-1} \right) \\
&= E(a_t + b_t y_t | Y_{t-1}) \\
&= E(a_t | Y_{t-1}) + E(b_t y_t | Y_{t-1}). \tag{28}
\end{aligned}$$

Consider  $E(b_t y_t | Y_{t-1})$ :

$$\begin{aligned}
E(b_t y_t | Y_{t-1}) &= \int \int b_t y_t f(b_t, y_t | Y_{t-1}) dy_t db_t \\
&= \int \int b_t y_t \mathcal{N}_{b_t}(\mu_{b,t}, \sigma_{b,t}^2) \mathcal{N}_{y_t}(\mu'_{t-1} z_t, \nu^2 + z_t' \Sigma_{t-1} z_t) db_t dy_t. \tag{29}
\end{aligned}$$

As  $\mathcal{N}_{y_t}$  does not depend on  $b_t$  we can write it as

$$\begin{aligned}
E(b_t y_t | Y_{t-1}) &= \int y_t \mathcal{N}_{y_t}(\mu'_{t-1} z_t, \underbrace{v^2 + z_t' \Sigma_{t-1} z_t}_{\Omega_{y_t}}) \underbrace{\int b_t \mathcal{N}_{b_t}(\mu_{b,t}, \sigma_{b,t}^2) db_t}_{Eb_t = \mu_{b,t}} dy_t \\
&= \int \mu_{b,t} y_t \mathcal{N}_{y_t}(\mu'_{t-1} z_t, \Omega_{y_t}) dy_t \\
&= \int (\mu_{b,t-1} + \Sigma_{b,t}(y_t - a_{t-1} - b_{t-1} y_{t-1})) y_t \mathcal{N}_{y_t}(\mu'_{t-1} z_t, \Omega_{y_t}) dy_t \\
&= (\mu_{b,t-1} - \Sigma_{b,t}(a_{t-1} + b_{t-1} y_{t-1})) \underbrace{\int y_t \mathcal{N}_{y_t}(\mu'_{t-1} z_t, \Omega_{y_t}) dy_t}_{Ey_t} \\
&\quad + \Sigma_{b,t} \underbrace{\int y_t^2 \mathcal{N}_{y_t}(\mu'_{t-1} z_t, \Omega_{y_t}) dy_t}_{Ey_t^2 = \text{Var}(y_t) + (Ey_t)^2} \\
&= (\mu_{b,t-1} - \Sigma_{b,t}(a_{t-1} + b_{t-1} y_{t-1})) Ey_t + \Sigma_{b,t} \text{Var}(y_t) + \Sigma_{b,t} (Ey_t)^2. \quad (30)
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
E(b_t y_t | Y_{t-1}) &= (\mu_{b,t-1} + \Sigma_{b,t}(Ey_t - a_{t-1} - b_{t-1} y_{t-1})) Ey_t + \Sigma_{b,t} \text{Var}(y_t) \\
&= E(\mu_{b,t} | Y_{t-1}) E(y_t | Y_{t-1}) + \Sigma_{b,t} \Omega_{y_t}. \quad (31)
\end{aligned}$$

Then,

$$\begin{aligned}
E(\phi'_t z_{t+1} | Y_{t-1}) &= E(a_t | Y_{t-1}) + E(b_t y_t | Y_{t-1}) \\
&= E(\mu_{a,t} | Y_{t-1}) + E(\mu_{b,t} | Y_{t-1}) E(y_t | Y_{t-1}) + \Sigma_{b,t} \Omega_{y_t} \\
&= E(\mu_t | Y_{t-1})' E(z_{t+1} | Y_{t-1}) + \Sigma_{b,t} \Omega_{y_t} \\
&= \mu'_{t-1} \begin{pmatrix} 1 \\ \mu'_{t-1} z_t \end{pmatrix} + \Sigma_{b,t} \Omega_{y_t}. \quad (32)
\end{aligned}$$

Recall that

$$E_{t-1}^* y_t = E(a_{t-1} + b_{t-1} y_{t-1} + v_t | Y_{t-1}) = \mu'_{t-1} z_t. \quad (33)$$

Substituting these expressions into (1) under Bayesian learning we obtain

the following expression:

$$\begin{aligned}
y_t &= \alpha + \delta y_{t-1} + \beta_0 E_{t-1}^* y_t + \beta_1 E_{t-1}^* y_{t+1} + v_t, \\
&= \alpha + \delta y_{t-1} + \beta_0 \mu'_{t-1} z_t + \beta_1 \mu'_{t-1} \begin{pmatrix} 1 \\ \mu'_{t-1} z_t \end{pmatrix} + \beta_1 \Sigma_{b,t} \Omega_{y_t} + v_t, \\
&= \alpha + \delta y_{t-1} + \beta_0 (\mu_{a,t-1} + \mu_{b,t-1} y_{t-1}) \\
&\quad + \beta_1 (\mu_{a,t-1} + \mu_{b,t-1} (\mu_{a,t-1} + \mu_{b,t-1} y_{t-1})) + \beta_1 \Sigma_{b,t} \Omega_{y_t} + v_t. \tag{34}
\end{aligned}$$

Finally, rearranging this expression, we are ready to present our key equation. In particular, we conclude that the actual law of motion under Bayesian learning can be written as

$$\begin{aligned}
y_t &= [\alpha + (\beta_0 + \beta_1) \mu_{a,t-1} + \beta_1 \mu_{a,t-1} \mu_{b,t-1}] \\
&\quad + [\delta + \beta_0 \mu_{b,t-1} + \beta_1 \mu_{b,t-1}^2] y_{t-1} + \beta_1 \Sigma_{b,t} \Omega_{y_t} + v_t. \tag{35}
\end{aligned}$$

Except for the term  $\beta_1 \Sigma_{b,t} \Omega_{y_t}$ , the above expression is exactly analogous to what one would obtain under standard recursive least squares [as shown by equation (7) above] as analyzed by Evans and Honkapohja (2001) for the MSV solution, but with parameters in the RLS case being here represented by their means.

### 3.4 Remarks on the Bayesian ALM

We said that we chose equation (1) carefully. In particular, we made sure that a lagged endogenous variable was included with a non-zero coefficient  $\delta$ , and that a two-step ahead expectation was included with a non-zero coefficient  $\beta_1$ . By considering the actual law of motion under Bayesian learning, we can show clearly why both  $\delta \neq 0$  and  $\beta_1 \neq 0$  are necessary to see the differences between standard recursive learning and Bayesian learning. First, if  $\beta_1 = 0$ , then the term  $\beta_1 \Sigma_{b,t} \Omega_{y_t}$  drops out of the expression (35). Second, if  $\delta = 0$ , then there would be no term  $\Omega_{y_t}$ , as the MSV solution (2) would not depend on  $y_{t-1}$ , and so the agents would only need to estimate means.

To return to a standard recursive learning case, we would have to make two assumptions. One is that the agents use the standard recursive least

squares estimator instead of the Bayesian estimator, and the second is that agents treat parameter estimates as constants when using their PLM to form expectations. So, there are really two levels to Bayesian learning. One is that the agents use the Bayesian estimators  $\mu_a$  and  $\mu_b$ , and the second is that the agents treat the estimates as random variables, not constants, which gives rise to the term  $\beta_1 \Sigma_{b,t} \Omega_{y_t}$  in the actual law of motion (35). It is important to stress that even systems with Bayesian estimation only (e.g.,  $\beta_1 = 0$ ) do not produce an actual law of motion equivalent to the RLS case, because  $\mu_a$  and  $\mu_b$  are not treated as constants.

### 3.5 Alternative expressions for the ALM

In order to work with the expression (35), we can write it in an expanded fashion. First, consider  $\Sigma_{b,t} \Omega_{y_t}$ :

$$\begin{aligned} \Sigma_t^{-1} &= \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t' \\ &= \begin{pmatrix} \sigma_{a,t-1}^2 & \sigma_{ab,t-1} \\ \sigma_{ab,t-1} & \sigma_{b,t-1}^2 \end{pmatrix}^{-1} + \nu^{-2} \begin{pmatrix} 1 & y_{t-1} \\ y_{t-1} & y_{t-1}^2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\sigma_{b,t-1}^2}{A_{t-1}} + \nu^{-2} & -\frac{\sigma_{ab,t-1}}{A_{t-1}} + \nu^{-2} y_{t-1} \\ -\frac{\sigma_{ab,t-1}}{A_{t-1}} + \nu^{-2} y_{t-1} & \frac{\sigma_{a,t-1}^2}{A_{t-1}} + \nu^{-2} y_{t-1}^2 \end{pmatrix}, \end{aligned} \quad (36)$$

where  $A_{t-1} = \sigma_{a,t-1}^2 \sigma_{b,t-1}^2 - \sigma_{ab,t-1}^2$  is the determinant of  $\Sigma_{t-1}$ . Then,

$$\begin{aligned} \Sigma_t &= (\Sigma_t^{-1})^{-1} \\ &= \begin{pmatrix} \frac{\sigma_{a,t-1}^2}{A_{t-1} A_t^I} + \frac{\nu^{-2} y_{t-1}^2}{A_t^I} & \frac{\sigma_{ab,t-1}}{A_{t-1} A_t^I} - \frac{\nu^{-2} y_{t-1}}{A_t^I} \\ \frac{\sigma_{ab,t-1}}{A_{t-1} A_t^I} - \frac{\nu^{-2} y_{t-1}}{A_t^I} & \frac{\sigma_{b,t-1}^2}{A_{t-1} A_t^I} + \frac{\nu^{-2}}{A_t^I} \end{pmatrix}, \end{aligned} \quad (37)$$

where  $A_t^I = \det(\Sigma_t^{-1})$ . We defined  $\Sigma_{b,t}$  as

$$\Sigma_{b,t} = X \nu^{-2} \Sigma_t z_t = \nu^{-2} (\sigma_{ab,t} + \sigma_{b,t}^2 y_{t-1}), \quad (38)$$

with  $X = (0 \ 1)$ . Therefore,

$$\begin{aligned}
\Sigma_{b,t} &= \nu^{-2} \left[ \frac{\sigma_{ab,t-1}}{A_{t-1}A_t^I} - \frac{\nu^{-2}y_{t-1}}{A_t^I} + \left( \frac{\sigma_{b,t-1}^2}{A_{t-1}A_t^I} + \frac{\nu^{-2}}{A_t^I} \right) y_{t-1} \right], \\
&= \frac{\nu^{-2}}{A_{t-1}A_t^I} (\sigma_{ab,t-1} + \sigma_{b,t-1}^2 y_{t-1}), \\
&= \frac{\sigma_{ab,t-1} + \sigma_{b,t-1}^2 y_{t-1}}{\nu^2 + \sigma_{a,t-1}^2 + 2\sigma_{ab,t-1}y_{t-1} + \sigma_{b,t-1}^2 y_{t-1}^2}, \tag{39}
\end{aligned}$$

as

$$A_t^I = \frac{1}{\nu^2 A_{t-1}} (\nu^2 + \sigma_{a,t-1}^2 + 2\sigma_{ab,t-1}y_{t-1} + \sigma_{b,t-1}^2 y_{t-1}^2). \tag{40}$$

We are ultimately interested in  $\Sigma_{b,t}\Omega_{y_t}$ . Using

$$\begin{aligned}
\Omega_{y_t} &= \text{Var}(y_t|Y_{t-1}) = \nu^2 + z_t' \Sigma_{t-1} z_t \\
&= \nu^2 + \sigma_{a,t-1}^2 + 2y_{t-1}\sigma_{ab,t-1} + y_{t-1}^2\sigma_{b,t-1}^2, \tag{41}
\end{aligned}$$

we can express  $\Sigma_{b,t}\Omega_{y_t}$  as

$$\Sigma_{b,t}\Omega_{y_t} = \sigma_{ab,t-1} + \sigma_{b,t-1}^2 y_{t-1}. \tag{42}$$

Substituting this expression into the ALM yields

$$\begin{aligned}
y_t &= [\alpha + (\beta_0 + \beta_1)\mu_{a,t-1} + \beta_1\mu_{a,t-1}\mu_{b,t-1} + \beta_1\sigma_{ab,t-1}] \\
&\quad + [\delta + \beta_0\mu_{b,t-1} + \beta_1\mu_{b,t-1}^2 + \beta_1\sigma_{b,t-1}^2] y_{t-1} + v_t. \tag{43}
\end{aligned}$$

This is an  $AR(1)$  process, consistent with the perceived law of motion, given beliefs at date  $t$ . Using this alternative expression for the actual law of motion allows us to define a T-map in a convenient way.

## 4 Expectational stability

In this section we turn to an analysis of expectational stability. Agents have beliefs about the parameters in their PLM and update them using Bayes

rule. Conditional on information at time  $t$ , that is, the observed sequence of  $\{y_\tau\}_{\tau=1}^t = Y_t$ , their beliefs are given by

$$f(\phi|Y_t) = \mathcal{N}(\mu_t, \Sigma_t), \quad (44)$$

where  $\mu_t$  and  $\Sigma_t$  have the recursive form

$$\mu_t = \mu_{t-1} + \Sigma_t \nu^{-2} z_t (y_t - z_t' \mu_{t-1}), \quad (45)$$

$$\Sigma_t^{-1} = \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t', \quad (46)$$

where  $y_t$  in the first equation is given by the actual law of motion in equation (43) above. The evolution of the mean of the distribution is given by

$$\begin{aligned} \mu_t = & \mu_{t-1} + \Sigma_t \nu^{-2} z_t (\alpha + (\beta_0 + \beta_1) \mu_{a,t-1} + \beta_1 \mu_{a,t-1} \mu_{b,t-1} + \beta_1 \sigma_{ab,t-1} \\ & + [\delta + \beta_0 \mu_{b,t-1} + \beta_1 \mu_{b,t-1}^2] y_{t-1} + v_t - z_t' \mu_{t-1}). \end{aligned} \quad (47)$$

Define a T-map as

$$T_a(\mu, \Sigma) = \alpha + (\beta_0 + \beta_1) \mu_a + \beta_1 \mu_a \mu_b + \beta_1 \sigma_{ab} \quad (48)$$

$$T_b(\mu, \Sigma) = \delta + \beta_0 \mu_b + \beta_1 \mu_b^2 + \beta_1 \sigma_b^2. \quad (49)$$

Rewriting  $\Sigma_t = \frac{1}{t} R_t^{-1}$ , where

$$R_t = (1/t) \Sigma_0^{-1} + (1/t) \nu^{-2} Z_t' Z_t, \quad (50)$$

and defining  $S_{t-1} = R_t$ , we can represent the problem in the stochastic recursive form,<sup>4</sup>

$$\mu_t = \mu_{t-1} + t^{-1} \nu^{-2} S_{t-1}^{-1} z_t (z_t' (T(\mu_{t-1}, S_{t-2}) - \mu_{t-1}) - v_t), \quad (51)$$

$$\begin{aligned} S_t = & S_{t-1} + t^{-1} (\nu^{-2} z_{t+1} z_{t+1}' - S_{t-1}) \\ & + t^{-2} \left( -\frac{t}{t+1} \right) (\nu^{-2} z_{t+1} z_{t+1}' - S_{t-1}). \end{aligned} \quad (52)$$

---

<sup>4</sup>See Evans and Honkapohja (2001, Section 8.4) for technical conditions on the recursive stochastic algorithm.

Using the stochastic recursive algorithm we can approximate the above system with the ordinary differential equation

$$\begin{aligned} \frac{d\mu}{d\tau} &= h(\mu) \\ &= \lim_{t \rightarrow \infty} E\nu^{-2} S^{-1} z_t(z'_t(T(\mu, S) - \mu) - v_t) \\ &= \tilde{T}(\mu) - \mu \end{aligned} \tag{53}$$

as

$$\lim_{t \rightarrow \infty} T(\mu, S) = \tilde{T}(\mu) \tag{54}$$

with

$$\tilde{T}_a(\mu) = \alpha + (\beta_0 + \beta_1)\mu_a + \beta_1\mu_a\mu_b \tag{55}$$

$$\tilde{T}_b(\mu) = \delta + \beta_0\mu_b + \beta_1\mu_b^2. \tag{56}$$

Linearizing and computing the eigenvalues of  $\tilde{T}(\mu)$  at an equilibrium, we obtain the stability conditions

$$\begin{aligned} \beta_0 + \beta_1 + \beta_1\mu_b - 1 &< 0 \\ \beta_0 - 1 + 2\beta_1\mu_b &< 0 \end{aligned} \tag{57}$$

These conditions are identical to the ones shown by Evans and Honkapohja (2001) to govern expectational stability under recursive least squares. We conclude that the system under Bayesian learning has the same E-stability conditions as with classical recursive learning.

This is an important finding, as it shows that concerns about the stability of macroeconomic systems under learning are equally relevant under a Bayesian learning assumption as under a recursive learning assumption. In particular, the system with Bayesian learners could be locally stable or unstable—it may or may not converge locally to the rational expectations equilibrium if expectations were initially displaced a small distance away from the REE. And, in fact, for the system we study the expectational stability conditions are the same under the two assumptions.

The intuition for this result is straightforward, given the known connections between Bayesian and least squares estimation methodology. In particular, if the variance terms vanish as data accumulates, and if in addition the estimators converge to their means, then the system will in effect be the same one which is analyzed for expectational stability under recursive least squares. The agents using recursive learning are assuming no variance terms and replacing mean parameter estimates with constants at the outset.

Even though stability properties are not altered, we have already stressed that the actual law of motion will not be the same under a Bayesian assumption relative to RLS. This means transition paths will be altered under Bayesian learning relative to RLS. We now turn to this issue.

## 5 Dynamics

### 5.1 Approach and parameterization

To illustrate above findings we conduct numerical simulations based on a version of an example taken from Evans and Honkapohja (2001, Section 8.5). Our version is intended to illustrate differences in the two systems as clearly as possible. Accordingly, we consider again the model

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E_{t-1}^* y_t + \beta_1 E_{t-1}^* y_{t+1} + v_t, \quad (58)$$

with parameter values  $\alpha = 2$ ,  $\delta = 0.3$ ,  $\beta_0 = 0.5$ , and  $\beta_1 = -0.4$ . The two  $AR(1)$  MSV solutions are  $(\bar{a}_1, \bar{b}_1) = (1.86, 0.44)$  and  $(\bar{a}_2, \bar{b}_2) = (8.97, -1.69)$ . Clearly, only the first solution is stationary and, in accordance with (57), E-stable.

We compare transition paths generated by agents with three different learning procedures. First, the *recursive least squares* case serves as a benchmark. Our second case is *Bayesian learning*. And, in order to isolate the effect of prior beliefs on the transition path in the Bayesian learning case we also consider a third case, *passive Bayesian estimation*, in which estimates are treated not as realizations of random variables but as constants, just as in the standard recursive learning case. In addition, we consider alternative

priors, each with a different precision, for both Bayesian learners and passive Bayesian estimation agents.

The initial settings of parameters, in the case of recursive least squares, and priors, in cases of Bayesian learning and passive Bayesian estimation, are at the stationary solution  $(\bar{a}_1, \bar{b}_1)$ . The lagged value of  $y$  is equal to unconditional mean of  $y$ . For each parameterization, we conduct 1,000 simulations and report the median realization to characterize the typical dynamics.

## 5.2 Bayesian learning dynamics can differ from RLS

We start with a comparison of the evolution of the RLS and Bayesian learning systems. The theory predicts that even though the expectational stability conditions are the same, the dynamics will be different. Figure 1 illustrates this point. In the Figure, the horizontal dotted line represents the rational expectations value of the parameter. Parameters estimated with both recursive least squares and Bayesian learning converge to rational expectations equilibrium.<sup>5</sup> However, it is also evident that the dynamic paths of  $a_t$  and  $b_t$  differ—and that these differences decrease over time. Figure 2 depicts first 100 periods from the same simulation. In this figure the difference between the two learning procedures is more pronounced.

In both figures, the estimates of Bayesian learners are closer to the rational expectations values than the recursive least squares estimates.

## 5.3 Bayesian learning versus Bayesian estimation

As we mentioned earlier, there are two levels of Bayesian learning. One is that the agents use the Bayesian estimator  $\mu_a$  and  $\mu_b$ , and the second is that the agents treat the estimates as random variables, which gives rise to the term  $\beta_1 \Sigma_{b,t} \Omega_{y_t}$  in equation (35). In order to distinguish between these two versions we can compare recursive least squares and Bayesian learning to the third case, passive Bayesian estimation.

---

<sup>5</sup>The relatively slow convergence is typical result for learning of AR(1) processes. See the discussion in Evans and Honkapohja (2001).

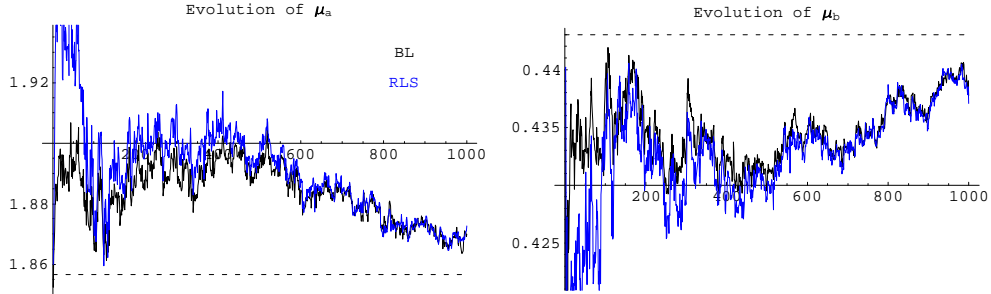


Figure 1: Bayesian learning vs RLS.

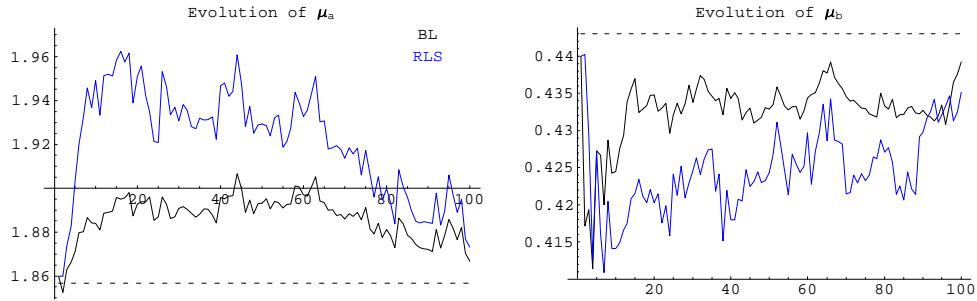


Figure 2: Bayesian learning vs RLS, first 100 periods

In Figure 3, we have added the simulated median path of estimated parameters with the passive Bayesian estimation (PBE). One advantage of plotting all three median trajectories is that we can decompose the Bayesian learning effect on learning dynamics into two components. The difference between PBE and RLS trajectories is the result of informative priors.<sup>6</sup> The alternative paths for Bayesian learning and PBE are the result, in turn, of the additional variance-covariance term in the Bayesian learning expression, stemming from (35).

<sup>6</sup>In the case of non-informative priors PBE and RLS are the same.

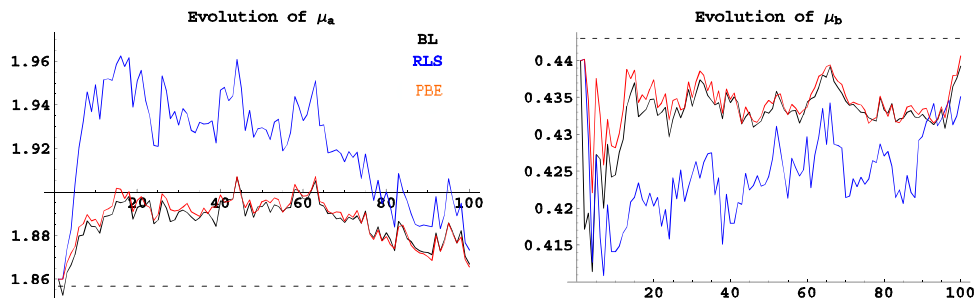


Figure 3: Bayesian learning and Bayesian estimation.

The striking feature of Figure 3 is that PBE and BL median trajectories are extremely close to one another, relative to the difference between these two trajectories and the trajectory of the recursive learning case. This suggests that the effects of priors are more significant in these examples than any contribution coming from the additional variance-covariance term.

## 5.4 Effects of priors

Bayesians have priors that may differ from an uninformative state, while standard recursive least squares does not. As Figure 3 illustrates, the precision of prior beliefs can be relatively more important for transition paths. Figure 4 depicts alternative trajectories of  $a_t$  and  $b_t$  for different prior variances.<sup>7</sup> The priors here are always centered at rational expectations values.<sup>8</sup> The increase in the precision of prior beliefs decreases the variability of the trajectory and moves it closer to rational expectations equilibrium. We stress, however, that this pattern is the result of prior beliefs being centered at rational expectations values. If the priors were centered at any other point, the increased precision of the prior would cause *slower* convergence to REE. We

<sup>7</sup>Since the variance is equal to inverse of precision,  $x = 4$  indicates variance of prior beliefs,  $\sigma$ , equal  $1/4$ .

<sup>8</sup>This removes a degree of freedom from the simulations. Since expectational stability is a local concept, this seems reasonable.

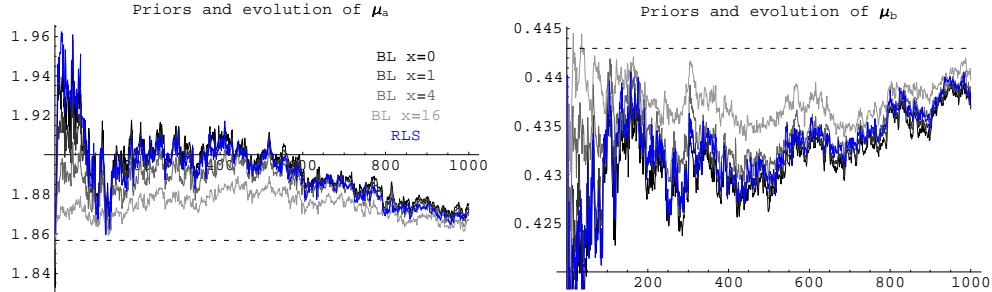


Figure 4: Effects of precision of priors.

think this point is well understood and we do not illustrate it here.

## 6 Conclusion

We have shown how to incorporate Bayesian learners into a standard linear recursive macroeconomic system, similar to ones studied by Evans and Honkapohja (2001). In order to illustrate the nature of the differences between Bayesian learning and a standard recursive least squares approach, we included a lagged endogenous variable in the system as well as a two-step ahead expectation. Without these features of the model, differences with RLS would not exist or would be more difficult to see. The analysis here is for an abstract linear scalar system, and it remains to be seen how these results would translate into commonly-studied macroeconomic models with microfoundations. Those systems would presumably have Bayesian decision-making as well as Bayesian updating.

A key result is that the system under Bayesian learning has the same expectational stability properties as the system under standard recursive learning. That is, expectational stability conditions are unaffected here by the introduction of Bayesian versus classical econometricians. Systems like this under Bayesian learning are just as likely or unlikely to meet expect-

tational stability requirements as equivalent systems under recursive least squares. Of the “two views” mentioned in the introduction, the suggestion that Bayesian estimation is a close relative of recursive least squares turned out to be the more prescient. Although agents here understand that they will be updating again in the future, this does not alter expectational stability findings.

However, we also show that the actual law of motion under Bayesian learning is in general not identical to the ALM under recursive least squares. This means that actual transition dynamics will differ under the two assumptions. This may be material for studies that wish to make quantitative statements about learning dynamics. We illustrated a few of these differences.

## References

- [1] Bray, M., and N. Savin. 1986. Rational expectations equilibria, learning, and model specification. *Econometrica* 54: 1129-1160.
- [2] Bullard, J. 1992. Time-varying and nonconvergence to rational expectations equilibrium under least squares learning. *Economics Letters* 40: 159-166.
- [3] Cogley, T., and T. Sargent. 2008. Anticipated utility and rational expectations as approximations of Bayesian decision making. *International Economic Review* 49(1): 185-221.
- [4] Evans, G., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton.
- [5] Evans, G., S. Honkapohja, and N. Williams. 2006. Generalized stochastic gradient learning. Manuscript, University of Oregon, Cambridge University, and Princeton University.
- [6] Guidolin, M., and A. Timmermann. 2007. Properties of equilibrium asset prices under alternative learning schemes. *Journal of Economic Dynamics and Control* 31(1): 161-217.

- [7] Kiefer, N., and Y. Nyarko. 1989. Optimal control of an unknown linear process with learning. *International Economic Review* 30(3): 571-586.
- [8] Kreps, D. 1998. Anticipated utility and dynamic choice. In D. P. Jacobs, E. Kalai, and M. Kamien, eds., *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983-1997*: 242-274.
- [9] 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.
- [10] McGough, B. 2003. Statistical learning with time-varying parameters. *Macroeconomic Dynamics* 7: 119-139.
- [11] Sargent, T., and N. Williams. 2005. Impacts of priors on convergence and escapes from Nash inflation. *Review of Economic Dynamics* 8: 360-391.
- [12] Wieland, V. 2000. Monetary policy, parameter uncertainty and optimal learning. *Journal of Monetary Economics* 46(1): 199-228.