Stability of Macroeconomic Systems with Bayesian Learners

James Bullard[*]
Federal Reserve Bank of St. Louis

Jacek Suda[†]
Washington University

30 October 2007
PRELIMINARY AND INCOMPLETE
Not for distribution

Abstract

We study macroeconomic systems in which expectations play an important role. Consistent with the recent literature on learning and expectations, we replace the agents in the economy with econometricians. Our econometricians are Bayesian learners. 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.

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 or the Federal Reserve System.

[†]Email: jrsuda@artsci.wustl.edu

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. This has led some observers to the conclusion that the learning literature may be somewhat "less rational" than it should be. Cogley and Sargent (forthcoming), for example, have noted that there is a "mild schizophrenia" embedded in the anticipated utility approach to learning that has become popular. In this paper, we take a first step toward addressing this criticism.

We study recursive learning in macroeconomic systems in a completely standard setting studied by Evans and Honkapohja (2001). However, our econometricians are Bayesian learners instead of least squares learners. The primary question we wish to address is whether we can describe local convergence properties of systems with Bayesian learners in the same terms as systems with recursive learning.

1.2 What we do

We consider a standard version of the generalized linear model of Evans and Honkapohja (2001). Instead of assuming ordinary recursive learning, we think of the private sector agents as being Bayesian econometricians. In certain circumstances, they will behave as if they are classical recursive learners, but in general, they will behave somewhat differently from classical econometricians. We highlight these differences and similarities.

1.3 Main findings

We find expectational stability conditions for systems with Bayesian learners.

We isolate cases where these conditions are identical to the conditions for non-Bayesian systems.

We document how the dynamics of Bayesian systems can differ from the dynamics of non-Bayesian systems.

We interpret these findings as follows. When we replace the rational expectations agents in a model with recursive 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 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 compromised. Equilibria which appeared to be stable might no longer be stable, for instance. The results in this paper suggest that this fear may be overblown. The expectational stability conditions for the systems with "rational learners" are not any different, at least in the classic cases studied 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 broad appeal, and that the assumption of recursive learning may be less restrictive than commonly believed.

The literature on Bayesian learning has left the impression that there would not be stability conditions attached in the case of Bayesian updating. But we show otherwise.

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.[1] 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.[2] He finds conditions under which such a system is expectationally stable. McGough also studies a Muth model with Kalman filter updating.

Cogley and Sargent (forthcoming) 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. They called this a form of "mild schizophrenia." The household is learning, but assumes that no learning will take place after today's update.[3] To address this, they allow the household to behave as a Bayesian decision-maker. They illustrate differences in decisions when households are modeled as Bayesian

[1] This is the same concern raised by Cogley and Sargent (forthcoming).

[2] Bullard (1992) also uses the Kalman filter to allow agents to take time-varying parameters into account.

[3] Kreps (1998) called this an *anticipated utility* model.

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.

Guidolin and Timmerman (2007) study a partial equilibrium 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 (forthcoming).

Evans, Honkapohja, and Williams (2006) study stochastic gradient learning. They show that under certain conditions the stochastic gain algorithm can approximate the Bayesian estimator. They have expectational stability conditions for the generalized stochastic gradient algorithm, which do not differ very much from those under standard recursive least squares.

In this paper, we think of the systems as describing private sector learning. However, some of the learning literature emphasizes policymaker learning with a rational private sector. For instance, Sargent and Williams (2004) 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.

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

2.1 A version of the general linear model

Evans and Honkapohja (2001) study a general linear model which can be viewed as a linear approximation to a rational expectations equilibrium of a microfounded model, such as a NK model or a RBC model. We study a somewhat less general, scalar version of their model given by

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E^*_{t-1} y_t + \beta_1 E^*_{t-1} y_{t+1} + v_t \tag{1}$$

where $v_t \sim \mathcal{N}(0, \nu^2)$. Here y_t is the state of the economic system, α, δ, β_0, and β_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 model.

The minimal state variable (MSV) solution is given by

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

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

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

$$\delta + \beta_0 \bar{b} + \beta_1 \bar{b}^2 = \bar{b} \ . \tag{4}$$

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),

$$y_t = a + b y_{t-1} + v_t. \tag{5}$$

The PLM then induces an actual law of motion (ALM) which is given below.

2.2 Real time Bayesian learning

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

$$\phi_0' = (a_0, b_0) \sim \mathcal{N}(\mu_0, \Sigma_0), \tag{6}$$

where $\mu_0' = \left(\mu_{a,0}, \mu_{b,0}\right)$ and

$$\Sigma_0 = \begin{bmatrix} \sigma_{a,0}^2 & \sigma_{ab,0} \\ \sigma_{ba,0} & \sigma_{b,0}^2 \end{bmatrix}, \tag{7}$$

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

$$y_t | Y_{t-1}, \phi_{t-1} \sim \mathcal{N}(a_{t-1} + b_{t-1} y_{t-1}, \nu^2), \tag{8}$$

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

$$\begin{aligned} f(Y_t|\phi_t) &= f(y_t|\phi_t, Y_{t-1}) f(Y_{t-1}|\phi_t) & (9) \\ &= f(y_t|\phi_t, Y_{t-1}) f(y_{t-1}|\phi_t, Y_{t-2}) \ldots f(y_2|\phi_t, y_1) f(y_1|\phi_t). & (10) \end{aligned}$$

Using these expressions we can represent a posterior distribution of ϕ_t as

$$\begin{aligned} f(\phi|Y_t) &\propto f(Y_t|\phi) f(\phi) & (11) \\ &\propto f(y_t|\phi, Y_{t-1}) f(y_{t-1}|\phi, Y_{t-2}) \ldots f(y_2|\phi, y_1) f(y_1|\phi) f(\phi). & (12) \end{aligned}$$

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

$$\begin{aligned} f(\phi|Y_t) &\propto \mathcal{N}(\phi' z_t) \cdots \mathcal{N}(\phi' z_1) \mathcal{N}(\phi_0) & (13) \\ f(\phi|Y_t) &= \mathcal{N}(\mu_t, \Sigma_t), & (14) \end{aligned}$$

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

$$\mu_t = \Sigma_t \left(\Sigma_0^{-1} \theta_0 + \nu^{-2} (Z_t' Y_t) \right), \tag{15}$$

and

$$\Sigma_t = \left(\Sigma_0^{-1} + \nu^{-2} (Z_t' Z_t) \right)^{-1}, \tag{16}$$

where Z_t is the history of z_t,

2.3 Recursive forms

Both μ_t and Σ_t can be written in the recursive form. For Σ_t,

$$\begin{aligned}
\Sigma_t^{-1} &= \Sigma_0^{-1} + \nu^{-2}(Z_t' Z_t) \\
&= \Sigma_0^{-1} + \nu^{-2} \sum_{i=1}^{t} z_i z_i' \\
&= \Sigma_0^{-1} + \nu^{-2} \sum_{i=1}^{t-1} z_i z_i' + \nu^{-2} z_t z_t' \\
&= \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t'. \qquad (17)
\end{aligned}$$

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

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

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

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

2.4 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

$$\begin{aligned}
y_t &= a_{t-1} + b_{t-1} y_{t-1} + v_t \qquad (19) \\
&= \phi_{t-1}' z_t + v_t, \qquad (20)
\end{aligned}$$

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

$$\begin{aligned} E^*_{t-1}y_t &= E(a_{t-1}+b_{t-1}y_{t-1}+v_t|Y_{t-1}) \\ &= \mu'_{t-1}z_t, && (21) \\ E^*_{t-1}y_{t+1} &= E(a_t+b_ty_t+v_{t+1}|Y_{t-1}) \\ &= E(\phi'_tz_{t+1}|Y_{t-1}). && (22) \end{aligned}$$

Notably, both y_t and b_t are random variables. We have to compute $E(\phi'_tz_{t+1}|Y_{t-1})$. We can write joint distribution of ϕ and y as

$$\begin{aligned} f(\phi_t,y_t|Y_{t-1}) &= \underbrace{f(\phi_t|Y_t)}_{\text{Posterior beliefs}} \cdot \underbrace{f(y_t|Y_{t-1})}_{\text{Posterior prediction}} && (23) \\ &= \mathcal{N}_\phi(\mu_t,\Sigma_t)\mathcal{N}_y(\mu'_{t-1}z_t,\nu^2+z'_t\Sigma_{t-1}z_t). && (24) \end{aligned}$$

To see the second term of (24), we write the distibution of y_{t+1} conditional on Y_t as

$$\begin{aligned} f(y_{t+1}|Y_t) &= \int f(y_{t+1}|Y_t,\phi_t)f(\phi_t|Y_t)\,d\phi_t && (25) \\ &= \int \mathcal{N}_y(z'_{t+1}\phi_t,\nu^2)\mathcal{N}_\phi(\mu_t,\Sigma_t)\,d\phi_t && (26) \\ &= \mathcal{N}_y(\mu'_tz_{t+1},\nu^2+z'_{t+1}\Sigma_tz_{t+1}), && (27) \end{aligned}$$

so that $f\,(y_t|Y_{t-1})$ is as given in (24).

The density function can be written as

$$f(\phi_t) = f\begin{pmatrix} a_t \\ b_t \end{pmatrix} = \mathcal{N}\left(\begin{bmatrix} \mu_{a,t} \\ \mu_{b,t} \end{bmatrix}, \begin{bmatrix} \sigma^2_{a,t} & \sigma_{ab,t} \\ \sigma_{ab,t} & \sigma^2_{b,t} \end{bmatrix}\right). \quad (28)$$

Also, using (18),

$$\begin{aligned} \mu_t &= \begin{pmatrix} \mu_{a,t} \\ \mu_{b,t} \end{pmatrix} = \mu_{t-1}+\Sigma_t\nu^{-2}z_t\,(y_t-z'_t\mu_t) && (29) \\ &= \begin{pmatrix} \mu_{a,t-1} \\ \mu_{b,t-1} \end{pmatrix} + \begin{pmatrix} \sigma^2_{a,t} & \sigma_{ab,t} \\ \sigma_{ab,t} & \sigma^2_{b,t} \end{pmatrix}\nu^{-2}\begin{pmatrix} 1 \\ y_{t-1} \end{pmatrix}(y_t-a_{t-1}-b_{t-1}y_{t-1}) && (30) \\ \mu_{a,t} &= \mu_{a,t-1}+\underbrace{\nu^{-2}(\sigma^2_{a,t}+\sigma_{ab,t}y_{t-1})}_{\Sigma_{a,t}}(y_t-a_{t-1}-b_{t-1}y_{t-1}) && (31) \\ \mu_{b,t} &= \mu_{b,t-1}+\underbrace{\nu^{-2}(\sigma_{ab,t}+\sigma^2_{b,t}y_{t-1})}_{\Sigma_{b,t}}(y_t-a_{t-1}-b_{t-1}y_{t-1}). && (32) \end{aligned}$$

We can write

$$
\begin{aligned}
f(b_t, y_t|Y_{t-1}) &= f(b_t|y_t, Y_{t-1})f(y_t|Y_{t-1}) && (33)\\
&= \mathcal{N}_b(\mu_{b,t}, \sigma^2_{b,t})\mathcal{N}_y(\mu'_{t-1}z_t, \nu^2 + z'_t\Sigma_{t-1}z_t) && (34)
\end{aligned}
$$

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:

$$
\begin{aligned}
E(\phi'_t z_{t+1}|Y_{t-1}) &= E\left((a_t, b_t)\begin{pmatrix}1\\ y_t\end{pmatrix}|Y_{t-1}\right) && (35)\\
&= E(a_t + b_t y_t|Y_{t-1}) && (36)\\
&= E(a_t|Y_{t-1}) + E(b_t y_t|Y_{t-1}). && (37)
\end{aligned}
$$

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

$$
\begin{aligned}
E(b_t y_t|Y_{t-1}) &= \int\int b_t y_t f(b_t, y_t|Y_{t-1})\, dy_t\, db_t && (38)\\
&= \int\int b_t y_t \mathcal{N}_{b_t}(\mu_{b,t}, \sigma^2_{b,t})\mathcal{N}_{y_t}(\mu'_{t-1}z_t, \nu^2 + z'_t\Sigma_{t-1}z_t) db_t dy_t && (39)
\end{aligned}
$$

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

$$
\begin{aligned}
E(b_t y_t|Y_{t-1}) &= \int y_t \mathcal{N}_{y_t}(\mu'_{t-1}z_t, \underbrace{\nu^2 + z'_t\Sigma_{t-1}z_t}_{\Omega_{y_t}}) \underbrace{\int b_t \mathcal{N}_{b_t}(\mu_{b,t}, \sigma^2_{b,t}) db_t}_{Eb_t = \mu_{b,t}}\, dy_t && (40)\\
&= \int \mu_{b,t}\, y_t \mathcal{N}_{y_t}(\mu'_{t-1}z_t, \Omega_{y_t})\, dy_t && (41)\\
&= \int \big(\mu_{b,t-1} + \Sigma_{b,t}(y_t - a_{t-1} - b_{t-1}y_{t-1})\big)\, y_t \mathcal{N}_{y_t}(\mu'_{t-1}z_t, \Omega_{y_t}) dy_t && (42)\\
&= \big(\mu_{b,t-1} - \Sigma_{b,t}(a_{t-1} + b_{t-1}y_{t-1})\big) \underbrace{\int y_t \mathcal{N}_{y_t}(\mu'_{t-1}z_t, \Omega_{y_t})\, dy_t}_{Ey_t} && (43)\\
&\quad + \Sigma_{b,t} \underbrace{\int y_t^2 \mathcal{N}_{y_t}(\mu'_{t-1}z_t, \Omega_{y_t})\, dy_t}_{Ey_t^2 = Var(y_t) + (Ey_t)^2} && (44)\\
&= \big(\mu_{b,t-1} - \Sigma_{b,t}(a_{t-1} + b_{t-1}y_{t-1})\big) Ey_t + \Sigma_{b,t}Var(y_t) + \Sigma_{b,t}(Ey_t)^2 && (45)
\end{aligned}
$$

Therefore, we obtain

$$E(b_t y_t|Y_{t-1}) = \left(\mu_{b,t-1} + \Sigma_{b,t}(Ey_t - a_{t-1} - b_{t-1}y_{t-1})\right) Ey_t + \Sigma_{b,t}Var(y_t) \quad (46)$$
$$= E(\mu_{b,t}|Y_{t-1})E(y_t|Y_{t-1}) + \Sigma_{b,t}\Omega_{y_t}. \quad (47)$$

Then,

$$E(\phi_t' z_{t+1}|Y_{t-1}) = E(a_t|Y_{t-1}) + E(b_t y_t|Y_{t-1}) \quad (48)$$
$$= E(\mu_{a,t}|Y_{t-1}) + E(\mu_{b,t}|Y_{t-1})E(y_t|Y_{t-1}) + \Sigma_{b,t}\Omega_{y_t} \quad (49)$$
$$= E(\mu_t|Y_{t-1})'E(z_{t+1}|Y_{t-1}) + \Sigma_{b,t}\Omega_{y_t} \quad (50)$$
$$= \mu_{t-1}' \begin{pmatrix} 1 \\ \mu_{t-1}' z_t \end{pmatrix} + \Sigma_{b,t}\Omega_{y_t}. \quad (51)$$

Recall that

$$E_{t-1}^* y_t = E(a_{t-1} + b_{t-1}y_{t-1} + v_t|Y_{t-1}) = \mu_{t-1}' z_t. \quad (52)$$

Substituting these expressions into (1) under Bayesian learning we obtain the following expression:

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E_{t-1}^* y_t + \beta_1 E_{t-1}^* y_{t+1} + v_t, \quad (53)$$
$$= \alpha + \delta y_{t-1} + \beta_0 \mu_{t-1}' z_t + \beta_1 \mu_{t-1}' \begin{pmatrix} 1 \\ \mu_{t-1}' z_t \end{pmatrix} + \beta_1 \Sigma_{b,t}\Omega_{y_t} + v_t, \quad (54)$$
$$= \alpha + \delta y_{t-1} + \beta_0(\mu_{a,t-1} + \mu_{b,t-1}y_{t-1}) \quad (55)$$
$$+ \beta_1 \left(\mu_{a,t-1} + \mu_{b,t-1}(\mu_{a,t-1} + \mu_{b,t-1}y_{t-1})\right) + \beta_1 \Sigma_{b,t}\Omega_{y_t} + v_t. \quad (56)$$

Finally, rearranging this expression, we conclude that the actual law of motion under Bayesian learning can be written as

$$y_t = \left[\alpha + (\beta_0 + \beta_1)\mu_{a,t-1} + \beta_1 \mu_{a,t-1}\mu_{b,t-1}\right] + \left[\delta + \beta_0 \mu_{b,t-1} + \beta_1 \mu_{b,t-1}^2\right] y_{t-1} + \beta_1 \Sigma_{b,t}\Omega_{y_t} + v_t. \quad (57)$$

Except for the term $\Sigma_{b,t}\Omega_{y_t}$, the above expression is analogous to standard recursive least squares as analyzed by Evans and Honkapohja (2001) for the MSV solution with parameters being represented by their means.

2.4.1 Remarks

Here we can see 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 (57). Second, if $\delta = 0$, then there would be no term Ω_{y_t}, as the MSV solution (2) would not depend on y_{t-1} so that agents would just be estimating means.

To get back 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 estimates as a constant. There are really two levels to Bayesian learning. One is that the agents use the Bayesian estimator μ_a and μ_b, and the second is that the agents treat the estimates as random variables, which gives rise to the term $\beta_1 \Sigma_{b,t} \Omega_{y_t}$.

2.5 Alternative expressions for the ALM

In order to work with the expression (57), we can write it in an expanded fashion. First, consider $\Sigma_{b,t} \Omega_{y_t}$:

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

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

Then,

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

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

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

with $X = (0 \quad 1)$. Theferore,

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

as

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

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

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

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

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

Putting this expression in the ALM yields,

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

Using this alternative expression for the actual law of motion allows us to define a T-map in a convenient way.

3 Expectational stability

3.1 General case

Agents have beliefs about 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

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

where μ_t and Σ_t have the recursive form

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

where y_t in the first equation is given by ALM above. The evolution of the mean of the distribution is given by

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

Define a T-map

$$\begin{aligned}T_a(\mu, \Sigma) &= \alpha + (\beta_0 + \beta_1)\mu_a + \beta_1 \mu_a \mu_b + \beta_1 \sigma_{ab} \\ T_b(\mu, \Sigma) &= \delta + \beta_0 \mu_b + \beta_1 \mu_b^2 + \beta_1 \sigma_b^2.\end{aligned}$$

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

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

and defining $S_{t-1} = R_t$, we can represent the problem in the stochastic recursive form.

$$\begin{aligned}&\mu_t = \mu_{t-1} + t^{-1}\nu^{-2} S_{t-1}^{-1} z_t (z_t'(T(\mu_{t-1}, S_{t-2}) - \mu_{t-1}) - v_t), \\ &S_t = S_{t-1} + t^{-1}(\nu^{-2} z_{t+1} z_{t+1}' - S_{t-1}) \\ &\quad + t^{-2}\left(-\frac{t}{t+1}\right)(\nu^{-2} z_{t+1} z_{t+1}' - S_{t-1}).\end{aligned}$$

See Evans and Honkapohja (2001, Section 8.4) for technical conditions on the recursive stochastic algorithm.

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

$$\begin{aligned}\frac{d\mu}{d\tau} &= h(\mu) = \lim_{t\to\infty} E\mathcal{H} \\ &= \lim_{t\to\infty} E\nu^{-2}S_{t-1}^{-1}z_t(z_t'(T(\mu,S)-\mu)-v_t) \\ &= \widetilde{T}(\mu)-\mu.\end{aligned}$$

as

$$\lim_{t\to\infty} T(\mu,S) = \widetilde{T}(\mu)$$

with

$$\begin{aligned}\widetilde{T}_a(\mu) &= \alpha + (\beta_0+\beta_1)\mu_a + \beta_1\mu_a\mu_b \\ \widetilde{T}_b(\mu) &= \delta + \beta_0\mu_b + \beta_1\mu_b^2.\end{aligned}$$

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

$$\begin{aligned}\beta_0+\beta_1+\beta_1\mu_b-1 &< 0 \\ \beta_0-1+2\beta_1\mu_b &< 0\end{aligned} \tag{58}$$

We conclude that we have the same E-stability conditions as with classical recursive learning.

Intuition for this result. The variance term vanishes as data accumulates, and the estimators converge to their means, which is the same as what the recursive least squares learner assumes at the outset.

Transition paths will be different.

4 Dynamics

4.1 Approach and parameterization

To illustrate above findings we conduct numerical simulations. We consider again the model

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E_{t-1} y_t + \beta_1 E_{t-1} y_{t+1} + v_t, \tag{1}$$

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 (58), 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.

4.2 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. It is clear that parameters estimated with both recursive least

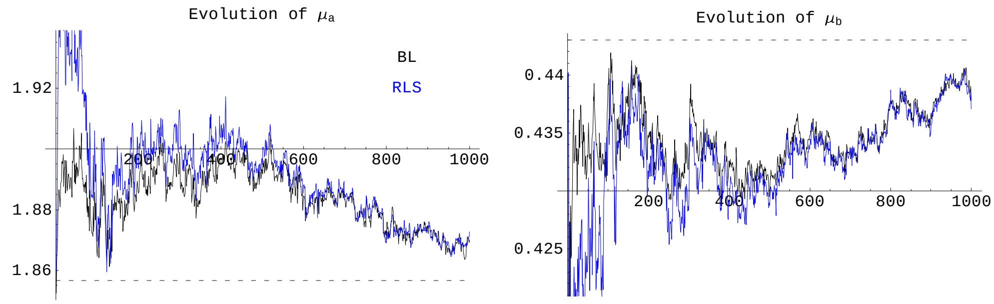

Figure 1: Bayesian learning vs RLS.

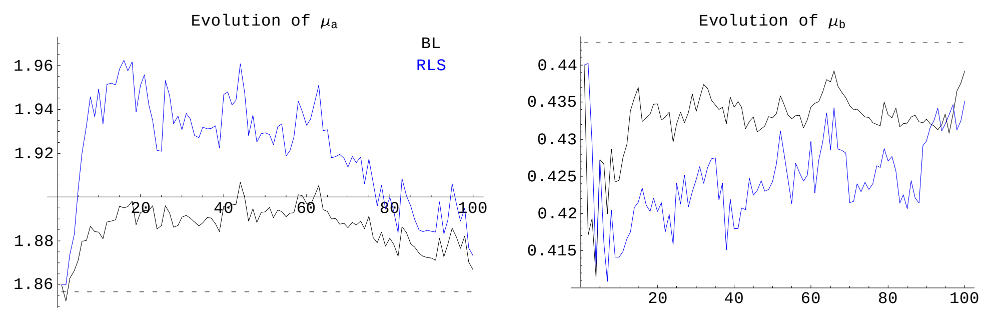

Figure 2: Bayesian learning vs RLS, first 100 periods

squares and Bayesian learning converge to rational expectations equilibrium.[4] However, it is also evident that the dynamic paths of a_t and b_t differ—but 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.

[4]The relatively slow convergence is typical result for learning of AR(1) processes. See the discussion in Evans and Honkapohja (2001).

4.3 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 μ_a and μ_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 (57). In order to distinguish between these two versions we can compare recursive least squares and Bayesian learning to the third case, passive Bayesian estimation.

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.[5] 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 (57).

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 trajectories are that 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.

4.4 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.[6] The priors here are always centered at rational expectations values. 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

[5] In the case of non-informative priors PBE and RLS are the same.

[6] Since the variance is equal to inverse of precision, $x = 4$ indicates variance of prior beliefs, σ, equal 1/4.

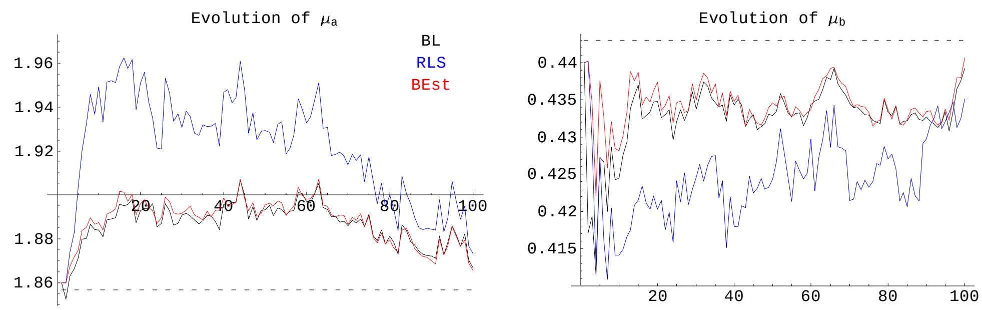

Figure 3: Bayesian learning and Bayesian estimation.

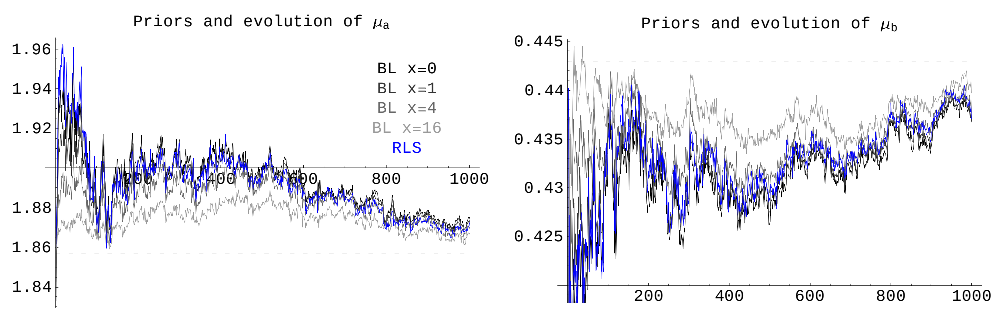

Figure 4: Effects of precision of priors.

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

5 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).

Expectational stability is not affected.

Bayesian systems will have some differences with RLS systems in the learning dynamics.

This is Bayesian updating, not really Bayesian decision making.

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. "title" *International Economic Review*, forthcoming.

[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] Kiefer, N., and Y. Nyarko. 1989. Optimal control of an unknown linear process with learning. *International Economic Review* 30(3): 571-586.

[7] McGough, B. 2003. Statistical learning with time-varying parameters. *Macroeconomic Dynamics* 7: 119-139.

[8] Wieland, V. 2000. Monetary policy, parameter uncertainty and optimal learning. *Journal of Monetary Economics* 46(1): 199-228.