Journal of Economic Dynamics & Control 62 (2016) 1–16

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

The stability of macroeconomic systems with Bayesian learners

James Bullard [a], Jacek Suda [b,*]

[a] Federal Reserve Bank of St. Louis, United States
[b] Narodowy Bank Polski, Poland

ARTICLE INFO

Article history:
Received 13 July 2015
Accepted 3 September 2015
Available online 8 October 2015

JEL classification:
D84
E00
D83

Keywords:
Expectational stability
Recursive learning
Learnability of rational expectations equilibrium
Bayesian learning

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.

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.

* Corresponding author.
E-mail addresses: bullard@stls.frb.org (J. Bullard), jacek.suda@nbp.pl (J. Suda).

http://dx.doi.org/10.1016/j.jedc.2015.09.004
0165-1889/

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 these issues 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. We think of this as a step toward rationality. However, the Bayesian econometricians in our model may still be viewed as boundedly rational rather than fully rational, because, while agents acknowledge that their beliefs will change over time they do not have an explicit model of this process. That is, they have a Bayesian perspective but they do not model their own learning behavior. In this sense we see our paper as one step closer toward rationality relative to the classical recursive learning case.

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.[1] They thought of this as a misspecification, a form of bounded rationality. They asked whether convergence to

[1] This is the same concern raised by Cogley and Sargent (2008).

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 (2008) study a Lucas tree 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.[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 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).

Wieland (2000a) studies optimal decision making problem under parameter uncertainty and Bayesian learning. He considers a Bayesian decision-maker controlling a linear stochastic process with constant unknown parameters.[4] He studies the value of optimal experimentation and shows numerically that myopic behavior (no experimentation) may result in mistaken beliefs about unknown parameters and as a result, in a bias in actions and outcomes. We consider a reduced form model and do not study the joint optimal decision and optimal estimation problem.

Guidolin and Timmermann (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 et al. (2010) 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 (2000b) adapts the framework of Kiefer and Nyarko (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

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

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 work with this general form in order to keep results comparable to Evans and Honkapohja (2001).[5]

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

[3] See the quote above.

[4] Beck and Wieland (2002) consider an optimal Bayesian learning and control problem with time-varying unknown parameters.

[5] Taylor's (1980) overlapping contract model is an example of a model that has this very form. A simple New-Keynesian model based on Woodford (2003) also fits this specification provided one assumes all expectations are taken using information available as of time $t-1$ rather than t. Rotemberg and Woodford (1997) provide one example of a New Keynesian model with an informational assumption of this type.

We have chosen this particular version of Evans and Honkapohja (2001), Eq. (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

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

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

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

and

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

We stress that there may be two solutions $\overline{b}$ which solve these equations. We assign a traditional perceived law of motion (PLM), which is consistent in form with the MSV solution (2),

$$y_t=a+by_{t-1}+v_t=\phi' z_t+v_t, \tag{5}$$

where $\phi=[a,b]'$ and $z_t=[1,y_{t-1}]'$. The agents use the PLM to form expectations, which then can be substituted into Eq. (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

$$E^*_{t-1}y_t=\widehat{a}_{t-1}+\widehat{b}_{t-1}y_{t-1}, \tag{6}$$

$$E^*_{t-1}y_{t+1}=E(\widehat{a}_t+\widehat{b}_t y_t+v_{t+1}|Y_{t-1})=\widehat{a}_{t-1}(1+\widehat{b}_{t-1})+\widehat{b}^2_{t-1}y_{t-1}, \tag{7}$$

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

$$y_t=[\alpha+(\beta_0+\beta_1)\widehat{a}_{t-1}+\beta_1\widehat{a}_{t-1}\widehat{b}_{t-1}]+[\delta+\beta_0\widehat{b}_{t-1}+\beta_1\widehat{b}^2_{t-1}]y_{t-1}+v_t. \tag{8}$$

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, Eq. (8), in the case of Bayesian learning in order to compare the two.

3. Real time Bayesian learning

3.1. Priors and posteriors

We now wish to assume instead 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 in Eq. (5).[6] They have priors which are given by

$$\phi_0'=(a_0,b_0)\sim\mathcal{N}(\mu_0,\Sigma_0), \tag{9}$$

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

$$\Sigma_0=\begin{bmatrix}\sigma^2_{a,0} & \sigma_{ab,0}\\ \sigma_{ab,0} & \sigma^2_{b,0}\end{bmatrix}, \tag{10}$$

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

$$y_t|Y_{t-1},\ \phi\sim\mathcal{N}(a+by_{t-1},\nu^2), \tag{11}$$

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

$$f(Y_t|\phi)=f(y_t|\phi,Y_{t-1})f(Y_{t-1}|\phi)=f(y_t|\phi,Y_{t-1})f(y_{t-1}|\phi,Y_{t-2})\ldots f(y_2|\phi,y_1)f(y_1|\phi). \tag{12}$$

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

$$f(\phi|Y_t)\propto f(Y_t|\phi)f(\phi), \tag{13}$$

that is, the posterior distribution is proportional to the distribution of Y_t times the prior distribution. Assuming $f(y_1|\phi)$ is

[6] We follow Evans and Honkapohja (2001) and assume that agents know the variance of the shocks, ν^2.

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

$$f(\phi|Y_t)=\mathcal{N}(\mu_t,\Sigma_t), \tag{14}$$

where

$$\mu_t=\Sigma_t\left(\Sigma_0^{-1}\mu_0+\nu^{-2}(Z_t'Y_t)\right), \tag{15}$$

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.[7]

Since we will consider recursive updating it is useful to express μ_t and Σ_t in a recursive form.

Lemma 1. *μ_t and Σ_t can be written in a recursive form as*

$$\mu_t=\mu_{t-1}+\Sigma_t\nu^{-2}z_t\left(y_t-z_t'\mu_{t-1}\right), \tag{17}$$

$$\Sigma_t^{-1}=\Sigma_{t-1}^{-1}+\nu^{-2}z_tz_t'. \tag{18}$$

Proof. See Appendix A.

We note that since $\Sigma_t=(\nu^2/t)R_t^{-1}$ the functional form of the recursions above are identical to those under recursive least squares:

$$\phi_t^{RLS}=\phi_{t-1}^{RLS}+t^{-1}R_t^{-1}z_t(y_t-\phi_{t-1}^{RLS\prime}z_t) \tag{19}$$

$$R_t=R_{t-1}+t^{-1}(z_tz_t'-R_{t-1}). \tag{20}$$

However, the values of the estimates will differ except for the case where Σ_0^{-1} is a null matrix—that is, the prior is diffuse.

3.2. 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 same PLM under learning in the standard analysis described above in Eq. (5), namely

$$y_t=a+by_{t-1}+v_t=\phi'z_t+v_t. \tag{21}$$

We now take expectations based on this PLM in order to substitute these into (1) to obtain the ALM. The necessary expectation terms are given by

$$E_{t-1}y_t=E(a+by_{t-1}+v_t|Y_{t-1})=\mu_{t-1}'z_t, \tag{22}$$

$$E_{t-1}y_{t+1}=E(a+by_t+v_{t+1}|Y_{t-1})=E(\phi'z_{t+1}|Y_{t-1}). \tag{23}$$

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

We first compute the joint distribution of ϕ and y. The density function of ϕ conditional on Y_t can be written as

$$f(\phi|Y_t)=f\left(\begin{matrix}a\\b\end{matrix}\middle|Y_t\right)=\mathcal{N}\left(\begin{bmatrix}\mu_{a,t}\\\mu_{b,t}\end{bmatrix},\begin{bmatrix}\sigma_{a,t}^2&\sigma_{ab,t}\\\sigma_{ab,t}&\sigma_{b,t}^2\end{bmatrix}\right). \tag{24}$$

Lemma 2. *The joint distribution of ϕ and y is given by*

$$f(\phi,y_t|Y_{t-1})=\mathcal{N}_\phi(\mu_t,\Sigma_t)\mathcal{N}_y(\mu_{t-1}'z_t,\nu^2+z_t'\Sigma_{t-1}z_t). \tag{25}$$

and, in particular, the joint distribution of b and y equals

$$f(b,y_t|Y_{t-1})=\mathcal{N}_b(\mu_{b,t},\sigma_{b,t}^2)\mathcal{N}_y(\mu_{t-1}'z_t,\nu^2+z_t'\Sigma_{t-1}z_t), \tag{26}$$

where

$$\mu_{b,t}=\mu_{b,t-1}+\nu^{-2}(\sigma_{ab,t}+\sigma_{b,t}^2y_{t-1})(y_t-\mu_{a,t-1}-\mu_{b,t-1}y_{t-1}). \tag{27}$$

Proof. See Appendix A.

Now we can find $E(by_t|Y_{t-1})$ and $E(\phi'z_{t+1}|Y_{t-1})$. As we have a joint distribution of both random variables we can compute these expectations directly.

[7] See Zellner (1971) for an introduction to Bayesian inference.

Lemma 3. *The expectations $E(by_t|Y_{t-1})$ and $E(\phi' z_{t+1}|Y_{t-1})$ are given by*

$$E(by_t|Y_{t-1}) = \mu_{b,t-1}E(y_t|Y_{t-1}) + \Sigma_{b,t}\Omega_{y_t} \tag{28}$$

and

$$E(\phi' z_{t+1}|Y_{t-1}) = \mu'_{t-1}\begin{pmatrix} 1 \\ \mu'_{t-1}z_t \end{pmatrix} + \Sigma_{b,t}\Omega_{y_t}, \tag{29}$$

where

$$\Sigma_{b,t} = \nu^{-2}(\sigma_{ab,t} + \sigma^2_{b,t}y_{t-1}) \tag{30}$$

$$\Omega_{y_t} = \nu^2 + z'_t\Sigma_{t-1}z_t. \tag{31}$$

Proof. See Appendix A.

We are now ready to present our key equation, the actual law of motion under Bayesian learning. Recall that our general linear model from Eq. (1) is given by

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E^*_{t-1}y_t + \beta_1 E^*_{t-1}y_{t+1} + v_t. \tag{32}$$

Substituting Eqs. (22) and (29) into this equation we can write the actual law of motion.

Proposition 4. *The actual law of motion under Bayesian learning can be written as*

$$y_t = \left[\alpha + (\beta_0 + \beta_1)\mu_{a,t-1} + \beta_1\mu_{a,t-1}\mu_{b,t-1}\right] + \left[\delta + \beta_0\mu_{b,t-1} + \beta_1\mu^2_{b,t-1}\right]y_{t-1} + \beta_1\Sigma_{b,t}\Omega_{y_t} + v_t. \tag{33}$$

Proof. See Appendix A.

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 Eq. (8) 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.

3.3. Remarks on the Bayesian ALM

We said that we chose Eq. (1) carefully. In particular, we made sure that a lagged endogenous variable was included with a non-zero coefficient δ, and that a two-step ahead expectation was included with a non-zero coefficient β_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 (33). 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}, 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 as we are describing it here. One is that the agents use the Bayesian estimators μ_a and μ_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 (33). 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 μ_a and μ_b are not treated as constants. Similarly, even if agents treat parameters in PLM as constants but still use Bayesian estimator to form their expectations—a passive Bayesian case—the actual law of motion will also differ from the RLS case.

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 $\{y_\tau\}^t_{\tau=1} = Y_t$, their beliefs are given by

$$f(\phi|Y_t) = \mathcal{N}(\mu_t, \Sigma_t), \tag{34}$$

where μ_t and Σ_t have the recursive form

$$\mu_t = \mu_{t-1} + \Sigma_t\nu^{-2}z_t(y_t - z'_t\mu_{t-1}), \tag{35}$$

$$\Sigma^{-1}_t = \Sigma^{-1}_{t-1} + \nu^{-2}z_tz'_t, \tag{36}$$

where y_t in the first equation is given by the actual law of motion in Eq. (33) above.

In order to work with the expression (33), we can write it in an expanded fashion.

Lemma 5. *The actual law motion (33) can be written as*

$$y_t = \left[\alpha + (\beta_0+\beta_1)\mu_{a,t-1} + \beta_1\mu_{a,t-1}\mu_{b,t-1} + \beta_1\sigma_{ab,t-1}\right] + \left[\delta + \beta_0\mu_{b,t-1} + \beta_1\mu_{b,t-1}^2 + \beta_1\sigma_{b,t-1}^2\right] y_{t-1} + v_t. \tag{37}$$

Proof. See Appendix A.

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.

The evolution of the mean of the distribution is given by

$$\mu_t = \mu_{t-1} + \Sigma_t \nu^{-2} z_t(\alpha + (\beta_0+\beta_1)\mu_{a,t-1} + \beta_1\mu_{a,t-1}\mu_{b,t-1} + \beta_1\sigma_{ab,t-1} + \left[\delta + \beta_0\mu_{b,t-1} + \beta_1\mu_{b,t-1}^2\right] y_{t-1} + v_t - z_t'\mu_{t-1}). \tag{38}$$

We can now state our main result.

Theorem 6. *Consider a macroeconomic system given by Eq. (1) with Bayesian learners. Assume agents have PLM given by* Eq. (5) *and they update their beliefs about the parameters in their PLM using Bayes rule according to Eqs.* (17) *and* (18)*. Then, this system has the same E-stability conditions as under classical recursive learning.*

Proof. Following Evans and Honkapohja (2001), we define a T-map as

$$T_a(\mu, \Sigma) = \alpha + (\beta_0+\beta_1)\mu_a + \beta_1\mu_a\mu_b + \beta_1\sigma_{ab} \tag{39}$$

$$T_b(\mu, \Sigma) = \delta + \beta_0\mu_b + \beta_1\mu_b^2 + \beta_1\sigma_b^2. \tag{40}$$

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

$$R_t = (1/t)\Sigma_0^{-1} + (1/t)\nu^{-2} Z_t' Z_t, \tag{41}$$

and defining $S_{t-1} = R_t$, we can represent the problem in the stochastic recursive form,[8]

$$\mu_t = \mu_{t-1} + t^{-1}\nu^{-2} S_{t-1}^{-1} z_t(z_t'(T(\mu_{t-1}, S_{t-2}) - \mu_{t-1}) - v_t), \tag{42}$$

$$S_t = S_{t-1} + t^{-1}(\nu^{-2} z_{t+1} z_{t+1}' - S_{t-1}) + t^{-2}\left(-\frac{t}{t+1}\right)(\nu^{-2} z_{t+1} z_{t+1}' - S_{t-1}). \tag{43}$$

The T-map is now defined in terms of μ and S as

$$T_a(\mu, S) = \alpha + (\beta_0+\beta_1)\mu_a + \beta_1\mu_a\mu_b + \beta_1\frac{s_{ab}}{t} \tag{44}$$

$$T_b(\mu, S) = \delta + \beta_0\mu_b + \beta_1\mu_b^2 + \beta_1\frac{s_b}{t}, \tag{45}$$

where s_{ab} and s_b are the corresponding entries in the S^{-1} matrix.

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

$$\frac{d\mu}{d\tau} = h(\mu) = \lim_{t\to\infty} E\nu^{-2} S^{-1} z_t(z_t'(T(\mu, S) - \mu) - v_t). \tag{46}$$

Using regularity condition $Ez_t z_t' \equiv M_z < \infty$ and $Ez_t v_t = 0$, we can rewrite this as

$$\frac{d\mu}{d\tau} = \nu^{-2} S^{-1} M_z\left(\lim_{t\to\infty} T(\mu, S) - \mu\right). \tag{47}$$

From the definition of the T-map and using $\lim_{t\to\infty} s_{ab}/t = \lim_{t\to\infty} s_b/t = 0$, we can define $\tilde{T}(\mu)$ as

$$\lim_{t\to\infty} T(\mu, S) = \tilde{T}(\mu) \tag{48}$$

with

$$\tilde{T}_a(\mu) = \alpha + (\beta_0+\beta_1)\mu_a + \beta_1\mu_a\mu_b \tag{49}$$

$$\tilde{T}_b(\mu) = \delta + \beta_0\mu_b + \beta_1\mu_b^2. \tag{50}$$

Therefore,

$$\frac{d\mu}{d\tau} = \nu^{-2} S^{-1} M_z\left(\tilde{T}(\mu) - \mu\right), \tag{51}$$

and the stability of the system is governed by the following equation:

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

$$\frac{d\mu}{d\tau} = \tilde{T}(\mu) - \mu. \tag{52}$$

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

$$\beta_0 + \beta_1 + \beta_1\mu_b - 1 < 0 \tag{53}$$

$$\beta_0 - 1 + 2\beta_1\mu_b < 0 \tag{54}$$

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.□

Theorem 6 presents 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.[9]

Even though asymptotic 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

$$y_t = \alpha + \delta y_{t-1} + \beta_0 E^*_{t-1} y_t + \beta_1 E^*_{t-1} y_{t+1} + v_t, \tag{55}$$

with parameter values $\alpha = 2$, $\delta = 0.3$, $\beta_0 = 0.5$, and $\beta_1 = -0.4$. The two $AR(1)$ MSV solutions are $(\overline{a}_1, \overline{b}_1) = (1.86, 0.44)$ and $(\overline{a}_2, \overline{b}_2) = (8.97, -1.69)$. We assume $v_t \sim \mathcal{N}(0, 0.5)$. Clearly, only the first solution is stationary and, in accordance with (53), 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 $(\overline{a}_1, \overline{b}_1)$. The lagged value of y is equal to unconditional mean of y. For each parameterization, we conduct 1000 simulations and report the mean 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. Figs. 1 and 2 illustrate this point. In these figures, 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.[10] However, it is also evident that the dynamic paths of a_t and b_t differ—and that these differences decrease over time. Figs. 3 and 4 depict

[9] The expectational stability of a system with passive Bayesian learners is governed by the same conditions.

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

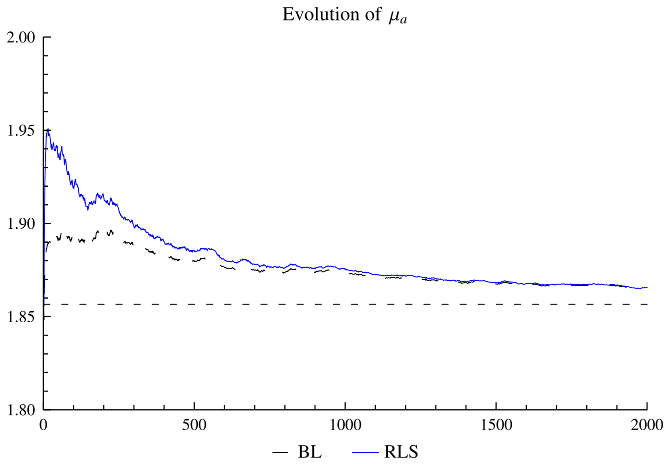

Fig. 1. Bayesian learning versus RLS—μ_a.

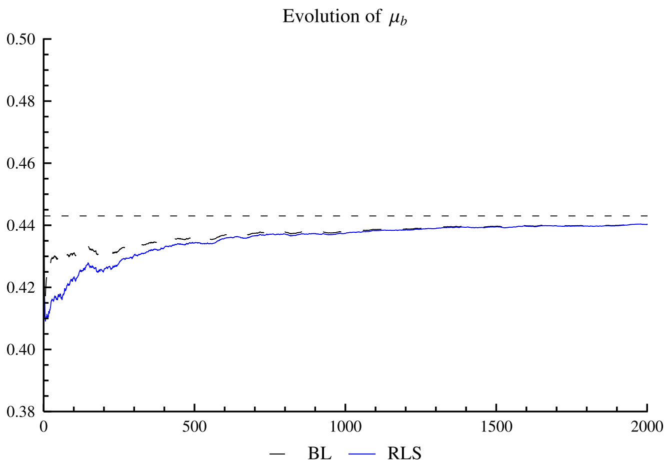

Fig. 2. Bayesian learning versus RLS—μ_b.

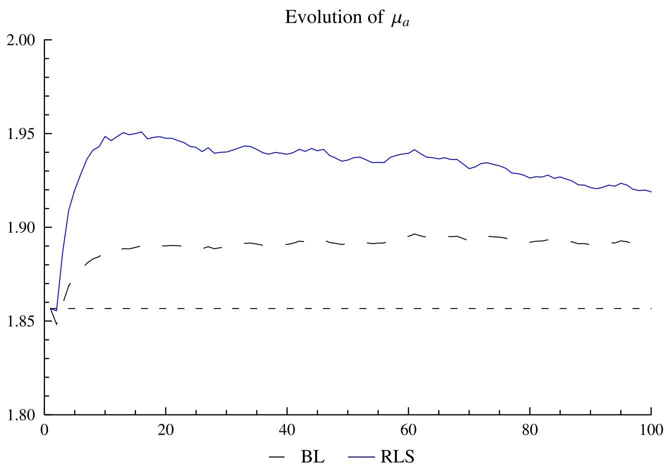

Fig. 3. Bayesian learning versus RLS, first 100 periods—μ_a.

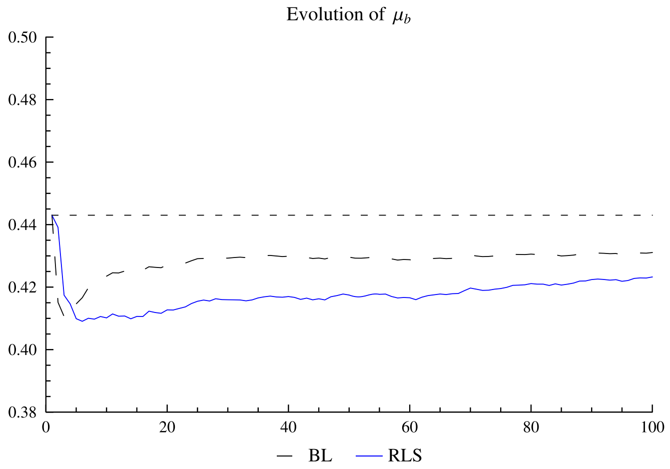

Fig. 4. Bayesian learning versus RLS, first 100 periods—μ_b.

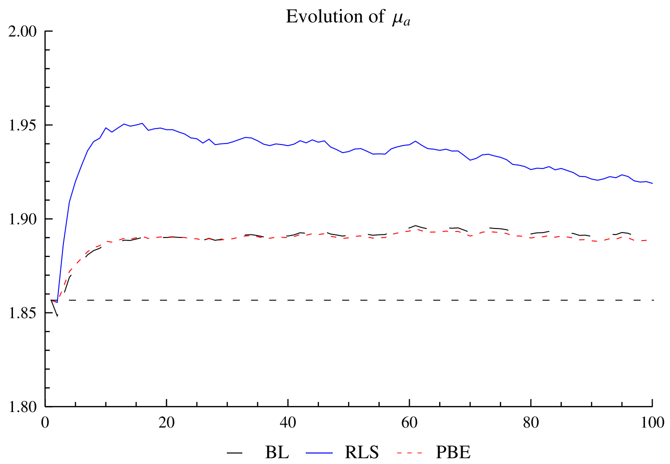

Fig. 5. Bayesian learning and Bayesian estimation—μ_a.

first 100 periods from the same simulation. In these figures 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 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 the actual law of motion (33). 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 Figs. 5 and 6, 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.[11] 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 (33).

The striking feature of Figs. 5 and 6 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.

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

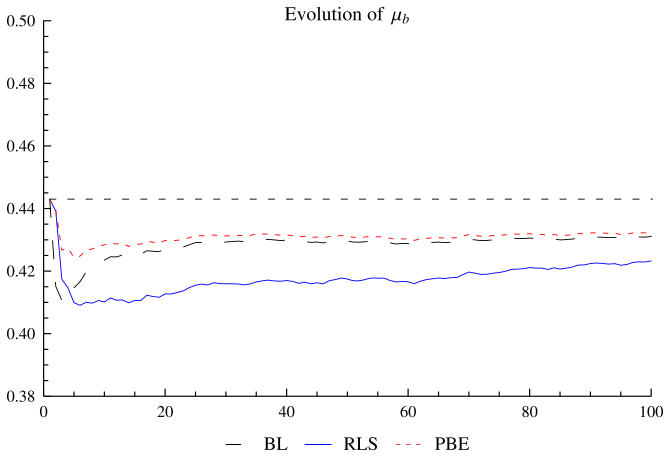

Fig. 6. Bayesian learning and Bayesian estimation—μ_b.

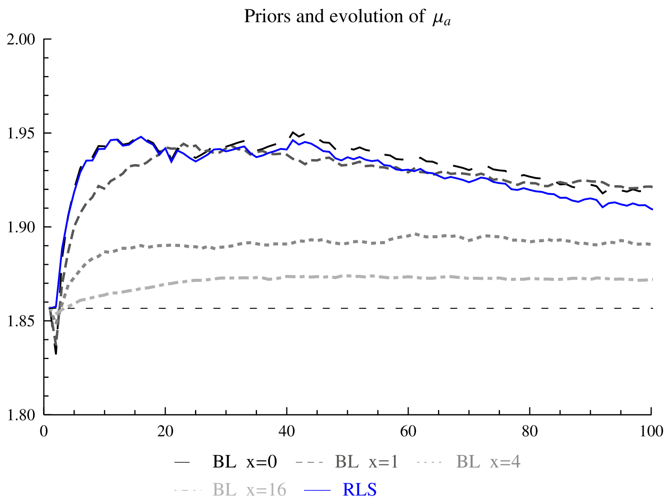

Fig. 7. Effects of precision on priors—μ_a.

5.4. *Effects of priors and variance*

Bayesians have priors that may differ from an uninformative state, while standard recursive least squares does not. As Figs. 7 and 8 illustrate, the precision of prior beliefs can be relatively more important for transition paths. Figs. 7 and 8 depict alternative trajectories of a_t and b_t for different prior variances.[12] The priors here are always centered at rational expectations values.[13] 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 think this point is well understood and we do not illustrate it here.

Figs. 9 and 10 show the effects of a larger shock variance on the evolution of estimated parameters. The differences between parameter estimates are increasing in the variance in this example.

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

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

[13] This removes a degree of freedom from the simulations. Since expectational stability is a local concept, this seems reasonable.

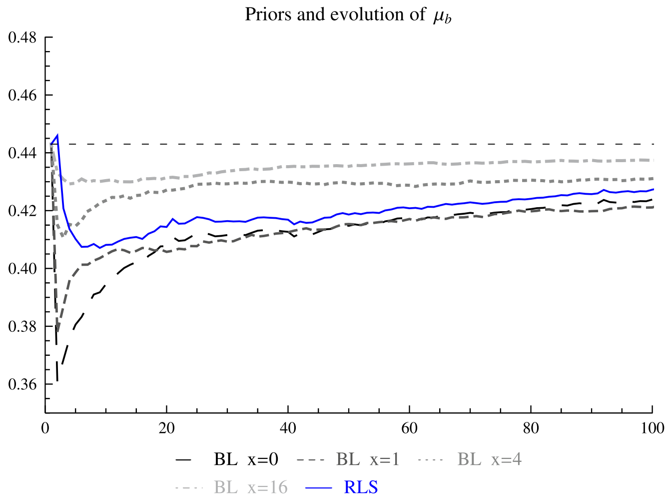

Fig. 8. Effects of precision on priors—μ_b.

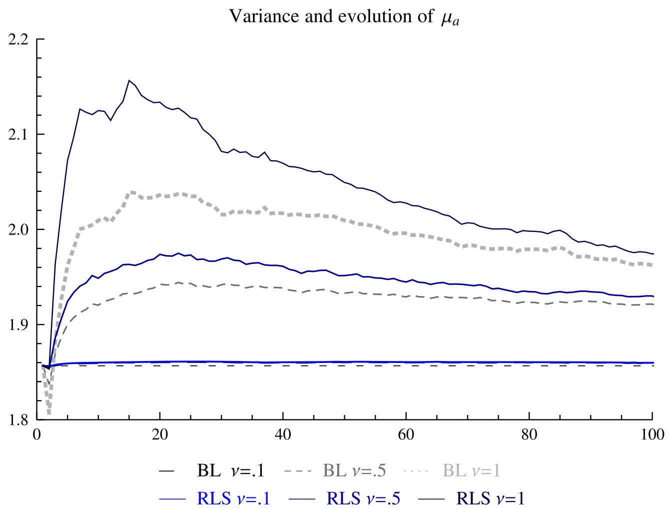

Fig. 9. Effects of increased variance in the shock—μ_a.

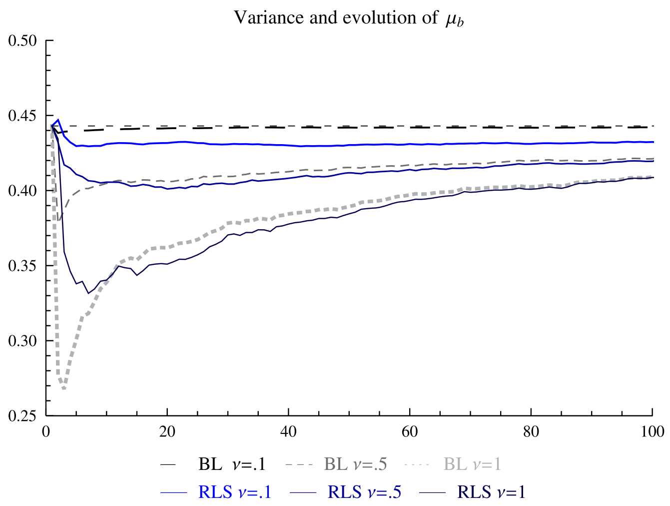

Fig. 10. Effects of increased variance in the shock—μ_b.

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 expectational 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.

Acknowledgments

We thank Sid Chib, James Morley, Yang Lu, anonymous referees, the editor, and the participants at the Learning Week, Banque de France and many other conferences and seminars for comments. Suda thanks the Olin Business School's Center for Research in Economics and Strategy (CRES) for funding.

Any views expressed are those of the authors and do not necessarily reflect the views of the Federal Open Market Committee or the Federal Reserve System.

Any views expressed are those of the authors and do not necessarily reflect the views of Narodowy Bank Polski.

Appendix A. Proofs

A.1. Proof of Lemma 1

Proof. Both μ_t and Σ_t can be written in a recursive form. For Σ_t, we can write

$$\Sigma_t^{-1} = \Sigma_0^{-1} + \nu^{-2}(Z_t' Z_t) = \Sigma_0^{-1} + \nu^{-2}\sum_{i=1}^{t} z_i z_i' = \Sigma_0^{-1} + \nu^{-2}\sum_{i=1}^{t-1} z_i z_i' + \nu^{-2} z_t z_t' = \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t'. \tag{56}$$

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

$$\begin{aligned}\mu_t &= \Sigma_t(\Sigma_{t-1}^{-1}\mu_{t-1} + \nu^{-2} z_t y_t),\\ &= \Sigma_t \Sigma_{t-1}^{-1}\mu_{t-1} + \Sigma_t \nu^{-2} z_t y_t,\end{aligned} \tag{57}$$

$$\begin{aligned}\mu_t - \mu_{t-1} &= (\Sigma_t \Sigma_{t-1}^{-1} - I)\mu_{t-1} + \Sigma_t \nu^{-2} z_t y_t,\\ \mu_t &= \mu_{t-1} + \Sigma_t\left((\Sigma_{t-1}^{-1} - \Sigma_t^{-1})\mu_{t-1} + \nu^{-2} z_t y_t\right),\end{aligned} \tag{58}$$

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

$$\mu_t = \mu_{t-1} + \Sigma_t\left(\nu^{-2} z_t y_t - \nu^{-2} z_t z_t' \mu_{t-1}\right) = \mu_{t-1} + \Sigma_t \nu^{-2} z_t\left(y_t - z_t'\mu_{t-1}\right). \square \tag{59}$$

A.2. Proof of Lemma 2

Proof. We can write joint distribution of ϕ and y as

$$f(\phi, y_t | Y_{t-1}) = \underbrace{f(\phi | Y_t)}_{\text{Posterior beliefs}} \cdot \underbrace{f(y_t | Y_{t-1})}_{\text{Posterior prediction}} \tag{60}$$

$$= \mathcal{N}_\phi(\mu_t, \Sigma_t)\mathcal{N}_y(\mu_{t-1}' z_t, \nu^2 + z_t' \Sigma_{t-1} z_t). \tag{61}$$

To see the second term of (61), we write the distribution of y_{t+1} conditional on Y_t as

$$f(y_{t+1}|Y_t) = \int f(y_{t+1}|Y_t,\phi)f(\phi|Y_t)\,d\phi = \int \mathcal{N}_y(\phi' z_{t+1},\nu^2)\mathcal{N}_\phi(\mu_t,\Sigma_t)\,d\phi = \mathcal{N}_y(\mu_t' z_{t+1},\nu^2 + z_{t+1}'\Sigma_t z_{t+1}), \tag{62}$$

so that $f(y_t|Y_{t-1})$ is as given in (61).

The density function can be written as

$$f(\phi|Y_t) = f\left(\begin{matrix} a \\ b \end{matrix}\middle| Y_t\right) = \mathcal{N}\left(\begin{bmatrix}\mu_{a,t} \\ \mu_{b,t}\end{bmatrix},\begin{bmatrix}\sigma_{a,t}^2 & \sigma_{ab,t} \\ \sigma_{ab,t} & \sigma_{b,t}^2\end{bmatrix}\right). \tag{63}$$

Also, using (17),

$$\mu_t = \begin{pmatrix}\mu_{a,t} \\ \mu_{b,t}\end{pmatrix} = \mu_{t-1} + \Sigma_t \nu^{-2} z_t (y_t - z_t'\mu_t) = \begin{pmatrix}\mu_{a,t-1} \\ \mu_{b,t-1}\end{pmatrix} + \begin{pmatrix}\sigma_{a,t}^2 & \sigma_{ab,t} \\ \sigma_{ab,t} & \sigma_{b,t}^2\end{pmatrix}\nu^{-2}\begin{pmatrix}1 \\ y_{t-1}\end{pmatrix}(y_t - \mu_{a,t-1} - \mu_{b,t-1}y_{t-1}) \tag{64}$$

$$\mu_{a,t} = \mu_{a,t-1} + \underbrace{\nu^{-2}(\sigma_{a,t}^2 + \sigma_{ab,t}y_{t-1})}_{\Sigma_{a,t}}(y_t - \mu_{a,t-1} - \mu_{b,t-1}y_{t-1}) \tag{65}$$

$$\mu_{b,t} = \mu_{b,t-1} + \underbrace{\nu^{-2}(\sigma_{ab,t} + \sigma_{b,t}^2 y_{t-1})}_{\Sigma_{b,t}}(y_t - \mu_{a,t-1} - \mu_{b,t-1}y_{t-1}). \tag{66}$$

We can write

$$f(b,y_t|Y_{t-1}) = f(b|y_t,Y_{t-1})f(y_t|Y_{t-1}) = \mathcal{N}_b(\mu_{b,t},\sigma_{b,t}^2)\mathcal{N}_y(\mu_{t-1}' z_t,\nu^2 + z_t'\Sigma_{t-1}z_t).\square \tag{67}$$

A.3. Proof of Lemma 3

Proof. Consider $E(by_t|Y_{t-1})$:

$$E(by_t|Y_{t-1}) = \int\int by_t f(b,y_t|Y_{t-1})\,dy_t\,db = \int\int by_t \mathcal{N}_b(\mu_{b,t},\sigma_{b,t}^2)\mathcal{N}_y(\mu_{t-1}' z_t,\nu^2 + z_t'\Sigma_{t-1}z_t)\,db\,dy_t. \tag{68}$$

As $\mathcal{N}_y$ does not depend on b we can write it as

$$\begin{aligned}
E(by_t|Y_{t-1}) &= \int y_t \mathcal{N}_y\left(\mu_{t-1}' z_t, \underbrace{\nu^2 + z_t'\Sigma_{t-1}z_t}_{\Omega_{y_t}}\right)\underbrace{\int b\mathcal{N}_b(\mu_{b,t},\sigma_{b,t}^2)\,db}_{E_t b = \mu_{b,t}}\,dy_t \\
&= \int \mu_{b,t}\, y_t \mathcal{N}_y(\mu_{t-1}' z_t,\Omega_{y_t})\,dy_t = \int (\mu_{b,t-1} + \Sigma_{b,t}(y_t - \mu_{a,t-1} - \mu_{b,t-1}y_{t-1}))y_t\mathcal{N}_y(\mu_{t-1}' z_t,\Omega_{y_t})\,dy_t \\
&= (\mu_{b,t-1} - \Sigma_{b,t}(\mu_{a,t-1} + \mu_{b,t-1}y_{t-1}))\underbrace{\int y_t\mathcal{N}_y(\mu_{t-1}' z_t,\Omega_{y_t})\,dy_t}_{E_{t-1}y_t} + \Sigma_{b,t}\underbrace{\int y_t^2\mathcal{N}_{y_t}(\mu_{t-1}' z_t,\Omega_{y_t})\,dy_t}_{E_{t-1}y_t^2 = Var_{t-1}(y_t) + (E_{t-1}y_t)^2} \\
&= (\mu_{b,t-1} - \Sigma_{b,t}(\mu_{a,t-1} + \mu_{b,t-1}y_{t-1}))E_{t-1}y_t + \Sigma_{b,t}Var_{t-1}(y_t) + \Sigma_{b,t}(E_{t-1}y_t)^2.
\end{aligned} \tag{69}$$

Recall that

$$E(y_t|Y_{t-1}) = \mu_{a,t-1} + \mu_{b,t-1}y_{t-1} = \mu_{t-1}' z_t.$$

Therefore, we obtain

$$E(by_t|Y_{t-1}) = (\mu_{b,t-1} + \Sigma_{b,t}(E(y_t|Y_{t-1}) - \mu_{a,t-1} - \mu_{b,t-1}y_{t-1}))E(y_t|Y_{t-1}) + \Sigma_{b,t}\,Var(y_t|Y_{t-1}) = \mu_{b,t-1}E(y_t|Y_{t-1}) + \Sigma_{b,t}\Omega_{y_t}. \tag{70}$$

Then,

$$E(\phi' z_{t+1}|Y_{t-1}) = E(a|Y_{t-1}) + E(by_t|Y_{t-1}) = \mu_{a,t-1} + \mu_{b,t-1}E(y_t|Y_{t-1}) + \Sigma_{b,t}\Omega_{y_t} = \mu_{t-1}'\begin{pmatrix}1 \\ \mu_{t-1}' z_t\end{pmatrix} + \Sigma_{b,t}\Omega_{y_t}. \tag{71}$$

A.4. Proof of Proposition 4

Proof. Substituting expressions (29) and (22) into (1) under Bayesian learning we obtain the following expression:

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

$$= \alpha + \delta y_{t-1} + \beta_0 \mu'_{t-1} z_t + \beta_1 \mu'_{t-1} \begin{pmatrix} 1 \\ \mu'_{t-1} z_t \end{pmatrix} + \beta_1 \Sigma_{b,t} \Omega_{y_t} + v_t,$$

$$= \alpha + \delta y_{t-1} + \beta_0 (\mu_{a,t-1} + \mu_{b,t-1} y_{t-1}) + \beta_1 \left(\mu_{a,t-1} + \mu_{b,t-1} (\mu_{a,t-1} + \mu_{b,t-1} y_{t-1}) \right) + \beta_1 \Sigma_{b,t} \Omega_{y_t} + v_t. \tag{72}$$

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

$$y_t = \left[\alpha + (\beta_0 + \beta_1) \mu_{a,t-1} + \beta_1 \mu_{a,t-1} \mu_{b,t-1} \right] + \left[\delta + \beta_0 \mu_{b,t-1} + \beta_1 \mu_{b,t-1}^2 \right] y_{t-1} + \beta_1 \Sigma_{b,t} \Omega_{y_t} + v_t. \tag{73}$$

A.5. *Proof of Lemma 5*

Proof. First, consider $\Sigma_{b,t} \Omega_{y_t}$:

$$\Sigma_t^{-1} = \Sigma_{t-1}^{-1} + \nu^{-2} z_t z_t' = \begin{pmatrix} \sigma_{a,t-1}^2 & \sigma_{ab,t-1} \\ \sigma_{ab,t-1} & \sigma_{b,t-1}^2 \end{pmatrix}^{-1} + \nu^{-2} \begin{pmatrix} 1 & y_{t-1} \\ y_{t-1} & y_{t-1}^2 \end{pmatrix} = \begin{pmatrix} \frac{\sigma_{b,t-1}^2}{A_{t-1}} + \nu^{-2} & -\frac{\sigma_{ab,t-1}}{A_{t-1}} + \nu^{-2} y_{t-1} \\ -\frac{\sigma_{ab,t-1}}{A_{t-1}} + \nu^{-2} y_{t-1} & \frac{\sigma_{a,t-1}^2}{A_{t-1}} + \nu^{-2} y_{t-1}^2 \end{pmatrix}, \tag{74}$$

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,

$$\Sigma_t = (\Sigma_t^{-1})^{-1} = \begin{pmatrix} \frac{\sigma_{a,t-1}^2}{A_{t-1} A_t^I} + \frac{\nu^{-2} y_{t-1}^2}{A_t^I} & \frac{\sigma_{ab,t-1}}{A_{t-1} A_t^I} - \frac{\nu^{-2} y_{t-1}}{A_t^I} \\ \frac{\sigma_{ab,t-1}}{A_{t-1} A_t^I} - \frac{\nu^{-2} y_{t-1}}{A_t^I} & \frac{\sigma_{b,t-1}^2}{A_{t-1} A_t^I} + \frac{\nu^{-2}}{A_t^I} \end{pmatrix}, \tag{75}$$

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

$$\Sigma_{b,t} = X \nu^{-2} \Sigma_t z_t = \nu^{-2} (\sigma_{ab,t} + \sigma_{b,t}^2 y_{t-1}), \tag{76}$$

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

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

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

$$= \frac{\sigma_{ab,t-1} + \sigma_{b,t-1}^2 y_{t-1}}{\nu^2 + \sigma_{a,t-1}^2 + 2 \sigma_{ab,t-1} y_{t-1} + \sigma_{b,t-1}^2 y_{t-1}^2}, \tag{77}$$

as

$$A_t^I = \frac{1}{\nu^2 A_{t-1}} \left(\nu^2 + \sigma_{a,t-1}^2 + 2 \sigma_{ab,t-1} y_{t-1} + \sigma_{b,t-1}^2 y_{t-1}^2 \right). \tag{78}$$

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

$$\Omega_{y_t} = Var(y_t | Y_{t-1}) = \nu^2 + z_t' \Sigma_{t-1} z_t = \nu^2 + \sigma_{a,t-1}^2 + 2 y_{t-1} \sigma_{ab,t-1} + y_{t-1}^2 \sigma_{b,t-1}^2, \tag{79}$$

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

$$\Sigma_{b,t} \Omega_{y_t} = \sigma_{ab,t-1} + \sigma_{b,t-1}^2 y_{t-1}. \tag{80}$$

Substituting this expression into the ALM yields

$$y_t = \left[\alpha + (\beta_0 + \beta_1) \mu_{a,t-1} + \beta_1 \mu_{a,t-1} \mu_{b,t-1} + \beta_1 \sigma_{ab,t-1} \right] + \left[\delta + \beta_0 \mu_{b,t-1} + \beta_1 \mu_{b,t-1}^2 + \beta_1 \sigma_{b,t-1}^2 \right] y_{t-1} + v_t. \square \tag{81}$$

References

Beck, G., Wieland, V., 2002. Learning and control in a changing economic environment. J. Econ. Dyn. Control 26 (9–10), 1359–1377.
Bray, M., Savin, N., 1986. Rational expectations equilibria, learning, and model specification. Econometrica 54, 1129–1160.
Bullard, J., 1992. Time-varying and nonconvergence to rational expectations equilibrium under least squares learning. Econ. Lett. 40, 159–166.
Cogley, T., Sargent, T., 2008. Anticipated utility and rational expectations as approximations of Bayesian decision making. Int. Econ. Rev. 49 (1), 185–221.
Evans, G., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.
Evans, G., Honkapohja, S., Williams, N., 2010. Generalized stochastic gradient learning. Int. Econ. Rev. 51 (1), 237–262.
Guidolin, M., Timmermann, A., 2007. Properties of equilibrium asset prices under alternative learning schemes. J. Econ. Dyn. Control 31 (1), 161–217.
Kiefer, N., Nyarko, Y., 1989. Optimal control of an unknown linear process with learning. Int. Econ. Rev. 30 (3), 571–586.
Kreps, D., 1998. Anticipated utility and dynamic choice. In: Jacobs, D.P., Kalai, E., Kamien, M. (Eds.), Frontiers of Research in Economic Theory: the Nancy L. Schwartz Memorial Lectures, 1983–1997, pp. 242–274.
Marcet, A., Sargent, T., 1989. Convergence of least squares learning mechanisms in self-referential linear stochastic models. J. Econ. Theory 48, 337–368.

McGough, B., 2003. Statistical learning with time-varying parameters. Macroecon. Dyn. 7, 119–139.
Rotemberg, J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy. In: NBER Macroeconomics Annual 1997, pp. 297–346.
Sargent, T., Williams, N., 2005. Impacts of priors on convergence and escapes from Nash inflation. Rev. Econ. Dyn. 8, 360–391.
Taylor, J., 1980. Aggregate dynamics and staggered contracts. J. Political Econ. 88, 1–23.
Wieland, V., 2000a. Learning by doing and the value of optimal experimentation. J. Econ. Dyn. Control 24 (4), 501–534.
Wieland, V., 2000b. Monetary policy, parameter uncertainty and optimal learning. J. Monet. Econ. 46 (1), 199–228.
Woodford, V., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton, NJ.
Zellner, A., 1971. Intoduction to Bayesian Inference in Econometrics. Wiley, Princeton University Press, Princeton, NJ.