Learnability of E–stable Equilibria [1]

Atanas Christev [2] and Sergey Slobodyan [3]

First Submitted 27 May 2010
This Version 18 August 2011

Abstract: If private sector agents update their beliefs with a learning algorithm other than recursive least squares, learnability of rational expectations equilibria (REE) is not guaranteed. Monetary policy under commitment, with a determinate and E–stable REE, may not imply robust learning stability of such equilibria if the RLS speed of convergence is slow. In this paper, we propose a refinement of E–stability conditions that select equilibria more robust to specification of the learning algorithm within the RLS/SG/GSG class. E–stable equilibria characterized by faster speed of convergence under RLS learning are learnable with SG or Generalized SG algorithms as well.

Journal of Economic Literature Classification Numbers: C62, D83, E17.

Keywords: Adaptive learning, Expectational Stability, Stochastic Gradient, Speed of Convergence

---

[1]We are grateful to Chrissy Giannitsarou, John Duffy and George Evans for the encouragement to pursue this project. We are also indebted to two anonymous referees and an Associate Editor of the journal for helpful comments. Financial support from the Royal Society of Edinburgh is gratefully acknowledged. All remaining errors are our own.

[2]Heriot-Watt University, Edinburgh and IZA, Bonn

[3]CERGE-EI, Prague

1. Introduction

Adaptive learning and expectational stability (E–stability) arise naturally in self-referential macroeconomic models. The literature on adaptive learning assumes that economic agents act as econometricians who run recursive regressions using historical data to inform their decisions. Adaptive learning, one description of boundedly-rational, real-world decision making processes, has an appealing feature: its asymptotic outcome may be consistent with rational expectations. Evans and Honkapohja (2001) provide the methodology and derive the conditions under which recursive learning dynamics converges to rational expectations equilibria. If economic agents use recursive least squares (RLS) learning to update their expectations of the future (or learn adaptively), then only E–stable REE can be the asymptotic outcomes of a real–time learning process. Equilibria, stable under a particular form of adaptive learning, are also called *learnable*.[4] Hence, E–stability is a necessary condition for RLS learnability.

Evans and Honkapohja (2001) also draw attention to the lack of general results on stability and convergence of different learning algorithms. Barucci and Landi (1997) and Heinemann (2000) demonstrate that E–stability may not be a sufficient condition for learnability if agents have adaptive algorithms other than RLS. Barucci and Landi (1997) show that an alternative learning mechanism, namely, stochastic gradient (SG) converges to REE but under different conditions than RLS learning.[5]

Furthermore,Giannitsarou (2005) provides examples, with a lagged endogenous variable, of E–stable equilibria which are not learnable under SG

[4]The possible convergence of learning processes and the E–stability criterion of REE dates back to DeCanio (1979) and Evans (1985). Marcet and Sargent (1989) first showed the conditions for convergence in a learning context using stochastic approximation techniques.

[5]Evans and Honkapohja (1998) find that E–stability and SG–stability are the same in a multivariate cobweb-type model.

learning. Evans et al. (2010) discuss additional conditions (related to the knowledge of certain characteristics of the REE) which ensure E–stable equilibria are learnable under SG and Generalized Stochastic Gradient (GSG) learning. They propose the GSG algorithm as an extension to model learning in the presence of uncertainty and parameter drift in agents' beliefs. The algorithm is a maximally robust learning rule. In addition, and more important, the authors illustrate that given a particular weighting matrix the conditions for GSG–stability are closely related to E–stability, and equivalent in a New Keynesian model of monetary policy with alternative interest rate rules.

Evans and Honkapohja (2003) and Bullard (2006), among others, establish E–stability criterion as a minimum requirement for the design of meaningful monetary policy. E–stability of the resulting REE is a desriable property of any monetary policy rule, claim Bullard and Mitra (2007); in effect equilibria ought to be learnable under RLS. Recently Tetlow and von zur Muehlen (2009) pose a problem of policy design in a world populated with agents who might learn using a misspecified model of the economy. They state that an equilibrium which is learnable for a wide range of possible specifications, even at a potential cost of welfare losses, is a valuable property of a monetary policy rule. Robustifying policy in this way ensures that learnability is achieved on the transition path to REE, without compromising convergence.

In this paper, we focus on the properties of E–stable equilibria which facilitate such a design problem allowing learnability under both RLS and GSG.[6] We propose a refinement of E–stability conditions that select equilibria more robust to specification of the learning algorithm within the RLS/SG/GSG

[6]In the literature, the concepts of E–stability and learnability are used interchangeably. Following Giannitsarou (2005) we consider learnability as a broader concept and distinguish between E–stability that is related to learnability under RLS and learnability under different learning algorithms.

class. We show that the (mean–dynamics) speed of convergence under RLS learning is an important component of such a refinement, because E–stable equilibria, characterized by a faster RLS speed of convergence, are likely to remain learnable under SG or GSG algorithms as well. The mean-dynamics speed of convergence, discussed in Ferrero (2007), can also have consequences for welfare, and is related to the asymptotic behavior of the agents' beliefs as demonstrated by Marcet and Sargent (1995).

We extend E–stability requirements using two additional criteria. First, we require that REE be learnable under a broad set of learning algorithms of the RLS/SG/GSG class. In some sense, this allows us to choose a subset of REE with properties such that they remain learnable even if agents' learning process is misspecified asymptotically relative to RLS. Second, the speed of convergence under RLS should be fast not only to aid learnability, but also to ensure a fairly quick return of agents' beliefs towards the REE even after a small disturbance or deviation.[7] We confirm that the two additional criteria are related and can be met simultaneously.

2. E–Stability and Learnability Revisited

It is well established in Evans and Honkapohja (2001) and elsewhere that the convergence of the RLS algorithm is closely related to E–stability. The equilibrium is said to be E–stable if a stationary point $\overline{\mathit{\Phi}}$ of the following ordinary differential equation (ODE) is asymptotically stable:

$$\frac{d\mathit{\Phi}}{d\tau} = T(\mathit{\Phi}) - \mathit{\Phi}. \tag{1}$$

$\overline{\mathit{\Phi}}$ corresponds to the rational expectations equilibrium of a forward–looking model. T is the mapping from perceived law of motion (PLM) to actual law of motion (ALM), and Φ is a vector of the parameters of interest. The

[7] Ferrero (2007) examines the welfare consequences of slow adjustment of inflationary expectations to their REE values.

differential equation (1) describes the behavior of approximating, or "mean" dynamics in continuous "notional" or meta- time.[8] Its equilibrium point is asymptotically stable if the Jacobian of (1) evaluated at $\overline{\Phi}$,

$$J = DT(\Phi)\,|_{\Phi=\overline{\Phi}} - I,$$

has only eigenvalues with negative real parts.[9]

If, instead of using RLS, economic agents rely on SG learning, the convergence of the mean dynamics of the learning process is governed by the following ODE,

$$\frac{d\Phi}{d\tau} = M(\Phi) \cdot (T(\Phi) - \Phi)\,, \tag{2}$$

where $M(\Phi)$ is a symmetric and positive-definite matrix of second moments of the state variables used by agents in forming their forecasts.

The RE equilibrium $\overline{\Phi}$ is still a stationary point of (2). It is learnable if $\overline{\Phi}$ is the locally asymptotically stable equilibrium of the ODE (2), which obtains when all eigenvalues of $M(\overline{\Phi}) \cdot J$ have negative real parts. Barucci and Landi (1997) first provided a proof of this result. It is important to note that the conditions which establish the analogue of the E–stability condition in this case are different from those obtained under RLS.

If the agents update their beliefs with a Generalized SG learning (GSG) algorithm instead, learnability is related to the negative real parts of all eigenvalues of the matrix

$$\Gamma M(\overline{\Phi}) \cdot J, \tag{3}$$

where we explicitly restrict our attention to weighting (symmetric) positive definite matrices Γ such that $\Gamma M(\overline{\Phi})$ is also arbitrary symmetric, positive

[8]There are technical conditions other than the stability of the approximating mean dynamics which ought to be satisfied for convergence of the real–time dynamics under learning, see Evans and Honkapohja (2001) Chapter 6, Sections 6.2 & 6.3. We assume that these conditions are always satisfied, and claim that learnability is obtained when the equilibrium is stable under the approximating mean dynamics.

[9]The exceptional cases where we observe eigenvalues with zero real part do not typically arise in economic models.

definite. The class of such matrices includes $\Gamma = I$ (classic SG) and $\Gamma = M(\overline{\Phi})^{-1}$ (GSG asymptotically equivalent to RLS), a linear combination of them, or such matrix Γ that has the same eigenspace as $M(\overline{\Phi})$ (GSG more generally).[10] This fact is well documented and illustrated in Evans et al. (2010).

The problem of a correspondence between E–stability and GSG–stability for all Γ is, therefore, equivalent to the following linear algebraic problem. Given a matrix J with all its eigenvalues to the left of the imaginary axis, can we guarantee that no eigenvalue of $\Gamma M(\overline{\Phi}) \cdot J$ becomes positive? This problem is well known and is referred to as H–stability, and discussed earlier in Arrow (1974), Johnson (1974b), Johnson (1974a), and Carlson (1968). A sufficient condition for H–stability that is easy to check exists: matrix J is *H*–stable if its symmetric part, $\frac{1}{2}\left(J + J^T\right)$, is stable. Such a matrix is called negative quasi–definite. It is rather difficult to interpret this condition meaningfully from an economic point of view. Again Evans et al. (2010) provide an economic example and an extended discussion of GSG learning and H–stability.

While the convergence of the adaptive learning algorithms has been extensively studied, the transition, along the learning path, towards the equilibrium REE of interest is less well understood. Our starting point of reference is the results in Benveniste et al. (1990) and Marcet and Sargent (1995) who first identified the behavior of the speed of convergence (how fast or slow agents' beliefs approach a REE point) and analyzed the asymptotic properties of the fixed point under RLS learning. For the purposes of the paper, we use the term "speed of convergence" to mean the minimum of the real parts of the eigenvalues of the linearized E–stability ODE. This value governs the

[10] An orthonormal basis is the eigenspace of the identity matrix. Evans et al. (2010), p. 246, choose to transform variables using Cholesky transformation, and hence allow *any* matrix Γ to have the same eigenspace as M. Diagonal matrices, as in their Proposition 5, p. 247, also have the same eigenspace consisting of coordinate unit vectors.

speed of convergence of the mean dynamics under RLS learning.

For the linearized E–stability ODE

$$\frac{d\Phi}{d\tau} = J \cdot \Phi, \tag{4}$$

where all eigenvalues of J are distinct and have only negative real parts, the solution will be given as a linear combination of terms of the form $C_i \cdot e^{\lambda_i \cdot t}$, where λ_is are the eigenvalues of J and C_i are arbitrary constants. In the long run, the solution is dominated by the term which corresponds to the smallest $\mid Re(\lambda_i) \mid$, the absolute value of the real part of the eigenvalue. [11] In the context of adaptive learning, this speed of convergence determines how fast the approximating mean dynamics described by the ODE in (1) approaches the REE asymptotically. Under standard decreasing-gain RLS learning, the speed of convergence is time-varying, subsiding as time evolves, and changes along the mean dynamics with the parameter estimates of the perceived law of motion, $\hat{\Phi}$.

The behavior along the transition path and the importance of short-run deviations away from the REE were illustrated by Evans and Honkapohja (1993) and Marcet and Sargent (1995). Ferrero (2007) further argued that the speed of convergence can be considered as an important policy variable in the design of monetary policy. An open question is how important is the RLS speed of convergence for learnability under alternative learning algorithms.

The concepts of GSG–learning stability for all Γ and the speed of convergence under RLS appear distinct and far apart. However, it turns out that there is a close connection between the two. Consider, for example, the model in Sections 2 and 3 of Giannitsarou (2005). The reduced form of this

[11] With RLS learning, the mean-dynamics ODE which governs estimates of the second moments matrix of regressors, R, is given by $\frac{dR}{d\tau} = M(\Phi) - R$. The Jacobian of this part of the mean-dynamics ODE equals $-I$, with all eigenvalues equal to -1, see Evans and Honkapohja (2001) p.234-235.

univariate model is given by

$$\begin{aligned} y_t &= \lambda y_{t-1} + \alpha E_t^* y_{t+1} + \gamma w_t, \\ w_t &= \rho w_{t-1} + u_t. \end{aligned}$$

In this model, $|\rho| < 1$ and $u_t \sim N(0, \sigma_u^2)$. The equilibrium of the model, with the same parameter values as in the paper: $\gamma = \sigma_u = 0.5$, $\rho = 0.9$, is E–stable, and therefore learnable under RLS. Both eigenvalues are real for all values of (α, λ) for which the solution $\overline{\Phi}_-$ is stationary and E–stable.

The E–stability ODE for this model is given by equation (1), where the mapping T is defined by equation (5) of Giannitsarou (2005), page 277, and the vector Φ is two–dimensional. Figures 1 and 2 summarize the negative quasi–definiteness for the corresponding Jacobian and the speed of convergence of the mean dynamics, respectively, as a function of the parameters α and λ. These figures clearly indicate that negative quasi–definiteness obtains in regions in the parameter space where the convergence speed is higher. Figure 2 is a contour plot of the speed of convergence which increases towards the lower left corner of the graph, a region where the negative quasi–definiteness of the Jacobian is also observed.

In the region of the parameter space where the Jacobian is not negative quasi–definite, we expect that a matrix Γ exists such that $\Gamma M(\overline{\Phi}) \cdot J$ is not stable. Therefore, the GSG learning algorithm that corresponds to this Γ does not result in an approximating dynamics converging to the REE. This conjecture stands correct: Giannitsarou (2005) shows that the equilibrium achieved under the SG learning algorithm (for which Γ equals the identity matrix) cannot be learned for a small set of parameter values. Figure 2 illustrates that for these parameter values the equilibrium is SG-unstable, and the speed of convergence is close to zero. On the other hand, for parameter values that correspond to a negative quasi–definite Jacobian, the speed of convergence is large and is never less than 0.35. Given that the negative quasi–definiteness is a sufficient condition for H–stability, we are guaranteed that GSG learning with any choice of Γ induces a learnable equilibrium. In particular, SG–learning will always converge for this set of parameter values.

[Figure 1 and Figure 2 about here]

This analysis indicates that there appears to exist a close relationship between SG–learning stability and the speed of convergence of the mean dynamics under RLS learning. The paper studies the nature of this relationship and addresses the following questions:

1. Why is SG–instability associated with a lower speed of convergence under RLS?

2. In contrast, why do conditions which guarantee fast convergence to REE, also seem to ensure SG– and GSG–stability for any Γ?

3. Are the answers to (1) and (2) general and applicable enough to a wide variety of self–referential models?

In what follows, we provide a two–dimensional geometric interpretation of the case when E–stability holds but SG–stability is not achieved (i.e., a matrix J is stable but not H–stable), and relate this finding to the speed of convergence of expectations to their REE values under RLS learning. In other words, when does E–stability fail to imply learnability for some GSG learning algorithm?

3. A Geometric Interpretation of Learnability

Provided J is a stable matrix , when is $\Omega \cdot J$ stable? We propose a simple two-dimensional geometric approach to answer this question. To preview our results: we study the eigenvalues of a matrix J, but not its components, and then relate those values to the speed of convergence of the mean dynamics under recursive least-squares learning. These findings have an intuitive and meaningful interpretation in a wide variety of adaptive learning models.

Now suppose that the 2×2 matrix J has only eigenvalues with negative real parts. This matrix is the Jacobian of the ODE in (1) for some adaptive learning model. J is asymptotically stable, therefore, the equilibrium associated with the model is E–stable and learnable under RLS.

The eigenvalue problem of J can be written as

$$J \cdot V = V \cdot \Lambda, \tag{5}$$

where V is the matrix with columns containing the eigenvectors of J, and Λ is diagonal with the corresponding eigenvalues λ_i on the main diagonal. If the eigenvectors are linearly independent, associated with distinct eigenvalues, the matrix J can be diagonalized as $J = V \Lambda V^{-1}$.[12]

Learnability of the equilibrium with GSG learning is determined by the eigenvalues of $\Omega \cdot J$. For our purposes, Ω is assumed to be symmetric and thus can be written as $\Omega = P \Delta P^T$, where Δ is diagonal with main elements the eigenvalues of Ω. [13] The eigenvalue problem for $\Omega \cdot J$ can be written as:

$$P \Delta P^T \cdot V \Lambda V^{-1} \cdot \widetilde{V} = \widetilde{V} \cdot \widetilde{\Lambda}, \tag{6}$$

where the columns of $\widetilde{V}$ are the eigenvectors of $\Omega \cdot J$ and $\widetilde{\Lambda}$ is a diagonal matrix with the eigenvalues of $\Omega \cdot J$ as the main entries.

Next pre–multiply (6) by P^{-1} and define $\overline{V} = P^{-1} \widetilde{V}$ to get

$$\Delta \cdot P^T V \Lambda V^{-1} P \cdot \overline{V} = \Delta \widetilde{J} \cdot \overline{V} = \overline{V} \cdot \widetilde{\Lambda}. \tag{7}$$

It is clear that the matrix $\widetilde{J} = P^T V \Lambda V^{-1} P$ has the same eigenvalues as J, i.e., the values on the main diagonal of Λ. Geometrically, if J represents a linear mapping in a two–dimensional space, then $\widetilde{J}$ represents the same mapping in new coordinates, given by the two orthogonal eigenvectors of Ω.

[12]Evans and Honkapohja (2001), page 90, explain, the set of diagonalizable matrices is "generic" in the sense that it constitutes a subset of the set of all $n \times n$ matrices. In addition, the set of digonalizable matrices "commutes" if and only if the set is simultaneously diagonalizable, see Horn and Johnson (1985), page 50 and 55-56.

[13]The eigenvectors of a symmetric matrix are orthogonal, and so $P^{-1} = P^T$.

We work in the new coordinates and replace the problem of seeking conditions on the eigenvalues and eigenvectors of J such that $\Omega \cdot J$ has a positive eigenvalue (i.e., turns unstable) with the equivalent problem concerning $\widetilde{J}$ and $\Delta\widetilde{J}$. To fix notation, let us order δ_1 and δ_2, the eigenvalues of Ω, so that the following is always true, $\left(\frac{\delta_2}{\delta_1}\right) > 1$. The eigenvalues of J and $\widetilde{J}$ are $-\lambda_1$ and $-\lambda_2$, and ordered so that $\frac{|\lambda_2|}{|\lambda_1|} > 1$.[14] Denote the eigenvectors of $\widetilde{J}$ corresponding to $-\lambda_1$ and $-\lambda_2$ as $v_1 = (v_{11}, v_{21})^T$ and $v_2 = (v_{12}, v_{22})^T$. Define $\Upsilon = \frac{v_{22}}{v_{21}}\frac{v_{11}}{v_{12}}$.

PROPOSITION 1: *Let $\lambda_{1,2}$ be real. The matrix $\Omega \cdot J$ has a positive eigenvalue and thus J is not H–stable if and only if the following conditions hold:*

(i) $0 < \Upsilon < 1$,

(ii) $\frac{\lambda_2}{\lambda_1} > \frac{1}{\Upsilon}$ *, and*

(iii)

$$\frac{\delta_2}{\delta_1} > \frac{\left(\frac{\lambda_2}{\lambda_1} - \Upsilon\right)}{\left(\frac{\lambda_2}{\lambda_1}\Upsilon - 1\right)} \tag{8}$$

Proof. See Appendix A. □

COROLLARY 1: *Let $\lambda_{1,2}$ be real. If either $\Upsilon < 0$ or $\Upsilon > \frac{\lambda_1}{\lambda_2}$, the matrix J is H–stable and the equilibrium is learnable for any GSG learning algorithm.*

Proof. See Appendix A. □

Proposition 1 constructs a counterexample of a matrix Ω such that $\Omega \cdot J$ is not stable. This means that agents who update their beliefs adaptively with the corresponding GSG algorithm cannot learn the REE, even though it is E–stable. A necessary condition for E–stable REE not to be learnable under GSG learning algorithm for any Γ is a positive Υ less than one. Geometrically,

[14] Given this choice of notation, the speed of convergence is equal to $|\lambda_1|$.

$0 < \Upsilon < 1$ implies that, after rotation into the system of coordinates defined by the eigenvectors of Ω, the two eigenvectors of J are in the same quadrant. When the eigenvectors are close to being collinear, finding Ω to satisfy this geometric condition is more likely, in the sense that the approach to the REE occurs relatively fast. [15]

It might be impossible to satisfy the assumptions of Proposition 1 if the eigenvectors are close to being orthogonal, as in this case Υ is either too small (less than $\frac{\lambda_1}{\lambda_2}$) or too large (above 1) when it is positive. If they are exactly orthogonal the positive Υ equals 0 or ∞. The necessary (and sufficient) condition described in (8) shows that, for a given Υ, instability of $\Omega \cdot J$ is more likely when either $\left(\frac{\lambda_2}{\lambda_1}\right)$ is large or $\left(\frac{\delta_2}{\delta_1}\right)$ is large or both. The latter occurs when the eigenvalues of Ω are highly unbalanced. If the agents update their beliefs using SG learning, this means that Ω is the covariance matrix of regressors with highly unequal diagonal terms. This result indicates how the speed of convergence plays an essential role in leading to (G)SG–instability in general: the slower you approach the REE the more likely it is to observe SG–instability, as exhibited in our simulations.

Moreover, the ratio $\left(\frac{\lambda_2}{\lambda_1}\right)$ will be large if $|\lambda_1|$ is very small, i.e., the speed of convergence under RLS is small. Increasing $|\lambda_1|$ will facilitate SG learnability as the condition in Proposition 1 becomes more difficult to fulfill. In particular, it will require much more unbalanced matrix Ω. The economic agents, who learn adaptively, are less likely to be using SG algorithm when the (co)variances of the regressors are extremely unequal. For example, the scaling invariance issue with SG identified by Evans et al. (2010) is more severe. Therefore, the speed–of–convergence criterion directly ensures that the set of equilibria learnable by agents using algorithms within the RLS/SG/GSG class is sufficiently large.

[15] Note that the angle between the eigenvectors of J is preserved under the rotation into the orthogonal coordinate system determined by the eigenvectors of Ω. Therefore, we use collinearity of the eigenvectors of J and $\widetilde{J}$ interchangeably.

Figure 3 plots the necessary and sufficient condition (8) for three values of Υ: 0.3, 0.1 and 0.05 (eigenvectors of $\widetilde{J}$ close to being orthogonal), respectively. The condition is satisfied in the area of the figure located above and to the right of the corresponding line. If v_1 and v_2 are almost collinear, SG–instability could be achieved for relatively mild ratios of the eigenvalues $\left(\frac{\lambda_2}{\lambda_1}\right)$ and $\left(\frac{\delta_2}{\delta_1}\right)$.

[Figure 3 about here]

Turning to the results in Giannitsarou (2005), we show that points in the parameter space for which the equilibrium is not learnable under SG satisfy the conditions of Proposition 1. In this case, the corresponding values of $\left(\frac{\lambda_2}{\lambda_1}\right)$ and $\left(\frac{\delta_2}{\delta_1}\right)$ are extreme and lie in the vicinity of 112 and $5.5 \cdot 10^5$, respectively. Such a high degree of imbalance in the matrix Ω emerges because Υ is very close to zero (the eigenvectors of J are almost orthogonal) throughout the whole parameter space. For these parameters the speed of convergence is slow (in the order of magnitude of 10^{-3}), and hence the ratio of the eigenvalues of J is very large. It is clear from this example that as the parameters of the model change, so do the associated eigenvalues. Hence, Proposition 1 establishes a link between GSG-stability conditions and the speed of convergence, both of which are influenced by the agents' estimates of the model parameters. We attempt to quantify this intuition via simulations in the next section.

For completeness, we next turn to the case when the eigenvalues of the stability matrix are complex. To fix notation, assume that J has two complex eigenvalues: $\nu \pm i\mu$, $\nu < 0$, and two complex eigenvectors $w_1 \pm iw_2$, where w_1 equals $(w_{11}, w_{21})^T$ and w_2 is $(w_{12}, w_{22})^T$. Define $\widetilde{W} = w_{11}w_{12} + w_{21}w_{22}$ and $|W| = w_{11}w_{22} - w_{12}w_{21}$. The following Proposition provides the necessary and sufficient condition for the instability of the matrix $\Omega \cdot J$:

Proposition 2: *Let $\lambda_{1,2}$ be complex. The matrix $\Omega \cdot J$ has a positive eigenvalue and thus J is not H–stable if and only if the following condition* *holds:*

$$\frac{\mu}{|\nu|}\frac{\widetilde{W}}{|W|}\frac{\left(\frac{\delta_2}{\delta_1}-1\right)}{\left(\frac{\delta_2}{\delta_1}+1\right)} > 1. \tag{9}$$

Proof. See Appendix B. □

Similarly, Proposition 2 demonstrates that a smaller ratio of the eigenvalues of Ω is conducive to the stability of $\Omega \cdot J$. Hence, the corresponding GSG learning algorithm generates a convergent dynamics. A higher speed of convergence, larger $|\nu|$, makes the necessary and sufficient condition described in (9) harder to fulfill, and thus increases the set of parameters for which equilibria are learnable. Orthogonality of w_1 and w_2 means that the condition in Proposition 2 cannot be satisfied. The intuition is similar to the real eigenvalue case: in both cases orthogonality of eigenvectors (real case) or their real and imaginary components (complex case) ensures that GSG–learning instability is impossible.

Furthermore, Propositions 1 and 2 construct examples and state that if the mean-dynamics Jacobian J is located farther away from the instability region (at least one eigenvalue with positive real part), it is harder to find a "disturbance" of J — pre–multiplication by a (symmetric) positive definite matrix — which will lead to instability. Since the E–stability conditions should be easier to satisfy, we guarantee robustness of the learning rule against such "disturbances." We establish that the speed of convergence facilitates this robustness property in the class considered here.

In this sense, Propositions 1 and 2 indicate that the second criterion we impose on all desirable REE — the high speed of convergence under RLS learning — is in accordance with the first criterion, namely the REE are learnable under a range of learning algorithms within the RLS/SG/GSG class. To illustrate further the alignment of these two criteria and the way in which they modify selection of monetary policy rules, we study a standard model of monetary policy under commitment with learning.

4. Monetary Policy under Commitment

4.1. The model environment

Following Evans and Honkapohja (2006), we start with a standard two-equation New Keynesian (NK) model:

$$x_t = -\varphi\left(i_t - \widehat{\pi}_{t+1}\right) + \widehat{x}_{t+1}, \tag{10a}$$

$$\pi_t = \lambda x_t + \beta\widehat{\pi}_{t+1} + u_t. \tag{10b}$$

Here x_t and π_t express the output gap and inflation in period t, and all variables with hats denote private sector expectations. i_t is the nominal interest rate, in deviation from its long run steady-state. The parameters φ and λ are positive and have the standard explanation and the discount factor is $0 < \beta < 1$. Our main interest is in the learning behavior of private sector agents, and we maintain the assumption that expectations may not be rational. We also assume the presence of only one shock to illustrate the results in this paper, and disregard the influence of the demand shock in the IS equation (10). The cost-push shock in (10b) is given by $u_t = \rho u_{t-1} + \epsilon_t$ where $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$ is independent white noise. In addition, $|\rho| < 1$.

We consider the expectations–based interest rate policy rule under commitment, using timeless perspective solution:

$$i_t = \phi_L x_{t-1} + \phi_\pi \widehat{\pi}_{t+1} + \phi_x \widehat{x}_{t+1} + \phi_u u_t. \tag{11}$$

The optimal values of the policy rule parameters, based on a standard loss function, are given in Evans and Honkapohja (2006), page 26, equation (15) (notice that the coefficient ϕ_g is assumed to be zero in our specification). Here we do not restrict our attention to optimal monetary policy. We fix the values of the policy parameters, ϕ_L and ϕ_u, at their optimal level, and treat the other two policy parameters as choice variables of the policy response of the monetary authority.

Under the assumed policy rule the model can be written as:

$$y_t = \begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} 1-\varphi\phi_x & \varphi(1-\phi_\pi) \\ \lambda(1-\varphi\phi_x) & \lambda\varphi(1-\phi_\pi)+\beta \end{bmatrix} \begin{bmatrix} \widehat{x}_{t+1} \\ \widehat{\pi}_{t+1} \end{bmatrix} + \begin{bmatrix} -\varphi\phi_L & 0 \\ -\varphi\lambda\phi_L & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} -\varphi\phi_u \\ 1-\lambda\varphi\phi_u \end{bmatrix} u_t \tag{12}$$

$$y_t = A\widehat{y}_{t+1} + CJ^T y_{t-1} + Bu_t$$

where $J = (1,0)^T$.

The MSV solution of this system can be expressed in the following way, with c and b both being vectors such that $c = (c^x, c^\pi)^T$, $b = (b^x, b^\pi)^T$:

$$y_t = cJ^T y_{t-1} + bu_t. \tag{13}$$

Using the method of undetermined coefficients we find the REE solution, where c^x solves the cubic equation:

$$c^x = -\varphi\phi_L + (1-\varphi\phi_x)(c^x)^2 + \frac{\lambda\varphi(1-\phi_\pi)(c^x)^2}{1-\beta c^x}.$$

The rest of the solution is provided in:

$$c^\pi = \frac{\lambda c^x}{1-\beta c^x}$$

and

$$b = \left(I - A\left(cJ^T + \rho I\right)\right)^{-1} B.$$

Next we turn to the conditions under which REE are learnable. We check for determinacy of the RE solution using the conditions derived in Evans and Honkapohja (2006), page 35, which are not reproduced here.

4.2. Determinacy and E–stability: the minimum requirements for desirable RE equilibria

Discussing E–stability, we follow the rest of the literature in assuming that the MSV solution obtained above is the PLM used by the private sector agents in the model. Let us re-write the model as:

$$y_t = A\widehat{y}_{t+1} + CJ^T y_{t-1} + Bu_t,$$
$$u_t = \rho u_{t-1} + \epsilon_t.$$

Calculate $\widehat{y_{t+1}}$, the non-rational expectations of the model (13), as:

$$\widehat{y}_{t+1} = E_t^*[cJ^T y_t + bu_{t+1}] \quad = \left[A\left(cJ^T\right)\left(cJ^T\right) + CJ^T\right] y_{t-1} + \left[A\left(cJ^T + \rho I\right) b + B\right] u_t.$$

Therefore, the T–map for the problem becomes:

$$T(b,c) = \left(A\left(cJ^T + \rho I\right) b + B, \left(c^T J\right) Ac + C\right).$$

This allows us to compute the (4×4) Jacobian matrix:

$$J = \begin{bmatrix} A\left(\overline{c}J^T + \rho I\right) - I_2 & \left(J^T \overline{b}\right) A \\ 0 & A\left(\overline{c}J^T + J^T\overline{c} \cdot I\right) - I_2 \end{bmatrix}.$$

4.3. SG–stability

The matrix, M_z, of the second moments of the state variables in the model used by the agents to forecast the inflation and the output gap is obtained from a two–variable VAR written as:

$$\begin{bmatrix} x_t \\ u_{t+1} \end{bmatrix} = \begin{bmatrix} c^x & b^x \\ 0 & \rho \end{bmatrix} \begin{bmatrix} x_{t-1} \\ u_t \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \epsilon_{t+1}.$$

Then, to study the SG–stability of the model, we examine the eigenvalues of $\Omega \cdot J$ which are now:

$$\Omega \cdot J = (M_z \otimes I_2)\, J.$$

There is no intercept in the learning model to reduce dimensionality, and it is not essential for our purposes (see Evans et al. (2010)).

We also study GSG learning stability in which the agents are updating their beliefs about the parameters in the model by making small errors around the outcomes achieved under RLS learning. We adopt Γ in (3) to be the inverse of a second moments matrix $\widetilde{M}_z$, which is compatible with the same model evaluated at parameter values obtained within the neighborhood of the calibrated parameters used in the simulation analysis. We assume these alternative learning algorithms to model agents' uncertainty about the second moments of the state variables which they need to know in order to be able to run regressions, and make GSG learning asymptotically equivalent to RLS learning. GSG–learning stability of the equilibria under this assumption is achieved when all eigenvalues of the matrix $((\widetilde{M}_z^{-1} \cdot M_z) \otimes I_2) \cdot J$ have negative real parts. We note that the (Cholesky) transformation of variables proposed by Evans et al. (2010) delivers the same E–stability ODE as running generalized Least Squares with $\Gamma = M_z^{-1}$. However, along the transition to the asymptotic equilibrium the dynamics in these two cases will be different. While the variable transformation in Evans et al. (2010) requires full knowledge of the covariance matrix of the regressors, we allow for agents' mistakes regarding its estimates.

Propositions 1 and 2 cannot be stated in the four dimensional case we have specified in the model, and therefore we analyze the stability of the system via simulations. Still, we expect that the main findings presented in the Propositions remain valid in the higher dimensions as well, namely that the lower speed of convergence of the mean dynamics under RLS will be associated with higher incidence of GSG–instability.

5. Learning Instability and Equilibria: Discussion

To analyze the link between GSG–stability for all Γ and the speed of convergence under RLS, we resort to simulations of the simple NK model under commitment for different values of the expectations-based policy rule parameters. We take this example as it gives robust learning stability argued in Evans and Honkapohja (2003). We do not perform exhaustive study of the possible monetary policy rules in this model. To be more precise, we keep the parameters ϕ_L and ϕ_u at their optimal values derived in Evans and Honkapohja (2006), but vary ϕ_π and ϕ_x in a sufficiently broad range. The theoretical results on expectation-based policy rules under commitment, namely determinacy and E–stability of the REE, for any parameter values, were derived only for optimal policy by Evans and Honkapohja (2006). Therefore, we proceed to check every point for determinacy and E–stability (i.e., checking the eigenvalues of J for a negative real part). In addition, we calculate the speed of convergence of the mean dynamics under RLS learning, as described in Section 2, and check for convergence of the SG learning algorithm, by evaluating the eigenvalues of $(M_z \otimes I_2) \cdot J$.

We calibrate our model using the following parameter values. They are the same as in Clarida et al. (1999) calibration: $\beta = 0.99$, $\varphi = 4$ and $\lambda = 0.075$, and also used in Evans and Honkapohja (2003). We assume different values for the persistence of cost–push shock, $\rho = 0.90$ (commonly used in the literature) and $\rho = 0.60$. We also perform sensitivity analysis and robustness checks for various combinations of parameters other than ρ, e.g. the various permutations of the parameter space Evans and McGough (2010). We vary the number of simulation runs and confirm that the results are not altered.

The results of our simulations with $\rho = 0.90$ are presented in Figure 4. It illustrates, for every pair of the policy parameters (ϕ_π, ϕ_x), whether the resulting REE is determinate, E–stable, and SG–stable.[16] We only plot the E–stable points. The black area represents all indeterminate equilibria. We see that the standard Taylor principle applies (see Llosa and Tuesta (2009) for the theoretical derivations). The points satisfying the Taylor principle

[16] We also check the sufficient condition for H–stability (symmetric part of J stable) but in the range of our calibration and policy parameters we found no points to satisfy the condition. This further shows how restrictive the negative quasi-definiteness of a matrix proves to be.

are further split into SG–stable (white area) and SG–unstable (grey area).[17] SG–instability is concentrated in areas where ϕ_π is relatively low; as evident, more active monetary policy is associated with SG–stability.

How does this result fare against our Propositions 1 and 2, which associate the robustness of learning stability under alternative algorithms with the higher convergence speed under RLS? In Figure 5 you see the link clearly. The association is shown by plotting contour levels of the speed of convergence for the same values of (ϕ_π, ϕ_x). All SG–unstable points have a low convergence speed. Moving to more active monetary policy under commitment, both the speed of convergence and the robustness of SG learnability increase.

To compare our results with those of Evans and Honkapohja (2006), we plot a black asterisk at the point corresponding to the optimal monetary policy for our calibrated values.[18] As expected, this policy delivers determinate and E–stable REE; however, notice that this policy is very close to both SG–stability and E–stability boundaries. This proximity raises an issue of robustness of the optimal monetary policy if the agents are making small mistakes in their learning process. Evans and McGough (2010) study optimal monetary policy in NK models with inertia. They show that such policy typically is located on (near) the boundary of a set in the space of policy parameters where an E–stable and determinate equilibrium obtains. Small mistakes in calculating the policy parameters thus could lead to E–instability, indeterminacy, or both. We consider a forward–looking model, with lagged endogenous variable, where inertia is introduced through monetary policy under commitment. We show that even if we select policy parameters well inside the E–stable and determinate region the outcome may turn unstable when the learning algorithm adopted by agents is SG, or GSG which is not asymptotically equivalent to RLS.

[17] We do not track SG–stability for indeterminate REEs.

[18] To derive the optimal policy values used in our simulations, we assume a relative weight equal to 0.02 on the output gap.

We perform the following experiment to study the robustness of the optimal or near–optimal monetary policy. We assume that the agents update their beliefs by running not RLS but instead use a GSG learning algorithm. If the agents knew exactly the second moments matrix M_z associated with the parameter values (including the optimal monetary policy parameters) of the model, they would run a GSG that used M_z^{-1} as a weighting matrix. This GSG algorithm would be asymptotically equivalent to RLS, delivering determinate and E–stable REE, as explained in Evans et al. (2010).

In the experiment, we also assume that agents face uncertainty regarding the second moments matrix. Given that the agents are *learning* second–order moments, such as correlations between future inflation and past output gap and the cost–push shock (and, therefore, they do not know them, at least away from the REE achieved asymptotically in infinite time), it seems natural to assume that their knowledge of other second–order moments is limited as well. Thus economic agents take the deep, structural parameters of the model to be somewhere in the neighborhood of the "true" parameter vector θ that we use in simulations. The agents would believe in $\widetilde{\theta}$ and use it to compute the second moments matrix $\widetilde{M}_z(\widetilde{\theta})$.[19] Then the agents would tend to use the matrix $\widetilde{M}_z^{-1}$ as a weighting matrix in their updating of beliefs. Hence, $\widetilde{M}_z^{-1} = \Gamma$ in equation (3). The condition for the convergence of this real–time learning process, as we explained, is given by all eigenvalues of the matrix $((\widetilde{M}_z^{-1} \cdot M_z) \otimes I_2) \cdot J$ being negative.

We draw realizations of agents' beliefs about the parameters, $\widetilde{\theta}$, from a distribution that is centered at the true parameters θ. The range of the distribution is comparable to the prior distributions usually found in the literature on estimated DSGE models, for example, Milani (2007). We nest

[19] We do not think that agents who learn adaptively and "know" exactly a wrong second moments matrix is an assumption that is any more restrictive than assuming they are endowed with perfect knowledge of M_z. Using this procedure, we intend to generate "perturbed" second moments matrices with correlation structure that is similar to the true one.

the true RLS learning in this procedure because you could argue that SG–learning is too different from the RLS (for example, it is not scale invariant, see Evans et al. (2010)) to be a realistic description of any actual updating process. Then, we check whether this GSG algorithm is learnable or not. By repeating this procedure one thousand times, we obtain an estimate of the probability of obtaining GSG instability for a given parameter pair, (ϕ_π, ϕ_x). [20,21]

The simulation exercise, once again, confirms the results described in Propositions 1 and 2. Figure 6 shows that, for the SG–unstable points with a low RLS mean-dynamics convergence speed, we generally observe high incidence of GSG learning instability with imperfect knowledge of the second moments matrix. For the lowest speed still consistent with determinate and E–stable REE, we observe up to 60% probability of GSG instability. The optimal monetary policy (black asterisk) is associated with about 20% probability of becoming SG unstable. This result lends support to the finding in Evans and McGough (2010) where optimal monetary policy is shown to be rendered E–unstable or indeterminate by small mistakes committed by adaptive agents.

The probability of observing a divergent GSG algorithm measures only how likely it is to find parameter draws such that the agents' misperceptions become strong enough to lead to expectational instability. How far these mistaken perceptions should be from the "truth" in order to generate a divergent algorithm? To answer this question, we take the matrix $\widetilde{M}_z^{-1} M_z$ and evaluate its largest eigenvalue.[22] We take this value as a measure of the mismatch between the "true" second moments matrix M_z and the agents'

[20]We assume that the agents keep parameters (ϕ_π, ϕ_x) the same but re–calculate (ϕ_L, ϕ_u).

[21]The exact values of the probability of obtaining a GSG algorithm that delivers a learnable equilibrium depend on the assumed distribution of $\widetilde{\theta}$. We are only interested in the direction in the parameter space in which this probability increases or decreases.

[22]If agents use RLS learning, this eigenvalue is equal to one.

erroneous beliefs $\widetilde{M}_z$. We further consider the minimum of this measure over those among the 1000 realizations that lead to divergent GSG algorithms, and plot their contour levels in Figure 7. The lower intensity of grey depicted to the right area of the figure is associated with the higher coutour levels. For example, the white area represents the strongest misperceptions about the true second moments matrix. The darkest area corresponds to the least mistmatch of perceptions that allow a divergent GSG algorithm.[23]

The results of this exercise point in the same direction. More active monetary policy precludes a divergent GSG learning algorithm, because the necessary mismatch of beliefs is stronger (the lighter areas in Figure 7). For points in the (ϕ_π, ϕ_x) space which correspond to almost zero probability of observing GSG–instability, the mismatch measure equals 10 or higher, with the few unstable points exhibiting a very large mismatch of beliefs.[24]

The simulations depicted in Figures 4 through 7 support the conditions established in Propositions 1 and 2. The faster the RLS speed of convergence, the harder it is to generate a divergent GSG-type algorithm. The higher speed of convergence corresponds to SG–learning stability, and a lower probability of finding a second moments matrix, $\widetilde{M}_z$, with bigger misperceptions necessary to generate a divergent GSG algorithm. Thus, the two 'refinements' to the concept of E-stability that we propose, the faster mean dynamics RLS speed of convergence and the greater robustness within a class of RLS/SG/GSG learning algorithms, do not need to generate trade-offs and can be satisfied simultaneously.

We summarize the results for $\rho = 0.9$ in Table 1. We also perform some sensitivity analysis of our simulation results. Table 2 presents the results for

[23]This exercise is in the spirit of Tetlow and von zur Muehlen (2009) who model agents with RLS learning but commit errors in the ALM. They study the minimum pertubation such that the resulting algorithm is divergent.

[24]In the simple case of a constant and iid shock, the value of 10 means agents perceive the shock as being 10 times more volatile than in "reality," which is indeed a severe misperception.

the same set of parameter values, but $\rho = 0.6$. In the presence of a highly persistent cost-push shock, a temporary increase in inflation might result in increased inflation expectations which would remain elevated for a prolonged period of time, and induce actual inflation persistence as well. The convergence speed under adaptive learning decreases with the higher persistence of the shocks, and thus any initial deviation in either actual inflation or expected inflation takes longer to die out. If the shock persistence is *lower* than in the baseline model calibration, in accordance with Propositions 1 and 2 and the results of this section, we expect higher RLS convergence speeds to lead to larger area of SG–stability, and lower probability of finding a GSG-unstable algorithm.

In Tables 1 and 2 the speed of convergence is a decreasing function of the cost-push shock persistence.[25] The area of SG-instability disappears completely for values of persistence less than 0.6 (Table 2, column 3). For $\rho = 0.6$, the probability of GSG-stability is everywhere above 0.99, and it becomes essentially unity for monetary policy with $\phi > 1.2$, which is less active than the optimal under REE. In the baseline calibration with persistence shock of 0.9, only a Taylor rule with ϕ_π as high as 3.5 guarantees GSG-stability.[26]

Finally, we comment on the known examples in this context where E–stability and SG–stability conditions are found to be equivalent. This is true in univariate, purely forward–looking, cobweb-type models discussed in Evans and Honkapohja (2001), page 37. The Jacobian J of the E–stability ODE is given by $(\alpha - 1)$ times the identity matrix. Pre–multiplying J with a positive definite $M(\Phi)$ therefore cannot affect stability, and thus SG–stability is equivalent to E–stability. Note that eigenvectors of an identity matrix are orthonormal. In a multivariate extension of the cobweb model studied by

[25]Compare the entries for the same ϕ_π and ϕ_x in Tables 1 and 2.

[26]Figures which illustrate the sensitivity to the shock persistence are not reported and are available upon request.

Evans and Honkapohja (1998), the eigenvectors of $\Omega \cdot J$ can be expressed as $m \otimes j$, where m is some eigenvector of the second moments matrix M_z and j is an eigenvector of J. Since the eigenvectors of a symmetric matrix are orthogonal, certain eigenvectors of $\Omega \cdot J$ are orthogonal as well. We conjecture that a variant of the geometric condition stated in Proposition 1 continues to hold in higher dimensions, namely the equivalent orthogonality condition is inconsistent with $\Omega \cdot J$ turning unstable while J is stable. The equivalence of E–stability and SG–stability conditions for univariate and multivariate cobweb-type models agree with our results.

6. Conclusion

While under recursive least-squares learning the dynamics of linear and some nonlinear models converge to E–stable rational expectations equilibria, recent examples argue that E–stability is not a sufficient condition for SG–stability. We establish that there is a close relationship between the learnability of E–stable equilibria and the speed of convergence of the RLS learning algorithm. In the 2×2 case, we give conditions which ensure that fast mean-dynamics speed of convergence implies learnability under a broad set of learning algorithms of the RLS/SG/GSG class. This is a refinement of the set of E–stable REE with properties such that learnability is achieved even if agents' learning is misspecified asymptotically relative to RLS.

In addition, we quantify the significance of the RLS speed of convergence for learnability under alternative learning algorithms. Evans and Honkapohja (2006) show that optimal monetary policy under commitment leads to expectational stability in private agents' learning. We provide evidence that such an E–stable REE might fail to obtain its GSG learning stability when agents have misperceptions about the true parameter values of the model. For the lowest speeds of convergence consistent with determinate and E–stable REE, we observe up to 60% probability of GSG instability. If the agents use an algorithm other than RLS, the optimal monetary policy under commitment is also associated with approximately 20% probability of being subject to expectational instability.

APPENDIX

A. Learning Instability: The Real Eigenvalues Case

First we examine the real eigenvalues case. We investigate the conditions under which $\Delta\widetilde{J}$ has a positive eigenvalue, and is therefore unstable. Since $\left|\Delta\widetilde{J}\right| = |\Delta| \cdot \left|\widetilde{J}\right| > 0$, instability can only appear if the trace of $\Delta\widetilde{J}$, denoted by $Tr(\Delta\widetilde{J})$, is strictly positive. Write the matrix of the eigenvectors of $\widetilde{J}$ and its inverse[27] as

$$V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}, \ V^{-1} = \frac{1}{|V|}\begin{bmatrix} v_{22} & -v_{12} \\ -v_{21} & v_{11} \end{bmatrix},$$

then the matrix $\widetilde{J}$ can be written as:

$$\widetilde{J} = V\Lambda V^{-1} = \frac{1}{|V|}\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \cdot \begin{bmatrix} -\lambda_1 & 0 \\ 0 & -\lambda_2 \end{bmatrix} \cdot \begin{bmatrix} v_{22} & -v_{12} \\ -v_{21} & v_{11} \end{bmatrix}.$$

We can establish that the diagonal elements of $\widetilde{J}$ are given by:

$$\widetilde{J}_{11} = \frac{\lambda_2 v_{12} v_{21} - \lambda_1 v_{11} v_{22}}{|V|}$$

and

$$\widetilde{J}_{22} = \frac{\lambda_1 v_{12} v_{21} - \lambda_2 v_{11} v_{22}}{|V|}$$

.

The trace of the $\Delta\widetilde{J}$ is thus equal to:

$$Tr(\Delta\widetilde{J}) = \delta_1 \frac{\lambda_2 v_{12} v_{21} - \lambda_1 v_{11} v_{22}}{|V|} + \delta_2 \frac{\lambda_1 v_{12} v_{21} - \lambda_2 v_{11} v_{22}}{|V|} > 0. \tag{A.1}$$

[27] The case of a non–invertible V is not generic, and we do not consider it here.

The condition (A.1) is equivalent to:

$$\delta_1 \frac{\lambda_2 v_{12} v_{21} - \lambda_1 v_{11} v_{22}}{|V|} + \delta_2 \frac{\lambda_1 v_{12} v_{21} - \lambda_2 v_{11} v_{22}}{|V|} > 0$$

$$\frac{\delta_2}{\delta_1} \frac{\lambda_1 v_{12} v_{21} - \lambda_2 v_{11} v_{22}}{|V|} - \frac{\lambda_1 v_{11} v_{22} - \lambda_2 v_{12} v_{21}}{|V|} > 0$$

$$\frac{\delta_2}{\delta_1} \frac{v_{12} v_{21} - \frac{\lambda_2}{\lambda_1} v_{11} v_{22}}{|V|} - \frac{v_{11} v_{22} - \frac{\lambda_2}{\lambda_1} v_{12} v_{21}}{|V|} > 0.$$

Now select the direction of the eigenvectors so that $v_{12}v_{21} > 0$, denote $\Upsilon = \frac{v_{22}}{v_{21}} \frac{v_{11}}{v_{12}} = \frac{v_{22}}{v_{12}} / \frac{v_{21}}{v_{11}}$, and observe that $|V|$ equals $v_{21}v_{12}(\Upsilon - 1)$. The trace condition (A.1) becomes

$$\frac{\frac{\delta_2}{\delta_1} v_{12} v_{21}}{v_{21} v_{12} (\Upsilon - 1)} \left(1 - \frac{\lambda_2}{\lambda_1} \Upsilon\right) - \frac{v_{12} v_{21}}{v_{21} v_{12} (\Upsilon - 1)} \left(\Upsilon - \frac{\lambda_2}{\lambda_1}\right) > 0$$

or

$$\frac{1}{\Upsilon - 1} \left[\frac{\delta_2}{\delta_1} \left(1 - \frac{\lambda_2}{\lambda_1} \Upsilon\right) - \left(\Upsilon - \frac{\lambda_2}{\lambda_1}\right)\right] > 0. \tag{A.2}$$

When $\Upsilon < 0$, this expression is clearly negative. Thus, learning instability requires that both eigenvectors of J after rotation into the coordinates defined by the eigenvectors of Ω are located in the same quadrant of the plane. This condition is impossible to meet if the two eigenvectors are orthogonal.

When $\Upsilon > 1$, the term in the square brackets in (A.2) is negative: it is a decreasing function of $\frac{\lambda_2}{\lambda_1}$, $\frac{\delta_2}{\delta_1}$, and Υ, reaching its maximum of 0 for $\frac{\lambda_2}{\lambda_1} = \frac{\delta_2}{\delta_1} = \Upsilon = 1$. Therefore, the whole expression (A.2) is negative for $\Upsilon > 1$.

When $0 < \Upsilon < 1$, learning instability requires

$$\frac{\delta_2}{\delta_1} \left(1 - \frac{\lambda_2}{\lambda_1} \Upsilon\right) < \Upsilon - \frac{\lambda_2}{\lambda_1} < 0.$$

This is possible only if $1 - \frac{\lambda_2}{\lambda_1} \Upsilon < 0$ or $\frac{\lambda_2}{\lambda_1} > \frac{1}{\Upsilon} > 1$, in which case the condition above can be rewritten as

$$\frac{\delta_2}{\delta_1} > \frac{\Upsilon - \frac{\lambda_2}{\lambda_1}}{1 - \frac{\lambda_2}{\lambda_1} \Upsilon} = \frac{\frac{\lambda_2}{\lambda_1} - \Upsilon}{\frac{\lambda_2}{\lambda_1} \Upsilon - 1}.$$

B. Learning Instability: The Complex Eigenvalues Case

In this case, the eigenvalues of $\widetilde{J}$ are given by $\nu \pm i\mu$, $\nu < 0$, and the corresponding eigenvectors are $w_1 \pm iw_2$. Following the same steps as in the real roots case, write

$$
\begin{aligned}
W &= \begin{bmatrix} w_{11} + iw_{12} & w_{11} - iw_{12} \\ w_{21} + iw_{22} & w_{21} - iw_{22} \end{bmatrix} \\
W^{-1} &= \frac{1}{2\,|W|} \begin{bmatrix} w_{22} + iw_{21} & -\left(w_{12} + iw_{11}\right) \\ w_{22} - iw_{21} & -\left(w_{12} - iw_{11}\right) \end{bmatrix} \\
|W| &= w_{11}w_{22} - w_{12}w_{21},
\end{aligned}
$$

$$
\begin{aligned}
W\Lambda W^{-1} &= \frac{1}{2\,|W|} \begin{bmatrix} w_{11} + iw_{12} & w_{11} - iw_{12} \\ w_{21} + iw_{22} & w_{21} - iw_{22} \end{bmatrix} \cdot \begin{bmatrix} \nu + i\mu & 0 \\ 0 & \nu - i\mu \end{bmatrix} \cdot \\
&\begin{bmatrix} w_{22} + iw_{21} & -\left(w_{12} + iw_{11}\right) \\ w_{22} - iw_{21} & -\left(w_{12} - iw_{11}\right) \end{bmatrix} \\
&= \frac{1}{2\,|W|} \begin{bmatrix} \left(\nu + i\mu\right)\left(w_{11} + iw_{12}\right) & \left(\overline{\nu + i\mu}\right)\left(\overline{w_{11} + iw_{12}}\right) \\ \left(\nu + i\mu\right)\left(w_{21} + iw_{22}\right) & \left(\overline{\nu + i\mu}\right)\left(\overline{w_{21} + iw_{22}}\right) \end{bmatrix} \cdot \\
&\begin{bmatrix} w_{22} + iw_{21} & -\left(w_{12} + iw_{11}\right) \\ \overline{w_{22} + iw_{21}} & -\left(\overline{w_{12} + iw_{11}}\right) \end{bmatrix}.
\end{aligned}
$$

The overline in the expressions denotes a complex conjugate. Finally, the two diagonal elements of $\widetilde{J}$ can be written as:

$$
\begin{aligned}
\widetilde{J}_{11} &= \frac{\operatorname{Re}\left[\left(\nu + i\mu\right)\left(w_{11} + iw_{21}\right)\left(w_{22} + iw_{21}\right)\right]}{|W|}, \\
\widetilde{J}_{22} &= -\frac{\operatorname{Re}\left[\left(\nu + i\mu\right)\left(w_{12} + iw_{22}\right)\left(w_{21} + iw_{11}\right)\right]}{|W|},
\end{aligned}
$$

which reduces to

$$\widetilde{J}_{11} = \nu - \mu \frac{w_{11}w_{12} + w_{21}w_{22}}{|W|},$$
$$\widetilde{J}_{22} = \nu + \mu \frac{w_{11}w_{12} + w_{21}w_{22}}{|W|}.$$

The trace of $\Delta\widetilde{J}$ then is given by

$$Tr(\Delta\widetilde{J}) = \nu\left(\delta_2 + \delta_1\right) + \mu \frac{w_{11}w_{12} + w_{21}w_{22}}{|W|}\left(\delta_2 - \delta_1\right) > 0 \tag{B.3}$$

and should be positive for the instability to occur.

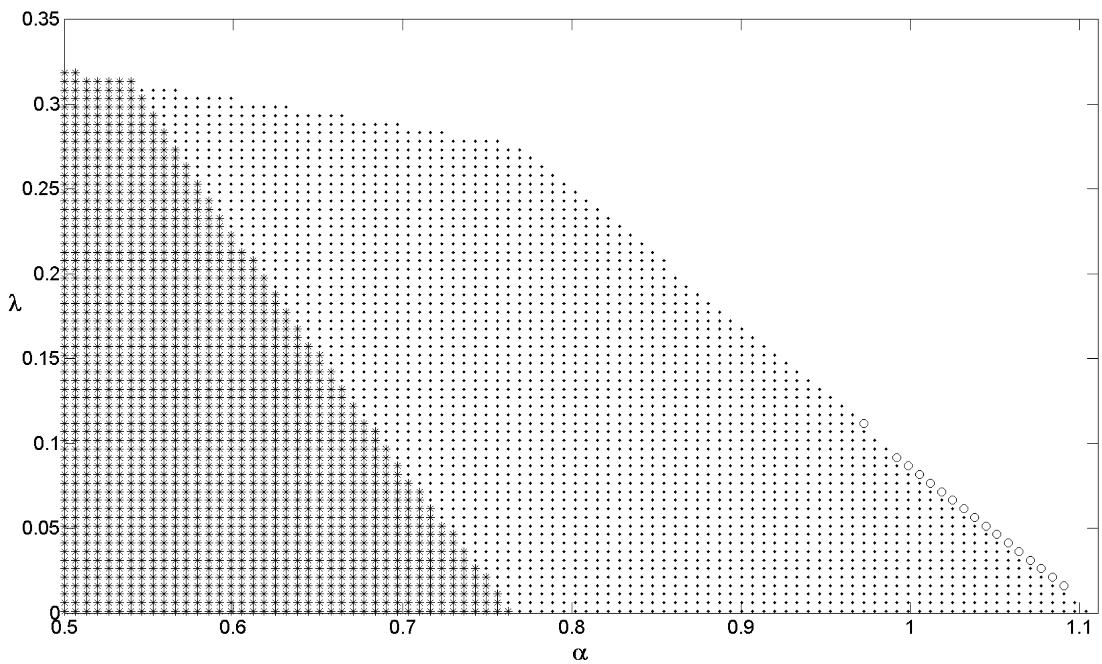
Let $\widetilde{W} = w_{11}w_{12} + w_{21}w_{22}$ and recall that ν is negative. Then (B.3) is equivalent to

$$\mu \frac{w_{11}w_{12} + w_{21}w_{22}}{|W|}\left(\delta_2 - \delta_1\right) > -\nu\left(\delta_2 + \delta_1\right),$$
$$\frac{\mu}{|\nu|}\frac{\widetilde{W}}{|W|}\frac{\frac{\delta_2}{\delta_1} - 1}{\frac{\delta_2}{\delta_1} + 1} > 1.$$

This expression allows us to evaluate and relate the speed of convergence, the real part of the eigenvalues in this case, and the conditions for learning instability. It is easy to show that if $w_1 \perp w_2$, then $\frac{\widetilde{W}}{|W|} = 0$ making (9) impossible to satisfy.

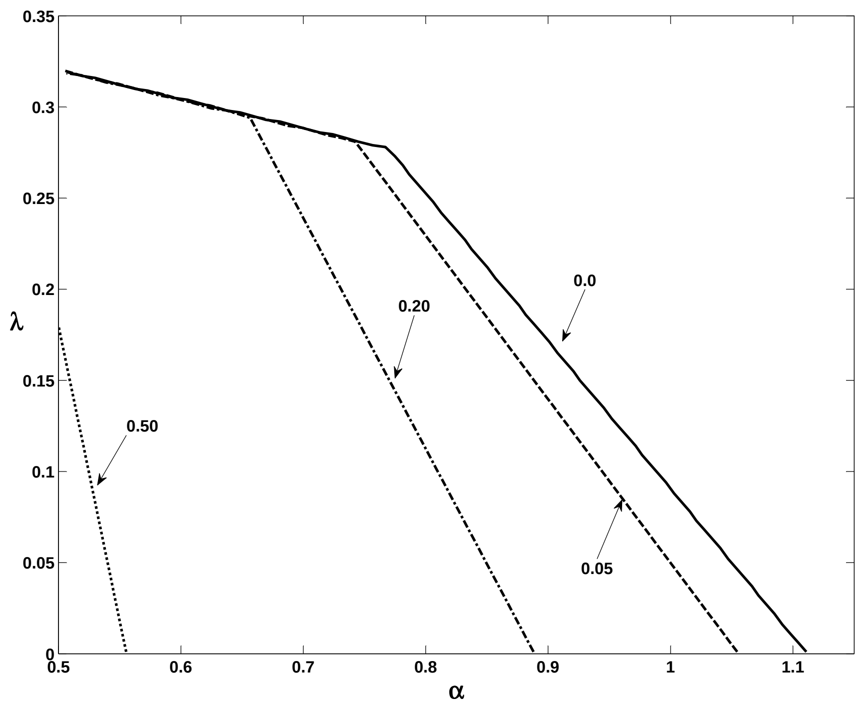
C. Figures and Tables

Figure 1: SG and H–stability



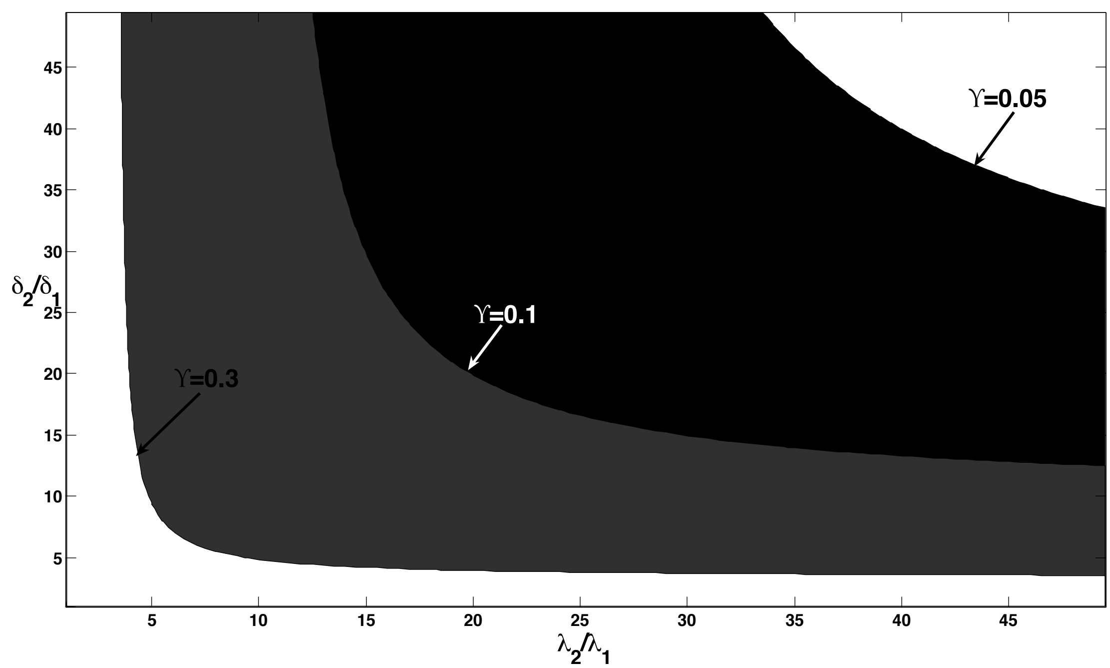
Note: The asterisks represent E–stable equilibria for which the sufficient condition of H–stability is satisfied. The dots show SG–stable equilibria which do not satisfy negative quasi-definiteness. The empty circles are SG-unstable equilibria.

Figure 2: The RLS mean-dynamics speed of convergence



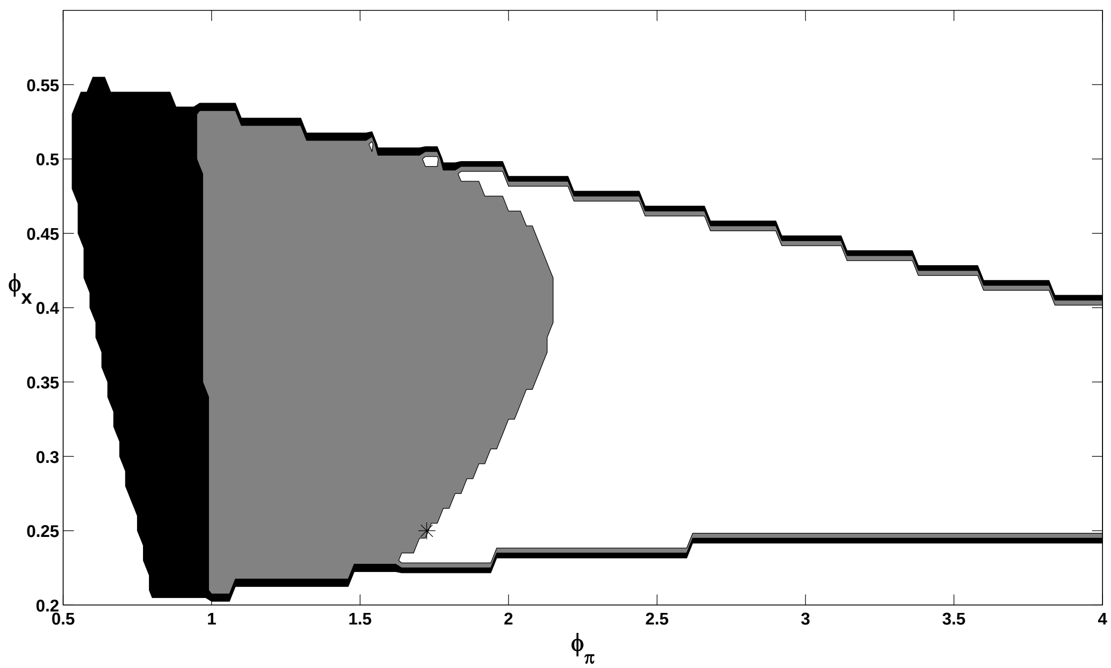
Note: The arrows point to the contour levels of the speed of convergence.

Figure 3: The necessary and sufficient conditions Proposition 1



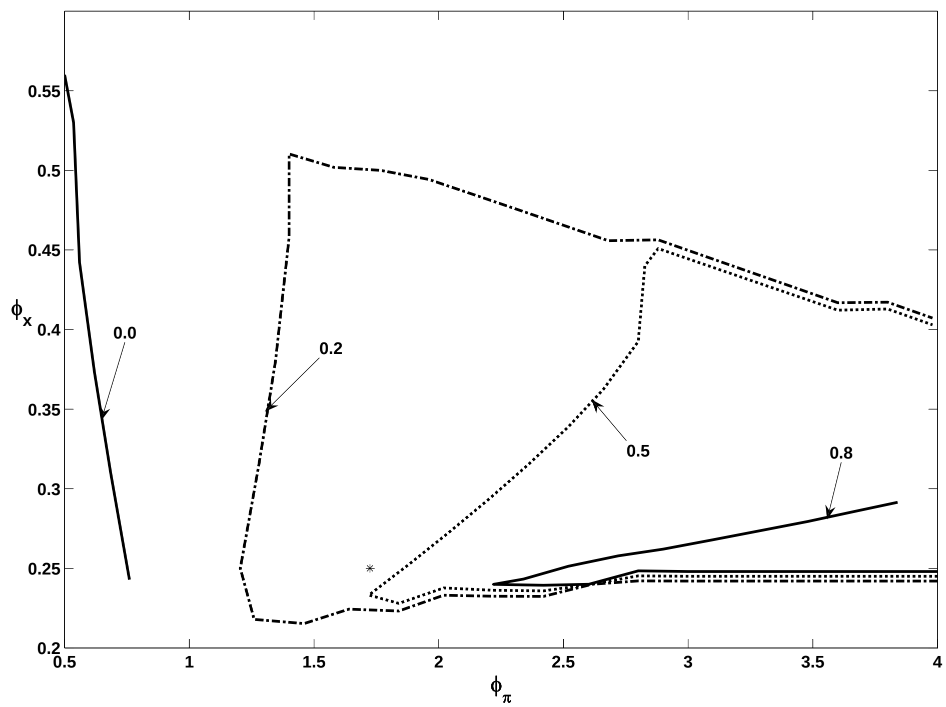
Note: The white area in the upper right corner is where (8) holds for $\Upsilon \geq 0.05$. The black area: (8) holds for $0.10 \geq \Upsilon \geq 0.05$. The grey area: (8) holds for $0.3 \geq \Upsilon \geq 0.10$.

Figure 4: Determinacy and SG–stability: Monetary policy under commitment $\rho = 0.9$



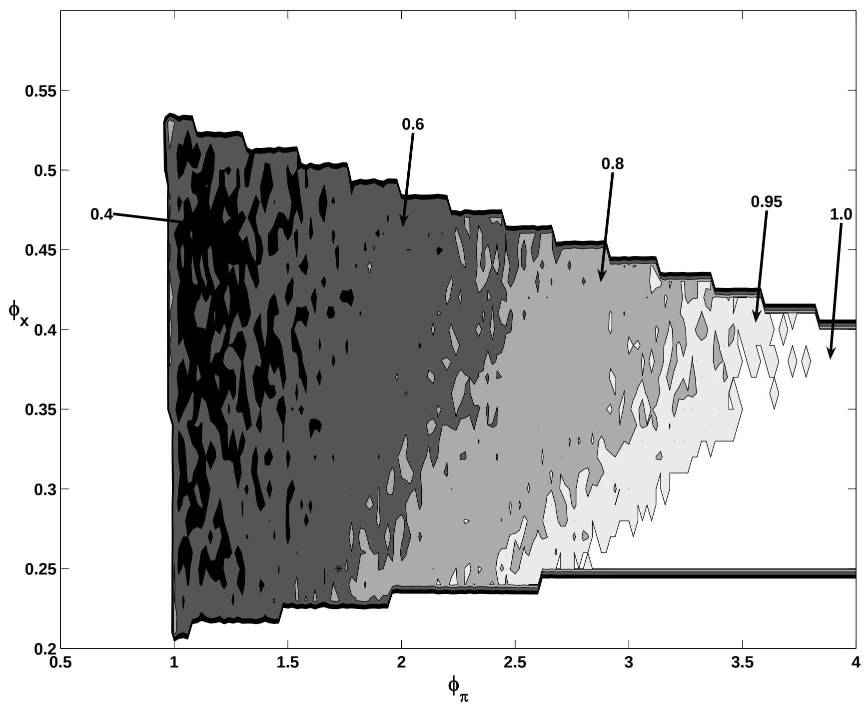
Note: All points within the outer contour are E–stable equilibria. The black area is indeterminate. The grey area is determinate but SG-unstable. The white area is determinate and SG-stable. The asterisks represents the optimal monetary policy under commitment.

Figure 5: The RLS mean-dynamics speed of convergence: Monetary policy under commitment $\rho = 0.9$



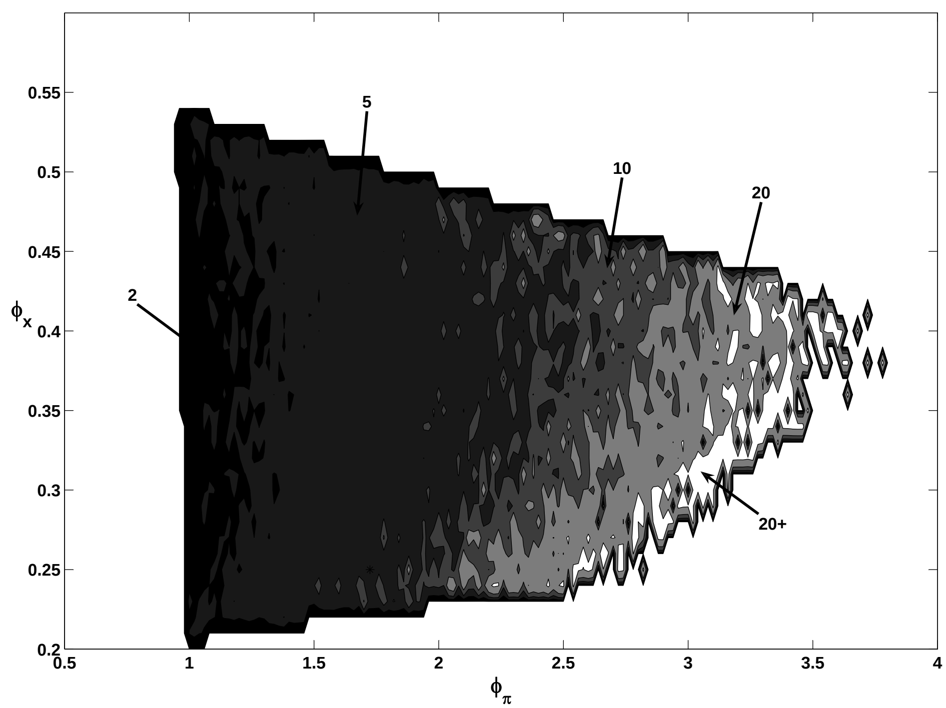
Note: The arrows point to the contour levels of the convergence speed. The asterisks represents the optimal monetary policy under commitment.

Figure 6: The probability of GSG stability: Monetary policy under commitment $\rho = 0.9$



Note: All points within the outer contour are determinate and E–stable equilibria. The probability of GSG instability is at least as large as the arrows indicate.

Figure 7: The minimum measure of misperceptions of GSG instability: Monetary policy under commitment $\rho = 0.9$



Note: All points within the outer contour are determinate and E–stable equilibria. The minimum measure of misperceptions of beliefs to get GSG instability is at least as large as the arrows indicate.

Table 1: Simulations results: Monetary policy under commitment $\rho = 0.9$

		SG-stab	Speed	GSG-stab Prob	Min Dist
$\phi_x = 0.25$					
	$\phi_\pi = 1.5$	-	0.33	0.73	2.71
	$\phi_\pi = 2.0$	+	0.56	0.83	4.15
	$\phi_\pi = 2.5$	+	0.80	0.93	16.00
$\phi_x = 0.35$					
	$\phi_\pi = 1.5$	-	0.25	0.64	2.17
	$\phi_\pi = 2.0$	-	0.38	0.74	3.59
	$\phi_\pi = 2.5$	+	0.48	0.84	6.56
$\phi_x = 0.45$					
	$\phi_\pi = 1.5$	-	0.22	0.70	3.79
	$\phi_\pi = 2.0$	-	0.33	0.71	5.10
	$\phi_\pi = 2.5$	+	0.43	0.76	5.72

Note: SG-stab: SG–stability. Speed: RLS convergence speed. GSG-stab Prob: the probability of GSG–stability. Min Dist: the minimum measure of misperceptions necessary to get GSG instability.

Table 2: Simulations results: Monetary policy under commitment $\rho = 0.6$

		SG-stab	Speed	GSG-stab Prob	Min Dist
$\phi_x = 0.25$					
	$\phi_\pi = 1.5$	+	0.59	1	N/A
	$\phi_\pi = 2.0$	+	0.77	1	N/A
	$\phi_\pi = 2.5$	+	0.97	1	N/A
$\phi_x = 0.35$					
	$\phi_\pi = 1.5$	+	0.51	1	N/A
	$\phi_\pi = 2.0$	+	0.60	1	N/A
	$\phi_\pi = 2.5$	+	0.68	1	N/A
$\phi_x = 0.45$					
	$\phi_\pi = 1.5$	+	0.49	1	N/A
	$\phi_\pi = 2.0$	+	0.57	1	N/A
	$\phi_\pi = 2.5$	+	0.64	1	N/A

Note: SG-stab: SG–stability. Speed: RLS convergence speed. GSG-stab Prob: the probability of GSG–stability. Min Dist: the minimum measure of misperceptions necessary to get GSG instability.

References

Arrow, K. J. (1974). Stability independent of adjustment speed. In G. Horwich and P. A. Samuelson (Eds.), *Trade, Stability, and Macroeconomics*, pp. 181–201. New York: Academic Press.

Barucci, E. and L. Landi (1997). Least mean squares learning in self-referential linear stochastic models. *Economics Letters 57*, 313–317.

Benveniste, A., M. Métivier, and P. Priouret (1990). *Adaptive Algorithms and Stochastic Approximations.* Berlin Heidelberg New York: Springer Verlag.

Bullard, J. (2006). The learnability criterion and monetary policy. *Federal Reserve Bank of St. Louis Review 88*, 203–217.

Bullard, J. and K. Mitra (2007). Determinacy, learnability, and monetary policy inertia. *Journal of Money, Credit and Banking 39*, 1177–1212.

Carlson, D. (1968). A new criterion for h–stability of complex matrices. *Linear Algebra and its Applications 1*, 59–64.

Clarida, R., J. Gali, and M. Gertler (1999). The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 1661–1707.

DeCanio, S. J. (1979). Rational expectations and learning from experience. *Quarterly Journal of Economics 94*, 47–57.

Evans, G. and B. McGough (2010). Implementing optimal monetary policy in new-keynesian models with inertia. *The B.E. Journal of Macroeconomics 10.*

Evans, G. W. (1985). Expectational stability and the multiple equilibria problem in linear rational expectations models. *Quarterly Journal of Economics 100*, 1217–1233.

Evans, G. W. and S. Honkapohja (1993). Adaptive forecasts, hysteresis, and endogenous fluctuations. *Economic Review Federal Reserve Bank of San Francisco 1*, 1–23.

Evans, G. W. and S. Honkapohja (1998). Stochastic gradient learning in the cobweb model. *Economics Letters 61*, 333–337.

Evans, G. W. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Pricenton, Oxford: Princeton University Press.

Evans, G. W. and S. Honkapohja (2003). Adaptive learning and monetary policy design. *Journal of Money, Credit and Banking 35*, 1045–1072.

Evans, G. W. and S. Honkapohja (2006). Monetary policy, expectations and commitment. *Scandinavian Journal of Economics 108*, 15–38.

Evans, G. W., S. Honkapohja, and N. Williams (2010). Generalized stochastic gradient learning. *International Economic Review 51*, 237–262.

Ferrero, G. (2007). Monetary policy, learning and the speed of convergence. *Journal of Economic Dynamics and Control 31*, 3006–3041.

Giannitsarou, C. (2005). E–stability does not imply learnability. *Macroeconomic Dynamics 9*, 276–287.

Heinemann, M. (2000). Convergence of adaptive learning and expectational stability: The case of multiple rational-expectations equilibria. *Macroeconomic Dynamics 4*, 263–288.

Horn, R. A. and C. R. Johnson (1985). *Matrix analysis*. Cambridge: Cambridge University Press.

Johnson, C. R. (1974a). Second, third, and fourth order d–stability. *Journal of Research of the National Bureau of Standards USA 78*, 11–13.

Johnson, C. R. (1974b). Sufficient conditions for d–stability. *Journal of Economic Theory 9*, 53–62.

Llosa, L.-G. and V. Tuesta (2009). Learning about monetary policy rules when the cost-channel matters. *Journal of Economic Dynamics and Control 33*, 1880–1896.

Marcet, A. and T. J. Sargent (1989). Convergence of least-squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory 48*, 337–368.

Marcet, A. and T. J. Sargent (1995). Speed of convergence of recursive least squares: Learning with autoregressive moving–average perceptions. In A. Kirman and M. Salmon (Eds.), *Learning and Rationality in Economics*, pp. 179–215. Oxford: Basil Blackwell.

Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics 54*, 2065–2082.

Tetlow, R. J. and P. von zur Muehlen (2009). Robustifying learnability. *Journal of Economic Dynamics and Control 33*, 296–316.