View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Learning Stability in Economies with Heterogenous Agents

Seppo Honkapohja
Kaushik Mitra

CESifo Working Paper No. 772
Category 6: Monetary Policy and International Finance
September 2002

An electronic version of the paper may be downloaded
- *from the SSRN website:* *www.SSRN.com*
- *from the CESifo website:* *www.CESifo.de*

CESifo *Working Paper No. 772*

LEARNING STABILITY IN ECONOMIES WITH HETEROGENOUS AGENTS

Abstract

An economy exhibits structural heterogeneity when the forecasts of different agents have different effects on the determination of aggregate variables. Various forms of structural heterogeneity can arise and we study the important case of economies in which agents' behavior depends on forecasts of aggregate variables and show how different forms of heterogeneity in structure, forecasts, and adaptive learning rules affect the conditions for convergence of adaptive learning towards rational expectations equilibrium. Results are applied to the market model with supply lags and a New Keynesian model of interest rate setting.

JEL Classification: D83, C62, E30.

Keywords: adaptive learning, expectations formation, stability of equilibrium, market model, monetary policy.

Seppo Honkapohja
Department of Economics
P.O. Box 54
FIN-00014 University of Helsinki
Finland
seppo.honkapohja@helsinki.fi

Kaushik Mitra
Department of Economics
Royal Holloway College
University of London
Egham Surrey, TW20 OEX
United Kingdom

1 Introduction

There has been a large amount of research into the implications of adaptive learning behavior in expectations formation for economic dynamics.[1] Paralleling general macroeconomics, most of the research that uses adaptive learning has been carried out in models with representative agents, i.e. in economies with *structural homogeneity.* In studies of adaptive learning the assumption of a representative agent is usually interpreted to mean that expectations and learning rules are also identical. These kinds of assumptions are made mostly for analytical convenience rather than for their realism. In this paper we reconsider stability of REE under adaptive learning when the economy exhibits a particular type of *structural heterogeneity,* in which the basic characteristics differ across consumers (and firms) and they thus respond to expectations of economy-wide aggregate variables in different ways. (This terminology is introduced in Chapter 2 of (Evans and Honkapohja 2001).)

In this kind of setting it is natural to assume that expectations of different agents can also differ. We will make the further distinction that heterogeneity in expectations can be due to different initial beliefs or the use of different learning algorithms by the agents. In contrast, structurally homogenous economies can exhibit heterogenous expectations (and this possibility is permitted in some studies, see the references below), but different agents respond to expectations in the same way in economies with structural homogeneity.

Our goal is to consider the stability of REE when both structural and expectational heterogeneity is present. The basic framework will be a multivariate linear model with two classes of agents. While the assumption of linearity is directly postulated for some models in the literature, it can be observed that most applied studies are in any case based on linearization.[2] The restriction to two classes of agents in the main analysis is done only for simplicity of exposition, and we will also state the stability conditions for economies with any finite number of different classes of agents. As already noted, heterogenous expectations can arise because of different initial beliefs or because the learning rules of the agents differ and we will analyze both possibilities.[3] We also take up the case in which one class of agents has continually RE while others are learning as this case occasionally appears in the literature.

The economy may be purely forward looking or it may also include lags of endogenous variables. Though we will work out the details in the forward looking framework,

[1] While different approaches to adaptive learning exist, probably the largest concentration of research has used what is called the statistical or econometric learning. Parameter updating is then assumed to be done using standard econometric methods such as recursive least squares estimation. (Evans and Honkapohja 2001) is a recent treatise on the subject. For overviews and surveys see e.g. (Evans and Honkapohja 1999), (Marimon 1997), (Sargent 1993) and (Sargent 1999).

[2] (Evans and Honkapohja 1995) and (Honkapohja and Mitra 2002a) show how learning stability in the linearized model implies stability in the original nonlinear model with sufficiently small shocks that are *iid* or a finite Markov chain, respectively.

[3] In independent work (Giannitsarou 2001) considers similar forms of heterogeneity under the restrictive assumption of structural homogeneity, so that the economy depends only on the average expectations of the agents.

the main results and the key analytical modifications for economies with lags will also be developed. We will use the general stability conditions in two economic applications: Muth's market model and a New Keynesian model of monetary policy. Our analysis is focused on models where different agents need to forecast a common vector of aggregate variables, which often arises in the literature. In other words, we will assume that information is symmetric between the agents. This is done for simplicity and brevity, though we conjecture that the approach can be generalized to models with informational asymmetries once the concept of equilibrium is suitably modified. Informational asymmetries are obviously a further source for heterogeneity in expectations (see e.g. (Evans and Honkapohja 2001), Chapter 13 and (Honkapohja and Mitra 2002b), Section 5 and also an early paper (Marcet and Sargent 1989a)). Naturally, further forms of heterogeneity can also be thought of. For example, some agents may be only concerned about "local" variables and the relevant local variables can differ across agents. Our analysis does not cover such cases, but the framework is still useful as testified by the applications.

In the earlier literature, the bulk of the work on econometric learning has assumed homogeneity in both expectations and structure, though there exist several studies that permit heterogenous expectations in a homogenous structure, see e.g. (Bray and Savin 1986), (Evans and Honkapohja 1997), (Evans, Honkapohja, and Marimon 2001) and (Giannitsarou 2001). In a non-stochastic setting (Grandmont 1998), Remark 2.3 suggests the use of average expectations in models with heterogenous expectations and structure. Heterogenous expectations are also present in some of the other approaches to adaptive learning.[4] Structural heterogeneity is permitted for a class of models in (Marcet and Sargent 1989a). Expectations are heterogenous in the Marcet and Sargent setup, but this arises solely from informational differences as different agents are assumed to use versions of recursive least squares (RLS) estimation as their learning algorithms. Moreover, Marcet and Sargent do not provide explicit stability conditions in terms of the structural parameters of the economy.[5]

2 The Framework

We consider a class of multivariate linear models where there are two types of agents (1 and 2) with different forecasts and with structural heterogeneity. The formal model is

[4] Other approaches to adaptive learning include the use of computational intelligence (see e.g. (Arifovic 1998)), models of discrete predictor choice (see e.g. (Brock and Hommes 1997) and (Brock and de Fontnouvelle 2000)) and eductive learning (see (Guesnerie 2002)). In addition, adaptive learning is usually a part of the so-called agent based models, see e.g. (LeBaron 2001).

[5] They employ a restrictive version of the stochastic approximation methodology by using the so-called projection facility. Its use has been criticized especially in connection with heterogenous expectations and differential information, see (Grandmont and Laroque 1991), (Grandmont 1998) and (Moreno and Walker 1994). Ways to avoid a projection facility are discussed in (Evans and Honkapohja 1998) and Chapter 6 of (Evans and Honkapohja 2001).

given by

$$y_t = \alpha + A_1 \hat{E}_t^1 y_{t+1} + A_2 \hat{E}_t^2 y_{t+1} + B w_t, \quad (1)$$
$$w_t = F w_{t-1} + \varepsilon_t. \quad (2)$$

Here y_t is $n \times 1$ vector of endogenous variables and w_t is k dimensional vector of exogenous variables that is assumed to follow a stationary VAR, so that ε_t is white noise. F is a diagonal matrix with all eigenvalues inside the unit circle.[6] For simplicity, it is assumed that F is known to the agents (if not, it could be estimated). Let $\lim_{t\to\infty} E w_t w_t' = M_w$, which is assumed to be positive definite. As for the matrices, A_1 is $n \times n$, A_2 is $n \times n$ while B is $n \times k$.

We let $\hat{E}_t^i y_{t+1}, i = 1, 2$, denote the (in general non-rational) expectations by agent i of the endogenous variables in the economy. Expectations without "^" refer to RE. Naturally, some of the endogenous variables may not be of interest to an agent i and in this case the relevant entries in the matrix A_i would be zero.

A key feature of model (1) is that both agents' characteristics and forecasts differ. If either agents or forecasts are identical, so that $A_1 = A_2$ or $\hat{E}_t^1 y_{t+1} = \hat{E}_t^2 y_{t+1}$, the model can be aggregated. In the former case the evolution of y_t depends only on average expectations, which has been analyzed in the earlier literature. In the latter case only the aggregate characteristics $A_1 + A_2$ matter and the model becomes homogenous.

In our analysis we will keep track of individual expectations as they will be stacked into vectors (instead of defining average expectations as suggested in (Grandmont 1998), Remark 2.3.) We prefer the method of stacking since the general framework is both multivariate and stochastic, and agents can have different types of algorithms for parameter updating. Our approach still allows us to relate the central results to the structure of the "average economy" as well as different schemes in parameter updating.

We will focus attention on the learnability of the fundamental or minimal state variable (MSV) solution to the class of models (1)-(2).[7] This REE takes the form

$$y_t = a + b w_t, \quad (3)$$

where the n vector a and $n \times k$ matrix b are to be computed in terms of the structural parameters of the model. We will show a bit later that the MSV solution is generically unique and it can be obtained by solving the following system of linear equations

$$a = \alpha + (A_1 + A_2)a$$
$$b = (A_1 + A_2)bF + B,$$

[6] Diagonality of F is usually without loss of generality since a non-diagonal matrix can very often be diagonalized. In that case the shocks w_t would be some linear transformations of the original fundamental shocks. Sometimes we will explicitly assume further that F is both diagonal and positive.

[7] As is well known, under certain conditions, known as indeterminacy of REE, there also exist other well behaved REE in forward looking models and these could also be studied for learnability. See e.g. (Evans and Honkapohja 2001), Part III for a discussion of the homogenous expectations case. The techniques developed in our paper can be extended to the study of learnability of the other types of REE under structural heterogeneity.

where the latter equation can be vectorized to yield a system of linear equations.

It should be noted that the framework is restrictive in that the model (1)-(2) is purely forward-looking. This is done solely to simplify the presentation of the theoretical results. In Section 6 we will extend the analysis to the case of lagged endogenous variables:

$$\begin{aligned} y_t &= \alpha + A_1 \hat{E}_t^1 y_{t+1} + A_2 \hat{E}_t^2 y_{t+1} + D y_{t-1} + B w_t, \\ w_t &= F w_{t-1} + \varepsilon_t. \end{aligned} \tag{4}$$

We also note that the corresponding static model

$$\begin{aligned} y_t &= \alpha + A_1 \hat{E}_{t-1}^1 y_t + A_2 \hat{E}_{t-1}^2 y_t + B w_t, \\ w_t &= F w_{t-1} + \varepsilon_t \end{aligned} \tag{5}$$

can be analyzed in the same way and the formal results directly apply to this case. Indeed, one of our economic applications will fit the form (5).

In the extension to $S > 2$ classes of agents the model becomes

$$y_t = \alpha + \sum_{s=1}^{S} A_s \hat{E}_t^s y_{t+1} + B w_t, \tag{6}$$

$$w_t = F w_{t-1} + \varepsilon_t. \tag{7}$$

We will summarize the convergence conditions for model (6)-(7) in of Section 4.3. We also note that it is straightforward to extend the stability results in this paper to models with proportions of agents of different types. In this setting the matrices in (1) would have the form $A_1 = \varkappa_1 \tilde{A}_1$ and $A_2 = \varkappa_2 \tilde{A}_2$ for $\varkappa_1, \varkappa_2 > 0, \varkappa_1 + \varkappa_2 = 1$. We do not provide explicit results on this last case, since it does not arise in our economic applications.

2.1 Economic Applications

Here we outline two economic models that fit our general setup.

Example 1 (Market model with structural heterogeneity)[8] The demand function for a single good is assumed to be linear and downward sloping, that is

$$d_t = l - k p_t + \varepsilon_t.$$

Here k, l are positive parameters and ε_t is a demand shock that follows the AR(1) process

$$\varepsilon_t = f \varepsilon_{t-1} + v_t,$$

where v_t is white noise with variance σ_v^2 and $|f| < 1$.

[8] The classic analysis of this model under RE and homogenous supplies was presented by (Muth 1961). Adaptive learning in the (homogenous) Muth model was studied by (Bray and Savin 1986) and (Fourgeaud, Gourieroux, and Pradel 1986). The model is sometimes called the cobweb model.

It is assumed that there are L classes of suppliers with different linear supply functions that depend on expected market price due to a production lag. Formally,

$$s_t^i = h_i + n_i \hat{E}_{t-1}^i p_t,\ i = 1, \dots, L. \tag{8}$$

where h_i, n_i are positive parameters and $\hat{E}_{t-1}^i p_t$ denotes the (in general non-rational) expectation of producer i about the market price. Expectations for period t are formed at the end of period $t-1$ before the demand shock ε_t is realized. We make the technical assumption that $fk^{-1}n_i + 1 > 0$ for all $i = 1, \dots, L$. From market clearing $d_t = \sum_{i=1}^{L} s_t^i$ we obtain the reduced form

$$p_t = k^{-1}(l - \sum_{i=1}^{L} h_i) - \sum_{i=1}^{L} k^{-1} n_i \hat{E}_{t-1}^i p_t + k^{-1}\varepsilon_t, \tag{9}$$

which is of the form (5).

The analysis of this model will be completed in Section 5, where we show that the model continues to be stable even under heterogenous learning. We will also extend the model to a case of supply externalities and show that sufficiently strong positive externalities can be a source of instability.

Example 2 (A model of monetary policy). Recent studies of monetary policy are often based on a model with representative consumer, monopolistic competition in product market and stickiness in price setting. In the literature both forward-looking and partly backward-looking or inertial setups have been used and we allow for inertia in both output and inflation. We thus consider the bivariate linear model suggested in Section 6 of (Clarida, Gali, and Gertler 1999):

$$z_t = -\phi(i_t - \hat{E}_t^P \pi_{t+1}) + (1-\theta)\hat{E}_t^P z_{t+1} + \theta z_{t-1} + g_t, \tag{10}$$

$$\pi_t = \lambda z_t + (1-\psi)\beta \hat{E}_t^P \pi_{t+1} + \psi \pi_{t-1} + u_t, \tag{11}$$

where z_t is the "output gap" i.e. the difference between actual and potential output, π_t is the inflation rate, i.e. the proportional rate of change in the price level from $t-1$ to t and i_t is the nominal interest rate. $\hat{E}_t^P \pi_{t+1}$ and $\hat{E}_t^P z_{t+1}$ denote *private sector* expectations of inflation and output gap next period. The parameters $0 \leq \psi, \theta < 1$ reflect inflation and output inertia, respectively. All the parameters in (10) and (11) are positive. $0 < \beta < 1$ is the discount rate of the representative firm.

u_t and g_t denote observable shocks that follow first order autoregressive processes:

$$\begin{pmatrix} u_t \\ g_t \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} u_{t-1} \\ g_{t-1} \end{pmatrix} + \begin{pmatrix} \hat{u}_t \\ \hat{g}_t \end{pmatrix}, \tag{12}$$

where $0 < |\mu| < 1, 0 < |\rho| < 1$ and $\hat{g}_t \sim iid(0, \sigma_g^2), \hat{u}_t \sim iid(0, \sigma_u^2)$. g_t represents shocks to government purchases as well as shocks to potential GDP. u_t represents any cost push shocks to marginal costs other than those entering through z_t.

The model is complete once an interest rate rule by the central bank, such as

$$i_t = \chi_0 + \chi_\pi \hat{E}_t^{CB} \pi_{t+1} + \chi_z \hat{E}_t^{CB} z_{t+1} + \chi_g g_t + \chi_u u_t, \tag{13}$$

is postulated. This rule is forward looking, i.e. depends on forecasts $\hat{E}_t^{CB} \pi_{t+1}, \hat{E}_t^{CB} z_{t+1}$ of inflation and output gap by the central bank. χ_i are parameters set by the central bank and they indicate how the bank responds to the values of the endogenous and exogenous variables. Interest rate rules such as (13) can arise from implementing optimal discretionary monetary policies (if $\theta = \psi = 0$), nominal GDP targeting or as a postulated Taylor-type instrument rule, see (Evans and Honkapohja 2002), (Mitra 2002) and (Bullard and Mitra 2002), respectively. (These papers consider the case where private and central bank forecasts are assumed to be identical.) Substituting (13) into (10) leads to a bivariate model of the form (4), or (1) if $\theta = \psi = 0$.

The above setting with private sector expectations and internal central bank forecasts very naturally involves heterogeneity in both expectations and economic structure. We will continue the analysis of this model in Sections 6.2 and 7.2.

3 Heterogenous Forecasts, E-Stability and RLS Learning

A mapping from the perceptions of the economic agents to the resulting temporary equilibrium of the economy has turned out to be the key relationship in the study of convergence of adaptive learning dynamics. In this section we develop the form of this mapping in the framework with heterogenous expectations and structure and establish the uniqueness of the MSV equilibrium.

It has been observed for a wide variety of different models that convergence of learning to REE (under homogenous forecasts and learning) obtains if and only if the REE satisfies certain stability conditions, known as E-stability conditions. In this section we extend the E-stability conditions for heterogenous forecasts. We then show that the same conditions govern convergence under actual real time learning as long as the two agents use learning algorithms that are asymptotically identical in a sense defined later.

3.1 E-stability Conditions

We assume that the two types of agents have different forecast functions, though they take the same parametric form. During the learning dynamics the agents have different beliefs about the parameters they are estimating, and these beliefs are adjusted over time. For given values of the parameters of the forecast function of each agent i, called the perceived law of motion (PLM) of agent i, one computes the actual law of motion (ALM) implied by the structure of the economy. E-stability is then determined by the differential equation in which the PLM parameters adjust in the direction of the ALM parameter values.

Define the vector of state variables $z_t = (1, w_t)'$ and the matrix of parameters $\varphi_i' = (a_i, b_i)$ with a_i being an n dimensional vector and b_i being an $n \times k$ matrix. Formally, we assume that the two agents have PLMs

$$y_t = a_1 + b_1 w_t = \varphi_1' z_t, \quad (14)$$
$$y_t = a_2 + b_2 w_t = \varphi_2' z_t, \quad (15)$$

with corresponding forecast functions

$$\hat{E}_t^1 y_{t+1} = a_1 + b_1 F w_t, \quad (16)$$
$$\hat{E}_t^2 y_{t+1} = a_2 + b_2 F w_t. \quad (17)$$

Note that the PLMs have the same form as the MSV solution (3), but in general a_i, b_i are not at their RE values. Inserting these forecasts into the model (1), one obtains the ALM

$$\begin{aligned} y_t &= \alpha + A_1 a_1 + A_2 a_2 + [(A_1 b_1 + A_2 b_2)F + B] w_t \\ &= [\alpha + A_1 a_1 + A_2 a_2, (A_1 b_1 + A_2 b_2)F + B] \begin{bmatrix} 1 \\ w_t \end{bmatrix} \\ &= T(\varphi_1', \varphi_2') z_t. \end{aligned} \quad (18)$$

The explicit form of the $T-$map in (18) is

$$a_1 \rightarrow \alpha + A_1 a_1 + A_2 a_2, \quad (19)$$
$$a_2 \rightarrow \alpha + A_1 a_1 + A_2 a_2, \quad (20)$$
$$b_1 \rightarrow (A_1 b_1 + A_2 b_2)F + B, \quad (21)$$
$$b_2 \rightarrow (A_1 b_1 + A_2 b_2)F + B. \quad (22)$$

We look at stability of the REE where the two agents have homogenous forecast functions, i.e., when $a_1 = a_2 = a$ and $b_1 = b_2 = b$. Before obtaining the E-stability conditions, we first show that this symmetric MSV solution is unique.

Proposition 1 *There exists a unique, symmetric equilibrium of the model (1)-(2) if the matrices $I_n - A_1 - A_2$ and $I_{nk} - F' \otimes (A_1 + A_2)$ are invertible.*

Here and in the rest of the paper I_m denotes the $m-$dimensional identity matrix. The proof of Proposition 1 is given in Appendix A.1.

We next formulate the differential equation defining E-stability

$$\frac{d\varphi_i}{d\tau} = T(\varphi_1', \varphi_2')' - \varphi_i, \ i = 1, 2. \quad (23)$$

The system involving $\dot{a}_1, \dot{a}_2$ is independent from the system for $\dot{b}_1$ and $\dot{b}_2$, and it can be written as

$$\begin{pmatrix} \dot{a}_1 \\ \dot{a}_2 \end{pmatrix} = \begin{pmatrix} A_1 - I_n & A_2 \\ A_1 & A_2 - I_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}. \quad (24)$$

The system for $\dot{b}_1, \dot{b}_2$ needs to be vectorized and it can be written as[9]

$$\begin{pmatrix} vec\dot{b}_1 \\ vec\dot{b}_2 \end{pmatrix} = \begin{pmatrix} F' \otimes A_1 - I_{nk} & F' \otimes A_2 \\ F' \otimes A_1 & F' \otimes A_2 - I_{nk} \end{pmatrix} \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} + \begin{pmatrix} vecB \\ vecB \end{pmatrix}. \quad (25)$$

We can now state the following proposition:

Proposition 2 *Consider the model (1)-(2) with the PLMs of the agents (14)-(15), their forecasts (16)-(17) and the ALM (18). The E-stability conditions extended for heterogenous expectations are the same as when the agents have homogenous forecasts. The symmetric equilibrium is E-stable if and only if the matrices* A_1+A_2-I *and* $F'\otimes(A_1+A_2)-I$ *have eigenvalues with negative real parts.*[10]

The Proof of Proposition 2 is in Appendix A.1. We remark that if F is a positive, diagonal matrix, the E-stability conditions simplify to condition that the eigenvalues of $A_1 + A_2 - I_n$ have negative real parts.

The next section will demonstrate that the stability of the system under certain forms of heterogenous learning rules obtains if and only if the above E-stability conditions are satisfied. In actual real time learning the two agents use versions of (generalized) recursive least squares in their updating of estimates of parameters which are relevant to their forecasting. However, the learning rules can start with different initial beliefs about the parameters so that they differ along the path. The E-stability conditions, therefore, govern convergence to REE even when we allow this (limited) form of heterogeneity in learning. The analysis thus shows that the stability conditions for the homogenous case are not as restrictive as they may seem - homogeneity in forecasting and learning is a good first approximation.

3.2 RLS Learning with Different Initial Beliefs

We now consider learning by agents in real time when they use versions of (generalized) recursive least squares in the updating of parameter estimates relevant to their forecasting. Assume that the perceived laws of motion (PLM) of agents 1 and 2 are, respectively,

$$y_t = a_{1,t} + b_{1,t}w_t = \varphi'_{1,t}z_t, \quad (26)$$
$$y_t = a_{2,t} + b_{2,t}w_t = \varphi'_{2,t}z_t, \quad (27)$$

where we note that the estimates of parameters, $\varphi'_{1,t}$ and $\varphi'_{2,t}$, are now time dependent.

The corresponding forecast functions are

$$\hat{E}^1_t y_{t+1} = a_{1,t} + b_{1,t}Fw_t, \quad (28)$$
$$\hat{E}^2_t y_{t+1} = a_{2,t} + b_{2,t}Fw_t. \quad (29)$$

[9] See (Evans and Honkapohja 2001), Chapter 10 for stability analysis of matrix valued systems requiring vectorization and computation of matrix differentials.

[10] Throughout the paper we ignore the non-generic cases where one or more relevant eigenvalues has a zero real part.

In this formulation the parameter estimates are assumed to depend on data up to $t-1$, but current observation on exogenous variables are allowed to be used in the forecasts. (This is typically done in the learning literature.) Using these forecasts, the ALM of y_t is then given (as before) by

$$y_t = T(\varphi'_{1,t}, \varphi'_{2,t})z_t, \tag{30}$$

where T is the map appearing in (18).

We assume in this section that both types of agents use versions of recursive least squares (RLS) but they can have different initial beliefs of the parameter estimates.[11] More specifically, agents 1 and 2 use the following learning algorithms

$$\begin{aligned}
\varphi_{1,t} &= \varphi_{1,t-1} + \gamma_{1,t}R_{1,t}^{-1}z_{t-1}(y_{t-1} - \varphi'_{1,t-1}z_{t-1})', & (31)\\
R_{1,t} &= R_{1,t-1} + \gamma_{1,t}(z_{t-1}z'_{t-1} - R_{1,t-1}), & (32)\\
\varphi_{2,t} &= \varphi_{2,t-1} + \gamma_{2,t}R_{2,t}^{-1}z_{t-1}(y_{t-1} - \varphi'_{1,t-1}z_{t-1})', & (33)\\
R_{2,t} &= R_{2,t-1} + \gamma_{2,t}(z_{t-1}z'_{t-1} - R_{2,t-1}). & (34)
\end{aligned}$$

In the system (31)-(34) the matrices $R_{i,t}, i = 1, 2$, are matrices of second moments of the state vector, which are needed to write down the estimation of parameters $\varphi_{i,t}, i = 1, 2$ when versions of least squares are employed.

Different initial beliefs can be accommodated by different initial conditions for the dynamics. The gain parameters $\gamma_{i,t} > 0$ indicate responsiveness of the change in parameter estimates to forecast errors and new data. They satisfy $\lim_{t\to\infty} \gamma_{i,t} = 0$ and $\sum \gamma_{i,t} = \infty$. RLS is the case where $\gamma_{it} = t^{-1}$. We allow for $\gamma_{1,t} \neq \gamma_{2,t}$ for the gain parameters of the learning rules and make the following assumption:

Assumption A: There exists a non-increasing positive sequence γ_t with properties:
(i) $\gamma_{i,t} \leq K_i\gamma_t$ for some constant $K_i > 0$,
(ii) $\sum \gamma_t = \infty$ and $\sum \gamma_t^p < \infty$ for some $p \geq 2$, and
(ii) $\limsup(1/\gamma_{t+1} - 1/\gamma_t) < \infty$.

We remark that these conditions on γ_t are commonly assumed in the literature.[12] Assumption A can allow various weighting schemes for data in later periods relative to early ones, see e.g. (Ljung and Söderström 1983) and (Marcet and Sargent 1989b).

However, we assume that asymptotically the gain sequences converge at the same rate, that is,

Condition 1: $\gamma_{1,t}\gamma_t^{-1} \to \delta$ and $\gamma_{2,t}\gamma_t^{-1} \to \delta$ as $t \to \infty$.

With these assumptions we have the following result:

[11] See Chapter 2 of (Evans and Honkapohja 2001) for an introductory discussion of RLS algorithms.

[12] We note that one can assume $K_i \leq 1$ without loss of generality. If γ_t satisfies Assumption A for $K_i > 1$, then one can construct another sequence $\tilde{\gamma}_t$ satisfying assumption A with the constant $K_i \leq 1, \forall i$.

Theorem 3 *Consider the model (1)-(2) with the PLMs (26)-(27), the forecasts (28)-(29), the learning algorithms (31)-(34), and the ALM (30). Assume furthermore that Assumption A and Condition 1 hold. If the symmetric equilibrium is E-stable, then the learning algorithms converge almost surely to this equilibrium from any initial conditions.*[13]

This result is in fact a special case of Theorem 4, see Section 4.1.

We remark that the initial conditions $\varphi_i, i = 1, 2$ can take any value, but for the moment matrices initial conditions should naturally be non-negative semidefinite matrices with positive diagonal elements.[14] We also note that under some (mild) regularity conditions, the RLS algorithm will converge to an E-unstable symmetric (MSV) solution with probability zero, see (Evans and Honkapohja 2001) for the details. The conclusion of this section is thus that convergence with some forms of heterogeneity in learning continues to be governed by the standard E-stability conditions.

As a technical remark we note that this kind of convergence and non-convergence results are formally established by deriving the so-called associated ordinary differential equation (ODE) of the stochastic recursive algorithm governing convergence and non-convergence of learning (see e.g. Chapters 6 and 7 of (Evans and Honkapohja 2001) or (Evans and Honkapohja 1998)). Moreover, above the ODE defining E-stability is directly linked to the associated ODE of the algorithm. More generally, the relationship between stability or instability in the associated ODE and the convergence or non-convergence of the algorithm also applies to other settings below and we will in part conduct our discussion using the ODEs. Moreover, below we will only state stability and convergence results, but it should be kept in mind that corresponding instability/nonconvergence results also exist.

4 Heterogeneity in Learning Algorithms

4.1 RLS Learning with Different Gain Sequences

We now analyze asymptotic heterogeneity in gain sequences in the learning algorithms. The formulation includes and generalizes the heterogeneities in the weights considered in the preceding section. Our formulation includes inertia in updating of forecast rules and independent random fluctuations in adaption speeds. Otherwise, we assume that the algorithms of the two agents are of the RLS type, i.e. they are given by (31)-(34).

The individual gain sequences are assumed to satisfy:

[13] While the theorem is a global result, it should be borne in mind that in specific applications the model may be a linearization around a steady state, and the study of learning is necessarily local in such settings.

[14] If at any time the resulting value of some moment matrix is singular (which almost surely can happen only a finite number of times), the algorithm must be modified accordingly, see (Evans and Honkapohja 2001), Chapter 6, Sections 6 and 7.

Condition 2: $\gamma_{i,t} = \hat{\gamma}_{i,t}\xi_{i,t}$, where the random gains $\hat{\gamma}_{i,t}$ are positive, independent of past information and across agents, and $\xi_{i,t}$ is a Bernoulli random variable equal to 0 with probability $\rho_{i,t} \in [0,1)$ and equal to 1 with probability $1-\rho_{i,t}$. $\xi_{i,t}$ is independent of past information and $\hat{\gamma}_{i,t}$. In addition, $\lim_{t\to\infty} E(\xi_{i,t}\hat{\gamma}_{i,t}/\gamma_t) = \delta_i > 0$, where the deterministic sequence γ_t satisfies Assumption A in the preceding section.

This condition allows for significant amounts of heterogeneity, including both inertia and random variation across agents, in the adaption speeds of the different agents. Heterogeneity in the formation of expectations is observed in experimental data, see for instance (Marimon and Sunder 1993) and (Evans, Honkapohja, and Marimon 2001). Possible randomness across agents in the degree of adaption is captured by $\hat{\gamma}_{i,t}$. Inertia is indicated by $\xi_{i,t}$ which can be either 0 or 1 for any agent in any time period. A similar formulation of heterogeneity in learning was suggested in (Evans, Honkapohja, and Marimon 2001). Effectively, the above condition means that the gain sequences can differ a lot between the agents and they converge (in mean) at different rates even asymptotically.

Formally, the dynamics continues to be given by the system (31)-(34), except that the gain sequences now have different properties. The technical analysis of the algorithm is outlined in Appendix A.1 and in this case the associated ODE is

$$\begin{aligned} d\varphi_1/d\tau &= \delta_1 S_1^{-1} M_z(T(\varphi_1',\varphi_2')' - \varphi_1), \\ dS_1/d\tau &= \delta_1(M_z - S_1), \\ d\varphi_2/d\tau &= \delta_2 S_2^{-1} M_z(T(\varphi_1',\varphi_2')' - \varphi_2), \\ dS_2/d\tau &= \delta_2(M_z - S_2), \end{aligned} \tag{35}$$

where the T map continues to be given by (18) and

$$lim_{t\to\infty} E z_{t-1} z_{t-1}' = M_z = \begin{pmatrix} 1 & 0 \\ 0 & M_w \end{pmatrix}. \tag{36}$$

(The existence of the limit follows from the assumption made in Section 2.) Note that M_w is a diagonal, positive definite matrix, since F in (2) was assumed to be diagonal. Since $S_1 \to M_z$ and $S_2 \to M_z$ stability is governed by the smaller differential equation

$$d\varphi_1/d\tau = \delta_1(T(\varphi_1',\varphi_2')' - \varphi_1), \tag{37}$$
$$d\varphi_2/d\tau = \delta_2(T(\varphi_1',\varphi_2')' - \varphi_2). \tag{38}$$

We first note that if $\delta_1 = \delta_2$ then the stability conditions obtained from (37)-(38) would be identical to the E-stability conditions, which proves Theorem 3.

Returning to the general case and rearranging (37)-(38), we get

$$\begin{aligned} \dot{\varphi}_1' &= [\delta_1(\alpha + A_1 a_1 + A_2 a_2 - a_1), \delta_1(A_1 b_1 F + A_2 b_2 F + B - b_1)], \\ \dot{\varphi}_2' &= [\delta_2(\alpha + A_1 a_1 + A_2 a_2 - a_2), \delta_2(A_1 b_1 F + A_2 b_2 F + B - b_2)]. \end{aligned}$$

We look at stability of the symmetric equilibrium where $a_1 = a_2 = a$ and $b_1 = b_2 = b$. The system for $\dot{a}_1$ and $\dot{a}_2$ can be written as

$$\begin{aligned}
\begin{pmatrix} \dot{a}_1 \\ \dot{a}_2 \end{pmatrix} &= \begin{pmatrix} \delta_1(A_1 - I_n) & \delta_1 A_2 \\ \delta_2 A_1 & \delta_2(A_2 - I_n) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \alpha \\ \delta_2 \alpha \end{pmatrix} \\
&= \begin{pmatrix} \delta_1 I_n & 0 \\ 0 & \delta_2 I_n \end{pmatrix} \begin{pmatrix} A_1 - I_n & A_2 \\ A_1 & A_2 - I_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \alpha \\ \delta_2 \alpha \end{pmatrix} \qquad (39) \\
&\equiv D_1 A \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \alpha \\ \delta_2 \alpha \end{pmatrix}.
\end{aligned}$$

where (39) defines the diagonal matrix D_1 and the matrix A. For stability we need the matrix $D_1 A$ to have eigenvalues with negative real parts.

The system for $\dot{b}_1$ and $\dot{b}_2$ needs to be vectorized as before to yield (ignoring constant terms)

$$\begin{aligned}
\begin{pmatrix} vec\dot{b}_1 \\ vec\dot{b}_2 \end{pmatrix} &= \begin{pmatrix} F' \otimes \delta_1 A_1 - \delta_1 I_{nk} & F' \otimes \delta_1 A_2 \\ F' \otimes \delta_2 A_1 & F' \otimes \delta_2 A_2 - \delta_2 I_{nk}) \end{pmatrix} \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} \\
&= \begin{pmatrix} \delta_1 I_{nk} & 0 \\ 0 & \delta_2 I_{nk} \end{pmatrix} \begin{pmatrix} F' \otimes A_1 - I_{nk} & F' \otimes A_2 \\ F' \otimes A_1 & F' \otimes A_2 - I_{nk} \end{pmatrix} \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} \\
&\equiv D_2 F_1 \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} \qquad (40)
\end{aligned}$$

where F_1 is defined as in (81). The eigenvalues of $D_2 F_1$ must have negative real parts for stability of the above system. We have thus obtained the following result:

Theorem 4 *Consider the model (1)-(2) under modified recursive least squares learning, given by (31)-(34), and assume Condition 2. If the matrices $D_1 A$ and $D_2 F_1$, given in (39) and (40), have eigenvalues with negative real parts, then the learning algorithms (almost surely) converge globally to the symmetric equilibrium.*

This theorem shows how the stability conditions are affected by δ_1 and δ_2 and the structure of the economy (matrices A_1, A_2 and F). In some cases it is possible to provide sufficient conditions for stability that do not depend on δ_1 and δ_2. For this we need the notion of D-stability.[15] A matrix A is said to be D-stable if the matrix DA has all eigenvalues with negative real parts for any positive diagonal matrix D. Using this concept one has the following corollary:

Corollary 5 *Consider the model (1)-(2) under RLS learning, given by (31)-(34), and assume Condition 2. If the matrices A and F_1, given in (80) and (81), are D-stable, then the learning algorithms (almost surely) converge globally to the symmetric equilibrium.*

The proof of this corollary is immediate from (39)-(40). Evidently, the requirement of D-stability is restrictive and, indeed, the monetary model of Example 2 does not satisfy this definition. However, the matrices in Example 1 do satisfy D-stability, as will be shown later in Section 5.

[15] This concept has been used earlier in the literature on Walrasian tatonnement dynamics, see (Arrow and McManus 1958), (Enthoven and Arrow 1956), and (Johnson 1974).

4.2 RLS Learning and SG Learning

We now consider the case when the agents are using quite different algorithms in their updating schemes. The broad aim is to consider settings where one class of agents is using a learning algorithm that is either more or less sophisticated than the algorithm used by the other class of agents. Specifically, we assume that there are two possible types of learning algorithms, the RLS and the stochastic gradient (SG) algorithms that the agents might use. The RLS algorithm is more commonly employed than SG in the literature. (SG algorithm is also called the least mean squares algorithm in the technical literature.)

The SG algorithm is computationally much simpler than the RLS algorithm; however the latter is more efficient from an econometric viewpoint since it uses information on the second moments of the variables. For parameter estimation of fixed exogenous stochastic processes, both the RLS and SG algorithms yield consistent estimates of parameters but the RLS, in addition, possesses some optimality properties. For instance, if the underlying shock process is *iid* normal, then the RLS estimator is minimum variance unbiased (see (Evans and Honkapohja 2001), Section 3.5 for a discussion and references to SG learning).

Formally, we assume that agent 1 updates the parameter estimates using an RLS algorithm while agent 2 updates using a stochastic gradient (SG) type algorithm. The SG algorithm is simpler than RLS as it does not make use of the matrix of second moments, see Chapter 3 of (Evans and Honkapohja 2001) for further discussion.

For agent 1 the algorithm is given by

$$\begin{aligned}\varphi_{1,t} &= \varphi_{1,t-1} + \gamma_t(\gamma_{1,t}\gamma_t^{-1})R_t^{-1}z_{t-1}(y_{t-1} - \varphi_{1,t-1}'z_{t-1})', &(41)\\ R_t &= R_{t-1} + \gamma_t(\gamma_{1,t}\gamma_t^{-1})(z_{t-1}z_{t-1}' - R_{t-1}), &(42)\end{aligned}$$

while for agent 2 it is given by

$$\varphi_{2,t} = \varphi_{2,t-1} + \gamma_t(\gamma_{2,t}\gamma_t^{-1})z_{t-1}(y_{t-1} - \varphi_{1,t-1}'z_{t-1})'. \quad (43)$$

In addition, we assume

Condition 3: $\lim_{t\to\infty}(\gamma_{1,t}\gamma_t^{-1}) \to 1$ and $\lim_{t\to\infty}(\gamma_{2,t}\gamma_t^{-1}) \to 1$.

The technical analysis of the algorithm (41), (42) and (43) is given in Appendix A.1. Stability is determined entirely by the small ODE

$$\begin{aligned}d\varphi_1/d\tau &= (T(\varphi_1', \varphi_2')' - \varphi_1),\\ d\varphi_2/d\tau &= M_z(T(\varphi_1', \varphi_2')' - \varphi_2).\end{aligned}$$

This immediately shows that the E-stability conditions are no longer sufficient for convergence of learning dynamics although they continue to be necessary. In particular, the moment matrix M_z affects the stability conditions.

By usual arguments,

$$\begin{aligned}\dot{\varphi}_1' &= [\alpha + A_1a_1 + A_2a_2 - a_1, (A_1b_1 + A_2b_2)F + B - b_1], \\ \dot{\varphi}_2' &= [\alpha + A_1a_1 + A_2a_2 - a_2, (A_1b_1 + A_2b_2)F + B - b_2]M_z \\ &= [\alpha + A_1a_1 + A_2a_2 - a_2, \{(A_1b_1 + A_2b_2)F + B - b_2\}M_w].\end{aligned} \tag{44}$$

For $\dot{a}_1, \dot{a}_2$ the equations are the same as (24) in the analysis of E-stability, while the system for $\dot{b}_1$ and $\dot{b}_2$ needs to be vectorized, which yields (ignoring constant terms)

$$\begin{aligned}\begin{pmatrix} vec\dot{b}_1 \\ vec\dot{b}_2 \end{pmatrix} &= \begin{pmatrix} F' \otimes A_1 - I_{nk} & F' \otimes A_2 \\ M_w'F' \otimes A_1 & M_w'F' \otimes A_2 - M_w' \otimes I_n \end{pmatrix} \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} \\ &= \begin{pmatrix} I_{nk} & 0 \\ 0 & M_w \otimes I_n \end{pmatrix} \begin{pmatrix} F' \otimes A_1 - I_{nk} & F' \otimes A_2 \\ F' \otimes A_1 & F' \otimes A_2 - I_{nk} \end{pmatrix} \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} \\ &\equiv D_wF_1 \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix},\end{aligned} \tag{45}$$

where D_w is the diagonal matrix in the second line of (45). We can then prove the following theorem:

Theorem 6 *Consider the model (1)-(2) where agent* 1 *uses recursive least squares (RLS) learning given by (41)-(42) and agent* 2 *uses the stochastic gradient algorithm (43). Assume, furthermore, Condition 3. If the matrices $A = A_1 + A_2$ and D_wF_1 have eigenvalues with negative real parts, then the learning dynamics (almost surely) converges globally to the symmetric equilibrium.*

We can also obtain a result analogous to Corollary 5 in Section 4.1. Since M_w is a diagonal, positive definite matrix, we have the analogy of Corollary 5:

Corollary 7 *Consider the model (1)-(2) where agent* 1 *uses recursive least squares (RLS) learning given by (41)-(42), agent* 2 *uses the stochastic gradient algorithm (43), and assume Condition 3. If A is stable (i.e. has eigenvalues with negative real parts) and F_1 is $D-$stable, then the learning dynamics (almost surely) converges globally to the symmetric equilibrium.*[16]

A common theme emerges from Corollaries 5 and 7. If both A $(= A_1 + A_2)$ and F_1 are $D-$ stable, then the learning rules converge globally to the symmetric equilibrium irrespective of whether they are characterized by differential gains asymptotically, as in Section 4.1, or are of different types as in this section.

[16] We note that that if the matrix M_w is not diagonal, then F_1 would need to be $S-$stable, see (Arrow and McManus 1958) for the definition of $S-$stability.

4.3 Remarks on the Model with $S > 2$ Classes of Agents

Here we note some extensions of the results to economies with more than two classes of agents and to global convergence of learning. Consider the model (6)-(7) with S classes of agents. The E-stability condition is that the eigenvalues of the matrices

$$\sum_{s=1}^{S} A_s - I_n \text{ and } F' \otimes \sum_{s=1}^{S} A_s - I_{nk}$$

have negative real parts. This is also the convergence condition in the case of heterogenous initial beliefs but identical learning rules. If agents' learning rules have different gain sequences $\gamma_{s,t}$, the stability condition is that the matrices

$$D_1 A \equiv \begin{pmatrix} \delta_1 I_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \delta_S I_n \end{pmatrix} \begin{pmatrix} A_1 - I_n & \cdots & A_S \\ \vdots & \ddots & \vdots \\ A_1 & \cdots & A_S - I_n \end{pmatrix} \text{ and} \tag{46}$$

$$D_2 F_1 \equiv \begin{pmatrix} \delta_1 I_{nk} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \delta_S I_{nk} \end{pmatrix} \begin{pmatrix} F' \otimes A_1 - I_{nk} & \cdots & F' \otimes A_S \\ \vdots & \ddots & \vdots \\ F' \otimes A_1 & \cdots & F' \otimes A_S - I_{nk} \end{pmatrix} \tag{47}$$

have eigenvalues with negative real parts, where $\delta_s = \lim_{t\to\infty} E(\gamma_{s,t}/\gamma_t)$. D-stability of A and F_1 continues to be a sufficient condition for stability. In the case where some agents use RLS rules and others SG rules the stability condition is that the real parts of the eigenvalues of the matrices

$$\sum_{s=1}^{S} A_s - I_n \text{ and}$$

$$QF_1 \equiv \begin{pmatrix} Q_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & Q_S \end{pmatrix} \begin{pmatrix} F' \otimes A_1 - I_{nk} & \cdots & F' \otimes A_S \\ \vdots & \ddots & \vdots \\ F' \otimes A_1 & \cdots & F' \otimes A_S - I_{nk} \end{pmatrix}$$

are negative, where $Q_s = I_{nk}$ or $M_w \otimes I_n$ if agent s is using RLS or SG, respectively. Stability of matrix A and D-stability of matrix F_1 are a sufficient condition for convergence in this case.

5 Application to the Market Model

We now apply our results to the market model in Example 1. Note that the market model is of the form (5) with $A_s = -k^{-1} n_s$, $s = 1, \ldots, L$.

Assume that the suppliers have PLMs of the form

$$\hat{E}^s_{t-1} p_t = a_s + b_s \varepsilon_{t-1}.$$

Substituting these into the ALM (9) we get the reduced form

$$\begin{aligned} p_t &= k^{-1}(l - \sum_{s=1}^{L} h_s) - \sum_{s=1}^{L} k^{-1} n_s (a_s + b_s \varepsilon_{t-1}) + k^{-1}\varepsilon_t \\ &= k^{-1}(l - \sum_{s=1}^{L} h_s - \sum_{s=1}^{L} n_s a_s) - k^{-1}(\sum_{s=1}^{L} n_s b_s - f)\varepsilon_{t-1} + k^{-1} v_t. \end{aligned}$$

The implied forecasts are $(i = 1, 2)$

$$\hat{E}_{t-1}^{i} p_t = k^{-1}(l - \sum_{s=1}^{L} h_s - \sum_{s=1}^{L} n_s a_s) - k^{-1}(\sum_{s=1}^{L} n_s b_s - f)\varepsilon_{t-1}$$

and the T map is then

$$\begin{aligned} a_i &\to k^{-1}(l - \sum_{s=1}^{L} h_s - \sum_{s=1}^{L} n_s a_s), \\ b_i &\to -k^{-1}(\sum_{s=1}^{L} n_s b_s - f). \end{aligned}$$

The fixed points of the T map above give us the REE solution and it is easy to show that the symmetric solution (where $a_1 = \ldots = a_L \equiv \bar{a}$ and $b_1 = \ldots = b_L = \bar{b}$) is unique.

5.1 E-Stability of the REE

We now consider E-stability of this symmetric equilibrium. Dropping constant terms, for E-stability we can consider the differential equations

$$\frac{d}{d\tau}\begin{pmatrix} a_1 \\ \vdots \\ a_L \end{pmatrix} = -k^{-1}\begin{pmatrix} n_1 & \cdots & n_L \\ \vdots & \ddots & \vdots \\ n_1 & \cdots & n_L \end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_L \end{pmatrix} - \begin{pmatrix} a_1 \\ \vdots \\ a_L \end{pmatrix}, \tag{48}$$

$$\frac{d}{d\tau}\begin{pmatrix} b_1 \\ \vdots \\ b_L \end{pmatrix} = -k^{-1}\begin{pmatrix} n_1 & \cdots & n_L \\ \vdots & \ddots & \vdots \\ n_1 & \cdots & n_L \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_L \end{pmatrix} - \begin{pmatrix} b_1 \\ \vdots \\ b_L \end{pmatrix}. \tag{49}$$

The eigenvalues of

$$-k^{-1}\begin{pmatrix} n_1 & \cdots & n_L \\ \vdots & \ddots & \vdots \\ n_1 & \cdots & n_L \end{pmatrix}$$

are obviously 0 and $-k^{-1}(\sum_{s=1}^{L} n_s)$. This proves the following result.

Proposition 8 *The symmetric equilibrium of the market model (9) under heterogenous forecasts is E-stable.*

5.2 Heterogenous Learning in the Market Model

Consider now the case when the suppliers have algorithms of the form (31)-(34) with the gain sequences satisfying Condition 2, i.e. they have differential gains asymptotically. The key matrices (46) and (47) for stability in this case are

$$\begin{pmatrix} -\delta_1(k^{-1}n_1+1) & -\delta_1 k^{-1}n_2 & \cdots & -\delta_1 k^{-1}n_L \\ -\delta_2 k^{-1}n_1 & -\delta_2(k^{-1}n_2+1) & \cdots & -\delta_2 k^{-1}n_L \\ \vdots & \vdots & \ddots & \vdots \\ -\delta_L k^{-1}n_1 & -\delta_L k^{-1}n_2 & \cdots & -\delta_L(k^{-1}n_L+1) \end{pmatrix}$$

and

$$\begin{pmatrix} -\delta_1(fk^{-1}n_1+1) & -\delta_1 fk^{-1}n_2 & \cdots & -\delta_1 fk^{-1}n_L \\ -\delta_2 fk^{-1}n_1 & -\delta_2(fk^{-1}n_2+1) & \cdots & -\delta_2 fk^{-1}n_L \\ \vdots & \vdots & \ddots & \vdots \\ -\delta_L Fk^{-1}n_1 & -\delta_L Fk^{-1}n_2 & \cdots & -\delta_L(Fk^{-1}n_L+1) \end{pmatrix}$$

We can apply Corollary 5, since the matrices A and F_1 in (46) and (47) reduce to

$$\begin{pmatrix} -(k^{-1}n_1+1) & -k^{-1}n_2 & \cdots & -k^{-1}n_L \\ -k^{-1}n_1 & -(k^{-1}n_2+1) & \cdots & -k^{-1}n_L \\ \vdots & \vdots & \ddots & \vdots \\ -k^{-1}n_1 & -k^{-1}n_2 & \cdots & -(k^{-1}n_L+1) \end{pmatrix}, \tag{50}$$

$$\begin{pmatrix} -(fk^{-1}n_1+1) & -fk^{-1}n_2 & \cdots & -fk^{-1}n_L \\ -fk^{-1}n_1 & -(fk^{-1}n_2+1) & \cdots & -fk^{-1}n_L \\ \vdots & \vdots & \ddots & \vdots \\ -fk^{-1}n_1 & -fk^{-1}n_2 & \cdots & -(fk^{-1}n_L+1) \end{pmatrix}. \tag{51}$$

Both of these matrices clearly have negative diagonals. Moreover, for column i of, say, the latter matrix (51) we compute the expression

$$|F_{1,ii}| - \sum_{j\neq i} \kappa\, |F_{1,ji}| = \left(\left|fk^{-1}n_i+1\right| - (L-1)\kappa\, |f|\, k^{-1}n_i\right) > 0$$

for some $\kappa > 0$ sufficiently small, which shows that matrices A and F_1 for the market model are quasi-dominant diagonal. (For the first matrix (50) set $f = 1$ in this argument.) The matrices are, therefore, totally stable and consequently $D-$stable (see e.g. pp.165-168 of (Quirk and Saposnik 1968) for these auxiliary concepts and results).

The same argument applies in the case of RLS and SG learning by the different types of agents. Thus we can state:

Proposition 9 *The symmetric equilibrium of the market model (9) is globally stable under learning*

(i) when the agents use RLS learning with differential gains (i.e. algorithm (31)-(34) under Condition 2);
or
(ii) when some suppliers use RLS and other suppliers use the SG algorithm.

These results show that stability of the symmetric REE in the standard market model under the assumption of homogenous forecasts and learning rules is not at all restrictive. This model continues to be stable in the presence of the use of heterogenous learning rules by different heterogenous suppliers of the good.[17]

5.3 The Market Model with Externalities

Before concluding this section, we consider an extension of the market model to incorporate externalities.[18] Suppose that individual supply functions (8) take the form

$$s_t^i = h_i + n_i \hat{E}_{t-1}^i p_t + r_i \sum_{i=1}^{L} s_t^i,\ i = 1, \ldots, L,$$

where the parameter r_i measures the size (and sign) of the externality from aggregate to individual supply. Straightforward calculations show that the reduced form (9) with externalities is

$$p_t = k^{-1}(l - \sum_{i=1}^{L} \tilde{h}_i) - \sum_{i=1}^{L} k^{-1} \tilde{n}_i \hat{E}_{t-1}^i p_t + k^{-1} \varepsilon_t,$$

where $\tilde{h}_i = h_i(1 - \sum_{i=1}^{L} r_i)^{-1}$ and $\tilde{n}_i = n_i(1 - \sum_{i=1}^{L} r_i)^{-1}$ for each i.

Proposition 10 *If the aggregate externality $r = \sum_{i=1}^{L} r_i$ is positive and sufficiently strong ($r > 1$), the REE of the market model can become unstable under learning, while stability continues to prevail when the externality is weak or negative ($r < 1$).*

To show this result, we note that if the aggregate externality r is negative or only weakly positive (so that $1 - r > 0$) the signs of each $\tilde{n}_i$ is the same as those of n_i and the earlier argument can be employed. In contrast, if the externality is positive and sufficiently strong (so that $k + \tilde{n}_i < 0$ for some i), then some diagonal elements of the matrix (50) become positive. In this case the matrix (50) fails a necessary condition for $D-$stability, see p.166 of (Quirk and Saposnik 1968). Thus, there may exist values for asymptotic individual gain parameters δ_i defined in Condition 2 in Section 4.1, such with these δ_i the market model is not stable under learning. In fact, it is easy to check that with $S = 2$, if $k + \tilde{n}_1 < 0$ and $k + \tilde{n}_2 > 0$, then the market model is unstable for *any* positive δ_1, δ_2.

[17] We remark that the (Lucas 1973) Aggregate Supply model, discussed e.g. in Chapter 2 of (Evans and Honkapohja 2001), can also be shown to be stable when different suppliers use heterogenous learning rules.

[18] This model has been studied by (Evans and Guesnerie 1993), Section 5, using an eductive approach to learning.

6 Lagged Endogenous Variables

6.1 General Analysis

The analysis discussed in the previous sections can be extended to cover models with lagged endogenous variables. There are two important differences to the case of purely forward looking model: (i) there can easily exist multiple REE that are of the MSV form and some of them may not be stationary, and (ii) the sense of convergence is only local. We develop the formal analysis only briefly since it goes through with only minor changes to what was done in the previous sections.

Consider the class of models (4) of Section 2. The MSV solutions are now of the form

$$y_t = a + bw_t + cy_{t-1}. \tag{52}$$

It should be noted that such a solution may or may not be stationary and we will keep track of this issue.

Agents forecast using the PLMs

$$\begin{aligned} y_t &= a_1 + b_1 w_t + c_1 y_{t-1} = \varphi_1' z_t \\ y_t &= a_2 + b_2 w_t + c_2 y_{t-1} = \varphi_2' z_t \end{aligned}$$

where $z_t = (1, w_t', y_{t-1}')'$ and $\varphi_i' = (a_i, b_i, c_i)$ for $i = 1, 2$ in this section. (We have kept the same general notation z_t and φ_i' for the state variable and parameters. This should not cause any confusion.) The corresponding forecast functions are

$$\begin{aligned} \hat{E}_t^i y_{t+1} &= a_i + b_i F w_t + c_i \hat{E}_t^i y_t \\ &= a_i + c_i a_i + c_i^2 y_{t-1} + (b_i F + c_i b_i) w_t, \end{aligned} \tag{53}$$

where we have assumed that the contemporaneous y_t is not available in the information set of the agents. (This assumption is often used in the literature.)

Inserting the forecasts (53) into the model (4), one obtains the ALM

$$\begin{aligned} y_t &= \alpha + A_1(a_1 + c_1 a_1) + A_2(a_2 + c_2 a_2) + (A_1 c_1^2 + A_2 c_2^2 + D) y_{t-1} + \\ &\quad [A_1(b_1 F + c_1 b_1) + A_2(b_2 F + c_2 b_2) + B] w_t \\ &\equiv T(\varphi_1', \varphi_2') z_t, \end{aligned}$$

where the T map is now given by

$$\begin{aligned} a_i &\rightarrow \alpha + A_1(a_1 + c_1 a_1) + A_2(a_2 + c_2 a_2), & (54) \\ b_i &\rightarrow A_1(b_1 F + c_1 b_1) + A_2(b_2 F + c_2 b_2) + B, & (55) \\ c_i &\rightarrow A_1 c_1^2 + A_2 c_2^2 + D. & (56) \end{aligned}$$

First, we consider the REE $(\bar{a}, \bar{b}, \bar{c})$ that are symmetric fixed points of the T map. From (56) is seen that $\bar{c}$ satisfies a quadratic matrix equation and thus multiple REE of the form (52) can easily arise. Given a value for $\bar{c}$, equation for the REE value $\bar{b}$ is obtained from a linear equation (55) with $c_1 = c_2 = \bar{c}$. This equation has often a unique solution. Likewise, the equation (54) determining $\bar{a}$ with $c_1 = c_2 = \bar{c}$ is also linear.

When learning, if agents use versions of RLS, their algorithms continue to be given by (31)-(34) and we assume that the gain sequences furthermore satisfy Condition 2. As before, local stability of learning dynamics is governed by the associated ordinary differential equation

$$\begin{aligned} d\varphi_1/d\tau &= \delta_1 S_1^{-1} M_z(\varphi_1, \varphi_2)(T(\varphi_1', \varphi_2')' - \varphi_1), \\ dS_1/d\tau &= \delta_1 (M_z(\varphi_1, \varphi_2) - S_1), \\ d\varphi_2/d\tau &= \delta_2 S_2^{-1} M_z(\varphi_1, \varphi_2)(T(\varphi_1', \varphi_2')' - \varphi_2), \\ dS_2/d\tau &= \delta_2 (M_z(\varphi_1, \varphi_2) - S_2). \end{aligned}$$

The key difference from earlier analysis is that convergence of learning is only local and the attention must be restricted at stationary REE of the form (52).[19] In the formal analysis the difference to Section 4.1 is that the moment matrix $M_z(\varphi_1, \varphi_2)$ now depends on φ_1 and φ_2. Nevertheless, it can be shown that local stability is finally governed by the smaller system

$$d\varphi_1/d\tau = \delta_1 (T(\varphi_1', \varphi_2')' - \varphi_1), \tag{57}$$
$$d\varphi_2/d\tau = \delta_2 (T(\varphi_1', \varphi_2')' - \varphi_2). \tag{58}$$

When $\delta_1 = \delta_2 = 1$ as in standard RLS, stability would be governed solely by the E-stability equations. In general, however, δ_1 and δ_2 affect stability.

One can also consider the case when agent 1 uses RLS and agent 2 the SG algorithm in their estimation, i.e., the learning algorithms are given by (41), (42), and (43). Replicating the arguments in Section 4.2, one can show that local stability is governed by the following system:

$$d\varphi_1/d\tau = (T(\varphi_1', \varphi_2')' - \varphi_1), \tag{59}$$
$$d\varphi_2/d\tau = M_z(\varphi_1, \varphi_2)(T(\varphi_1', \varphi_2')' - \varphi_2). \tag{60}$$

We collect these considerations in the following:

Theorem 11 *In the model (4) a stationary REE $(\bar{a}, \bar{b}, \bar{c})$ of the form (52) is locally stable under learning if it is a locally asymptotically stable equilibrium point of the differential equations*
(i) (57)-(58) when agents have different gains,
or
(ii) (59)-(60) when agent 1 uses RLS and agent 2 uses SG algorithm.

Since the model is linear, the stability conditions can be written explicitly in terms of the coefficient matrices. We will illustrate this in the next section.

[19] See Theorems 6.4 and 6.5 of (Evans and Honkapohja 2001) for the precise notions of local convergence. We note that the use of a "projection facility" would not be appropriate with heterogenous agents, see also footnote 6. We also remark that nonstationary cases can in principle be studied if the learning algorithms are based on suitably transformed analysis.

6.2 Application to the Model of Monetary Policy

We now apply Theorem 11 to obtain some new results for the model of monetary policy, Example 2 in Section 2.1. The companion paper (Honkapohja and Mitra 2002b) presents a systematic study of the stability conditions for the purely forward looking version of the model. Our interest here is to see whether inertia in output or inflation can affect the stability constraints for monetary policy.

To simplify the algebra we assume that the instrument rule for interest rate setting is the forward looking Taylor-type rule, (13). We assume that $\chi_z \geq 0$ and $\chi_\pi \geq 0$. The model can be written in the matrix form

$$\begin{aligned} \begin{pmatrix} z_t \\ \pi_t \end{pmatrix} &= \begin{pmatrix} -\phi \\ -\lambda\phi \end{pmatrix} \chi_0 + \begin{pmatrix} 1-\theta & \phi \\ \lambda(1-\theta) & (1-\psi)\beta + \lambda\phi \end{pmatrix} \begin{pmatrix} E_t^P z_{t+1} \\ E_t^P \pi_{t+1} \end{pmatrix} + \\ & \begin{pmatrix} -\phi\chi_z & -\phi\chi_\pi \\ -\lambda\phi\chi_z & -\lambda\phi\chi_\pi \end{pmatrix} \begin{pmatrix} E_t^{CB} z_{t+1} \\ E_t^{CB} \pi_{t+1} \end{pmatrix} + \\ & \begin{pmatrix} \theta & 0 \\ \lambda\theta & \psi \end{pmatrix} \begin{pmatrix} z_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} 1-\phi\chi_g & -\phi\chi_u \\ \lambda(1-\chi_g) & 1-\lambda\phi\chi_u \end{pmatrix} \begin{pmatrix} g_t \\ u_t \end{pmatrix}. \end{aligned} \quad (61)$$

We write this system in the general form

$$\begin{aligned} y_t &= \alpha + A^P \hat{E}_t^P y_{t+1} + A^{CB} \hat{E}_t^{CB} y_{t+1} + D y_{t-1} + B w_t, \\ w_t &= F w_{t-1} + v_t. \end{aligned} \quad (62)$$

where $y_t = (z_t, \pi_t)'$, $w_t = (u_t, g_t)'$ and A^P, A^{CB}, D, B denote the right hand matrices in (61), and F is the (diagonal) matrix appearing in (12), namely

$$F = \begin{pmatrix} \rho & 0 \\ 0 & \mu \end{pmatrix}.$$

For brevity, we focus only on the case where the private sector and the central bank have different gain sequences δ_P, δ_{CB} but use versions of RLS learning. The PLMs take the form $y_t = a_i + b_i w_t + c_i y_{t-1}, i = P, CB$. Dropping constants, the coefficient matrices for the linearized version of the small ODE (57)-(58) for the parameters $\{a_P, a_{CB}\}$, $\{vecb_P, vecb_{CB}\}$ and $\{vecc_P, vecc_{CB}\}$ take the form

$$\begin{pmatrix} \delta_P(A^P(I_2+\bar{c}_P) - I_2) & \delta_P A^{CB}(I_2+\bar{c}_{CB}) \\ \delta_{CB} A^P(I_2+\bar{c}_P) & \delta_{CB} A^{CB}(I_2+\bar{c}_{CB}) - I_2) \end{pmatrix},$$

$$\begin{pmatrix} \delta_P(F' \otimes A_P + I_2 \otimes A_P\bar{c}_P - I_4) & \delta_P(F' \otimes A_{CB} + I_2 \otimes A_{CB}\bar{c}_{CB}) \\ \delta_{CB}(F' \otimes A_P + I_2 \otimes A_P\bar{c}_P) & \delta_{CB}(F' \otimes A_{CB} + I_2 \otimes A_{CB}\bar{c}_{CB} - I_4) \end{pmatrix},$$

$$\begin{pmatrix} \delta_P(\bar{c}_P' \otimes A_P + I_2 \otimes A_P\bar{c}_P - I_4) & \delta_P(\bar{c}_{CB}' \otimes A_{CB} + I_2 \otimes A_{CB}\bar{c}_{CB}) \\ \delta_{CB}(\bar{c}_P' \otimes A_P + I_2 \otimes A_P\bar{c}_P) & \delta_{CB}(\bar{c}_{CB}' \otimes A_{CB} + I_2 \otimes A_{CB}\bar{c}_{CB} - I_4) \end{pmatrix}.$$

Stability under learning requires that all eigenvalues of these matrices have negative real parts.

When applying these conditions to the model of monetary policy (61), it is evident that theoretical results are not obtainable. However, the convergence conditions can still be applied in numerically calibrated models. We thus employ a calibration of the model $\phi = 1/0.157$, $\lambda = 0.024$, $\beta = 0.99$ suggested by (Woodford 1999), to which we append the following values for the policy parameters $\chi_z = 0$, $\chi_\pi = 1.1$. With these parameter values, the purely forward looking model ($\theta = \psi = 0$) with private sector and central bank forecasting is known to be stable, provided $\delta_{CB}/\delta_p \geq 0.87$, see the results in (Honkapohja and Mitra 2002b). We are interested in the effects of inertia in either output or inflation on the stability of the uniquely stationary solution. To do this we vary either θ or ψ from 0 to 1, keeping the other inertia parameter at 0.

We consider two cases: (i) δ_p/δ_{CB} is set at level for which the forward looking model is stable and (ii) δ_p/δ_{CB} is such that the forward looking model is unstable. The two tables below report whether stability or instability (S or U, respectively) prevails with different values of the inertia parameter.

Table 1. Inertia with a stable forward looking model

θ or ψ	0	0.1	0.2	0.3	0.4
z_t inertia only	S	U	U	S	S
π_t inertia only	S	U	S	S	S

Table 2. Inertia with an unstable forward looking model

θ or ψ	0	0.2	0.4	0.6	0.8
z_t inertia	U	U	U	S	S
π_t inertia	U	U	S	S	S

These numerical results suggest that (i) if the forward looking model is stable, then inertia in output or inflation can de-stabilize the economy and that (ii) sufficient inertia can yield stability when the purely forward looking model is unstable under learning. Table 1 uses $\delta_{CB}/\delta_p = .87$ and the uniquely stationary solution appears to be stable for values of (either) output or inflation inertia more than 0.4. Table 2 uses $\delta_{CB}/\delta_p = .8$ and again the uniquely stationary solution appears to be stable for values of (either) output or inflation inertia more than 0.8. Thus the influence of output and inflation inertia on stability is not straightforward and different possibilities can arise.

7 Case with Rational and Learning Agents

For brevity we conduct the analysis for the class of forward looking models (1)-(2) in Section 2 with two types of agents. We now assume that one type of agent has rational expectations (RE) while the other type of agent is learning. Such situations have occasionally been considered in the previous literature; for example (Sargent 1999), (Cho, Williams, and Sargent 2002) and (Carlstrom and Fuerst 2001) assume that private agents have RE and the Central Bank is learning .[20] We first provide general conditions for stability in this case and then apply them to the economic examples.

Consider the class of models (1)-(2) and assume now that agent of type 1 is learning via RLS, while agent of type 2 has RE at every point of time (even outside the equilibrium). Obviously, the MSV solutions continue to take the same form as before and we examine stability of this class of solutions. Assume that the agent of type 1 has the PLM and corresponding forecast

$$\begin{aligned} y_t &= a_1 + b_1 w_t = \varphi_1' z_t, \\ \hat{E}_t^1 y_{t+1} &= a_1 + b_1 F w_t. \end{aligned} \tag{63}$$

Agent 2 has RE and knows that agent 1 is learning and the influence of the learning on the actual outcome of the economy. He makes use of this knowledge in forming his own forecasts. Given the forecast of agent 1, the ALM of the economy is

$$y_t = \alpha + A_1 a_1 + A_1 b_1 F w_t + A_2 E_t^2 y_{t+1} + B w_t,$$

where we no longer use the "^" symbol for agent 2's forecast since he has RE. Agent 2 knows the above ALM and makes use of this to form his own forecast.

Guessing that the MSV solution for y_t takes the same form as before, his forecast is

$$E_t^2 y_{t+1} = a_2 + b_2 F w_t \tag{64}$$

and plugging this into the ALM yields

$$y_t = \alpha + A_1 a_1 + A_2 a_2 + (A_1 b_1 F + A_2 b_2 F + B) w_t.$$

The rational forecast for agent 2, given this ALM is,

$$E_t^2 y_{t+1} = \alpha + A_1 a_1 + A_2 a_2 + (A_1 b_1 F + A_2 b_2 F + B) F w_t \tag{65}$$

Given a_1, and b_1, we require agent 2 to have always RE at every point of time. This will be so if the coefficients in (64) and (65) are equal, i.e. if

$$a_2 = \alpha + A_1 a_1 + A_2 a_2, \tag{66}$$

$$b_2 = A_1 b_1 F + A_2 b_2 F + B. \tag{67}$$

One can then solve for a_2 and b_2 from (66)-(67) after vectorizing the latter equation. This yields

$$a_2 = (I - A_2)^{-1} (\alpha + A_1 a_1), \tag{68}$$

$$vec b_2 = (I - F' \otimes A_2)^{-1} (F' \otimes A_1 vec b_1 + vec B). \tag{69}$$

[20] (Evans, Honkapohja, and Sargent 1993) analysed the structure of equilibria with rational and boundedly rational agents in the standard overlapping generations model. See (Bomfim 2001) for references on other models of economies with heterogenous agents in terms of sophistication in forecasting.

Note that (68)-(69) determine the RE values of a_2 and b_2 as functions of a_1 and b_1. Solving (68)-(69) with $a_1 = a_2$ and $vecb_1 = vecb_2$ just gives the symmetric RE value for a_1, b_1.

The right-hand sides of (68)-(69) lead to the $T-$mapping

$$\begin{aligned} a_1 &\rightarrow A_2(I - A_2)^{-1}(\alpha + A_1 a_1) + \alpha + A_1 a_1, \\ vecb_1 &\rightarrow (F' \otimes A_2)(I - F' \otimes A_2)^{-1}(F' \otimes A_1 vecb_1 + vecB) + (F' \otimes A_1)vecb_1 + vecB. \end{aligned}$$

For E-stability, we proceed as before. Given the PLM of agent 1, stability of learning dynamics is governed by the ODE for a_1, i.e.

$$\dot{a}_1 = [A_1 + A_2(I - A_2)^{-1}A_1 - I]a_1 + \alpha + A_2(I - A_2)^{-1}\alpha \tag{70}$$

and that for b_1 given by

$$\begin{aligned} vec\dot{b}_1 &= [F' \otimes A_1 + (F' \otimes A_2)(I - F' \otimes A_2)^{-1}(F' \otimes A_1) - I]vecb_1 + \\ &\quad [(F' \otimes A_2)(I - F' \otimes A_2)^{-1} + I]vecB. \end{aligned} \tag{71}$$

The symmetric equilibrium will be globally stable under learning iff the differential equations (70)-(71) are globally asymptotically stable at the point, which proves the following proposition:

Proposition 12 *Consider the class of models (1)-(2), where agent 1 uses RLS learning and agent 2 has RE. The symmetric equilibrium of this model is globally stable under learning iff the eigenvalues of the matrices*

$$\begin{aligned} &A_1 + A_2(I - A_2)^{-1}A_1 - I, \\ &F' \otimes A_1 + (F' \otimes A_2)(I - F' \otimes A_2)^{-1}(F' \otimes A_1) - I, \end{aligned}$$

have negative real parts.

7.1 Application to the Market Model

As first application of Proposition 12 consider Example 1, the market model with two agents, for which $A_i = -k^{-1}n_i, i = 1, 2$. Suppose that agent 2 has rational expectations. In this case the stability conditions of Proposition 12 reduce to

$$\frac{-k^{-1}n_1}{1 + k^{-1}n_2} - 1, \frac{-fk^{-1}n_1}{1 + fk^{-1}n_2} - 1 < 0.$$

Thus we have:

Proposition 13 *When one class has rational expectations and the other uses RLS learning the market model with two classes of heterogenous supplies is globally stable under learning if* $fk^{-1}(n_{1+}n_2) + 1 > 0$.

We remark that the requirement $fk^{-1}(n_{1+}n_2) + 1 > 0$ is stronger than what was assumed before. However, it is still not very restrictive. For example, it is satisfied in the plausible case $f > 0$.

7.2 Application to Monetary Policy

We now apply Proposition 12 to our Example 2 on monetary policy when there is no inertia ($\theta = \psi = 0$). Appending the interest rule (13) to equations (10) and (11), the reduced form of the model takes the form

$$\begin{aligned} y_t &= \alpha + A^P \hat{E}_t^P y_{t+1} + A^{CB} \hat{E}_t^{CB} y_{t+1} + B w_t, \\ w_t &= F w_{t-1} + v_t. \end{aligned} \tag{72}$$

where

$$\begin{aligned} A^P &= \begin{pmatrix} 1 & \phi \\ \lambda & \beta + \lambda\phi \end{pmatrix}, \; A^{CB} = \begin{pmatrix} -\phi\chi_z & -\phi\chi_\pi \\ -\lambda\phi\chi_z & -\lambda\phi\chi_\pi \end{pmatrix}, \\ B &= \begin{pmatrix} 1 - \phi\chi_g & -\phi\chi_u \\ \lambda(1 - \chi_g) & 1 - \lambda\phi\chi_u \end{pmatrix}, \end{aligned}$$

which is the same as (61) or (62) with $\theta = \psi = 0$, i.e. $D = 0$.

7.2.1 Central Bank Has RE While Private Sector is Learning

When the central bank has RE and the private sector is learning, the two matrices in Proposition 12 reduce to

$$\begin{aligned} & A^P + A^{CB}(I - A^{CB})^{-1} A^P - I \\ & = (1 + \varphi\chi_z + \lambda\varphi\chi_\pi)^{-1} \begin{pmatrix} -\phi(\chi_z + \lambda\chi_\pi) & \phi(1 - \beta\chi_\pi) \\ \lambda & -[(1-\beta)(1 + \phi\chi_z) + \lambda\phi(\chi_\pi - 1)] \end{pmatrix} \end{aligned} \tag{73}$$

and

$$F' \otimes A^P + (F' \otimes A^{CB})(I - F' \otimes A^{CB})^{-1}(F' \otimes A^P) - I = \begin{pmatrix} B_\rho & 0 \\ 0 & B_\mu \end{pmatrix} \tag{74}$$

where

$$\begin{aligned} B_\rho = {} & [1 + \rho(\phi\chi_z + \lambda\phi\chi_\pi)]^{-1} \\ & \begin{pmatrix} -[1 - \rho + \rho\phi(\chi_z + \lambda\chi_\pi)] & \rho\phi(1 - \beta\rho\chi_\pi) \\ \lambda\rho & -[1 - \beta\rho + (1 - \beta\rho)\rho\phi\chi_z + \lambda\rho\phi(\chi_\pi - 1)] \end{pmatrix}, \end{aligned} \tag{75}$$

and B_μ takes the same form as B_ρ with μ replacing ρ. The trace and determinant of (73) are, respectively,

$$\begin{aligned} & -(1 + \phi\chi_z + \lambda\phi\chi_\pi)^{-1}[1 - \beta + (2 - \beta)\phi\chi_z + \lambda\phi(2\chi_\pi - 1)], \\ & \phi(1 + \phi\chi_z + \lambda\phi\chi_\pi)^{-1}[(1 - \beta)\chi_z + \lambda(\chi_\pi - 1)]. \end{aligned}$$

It is easy to check that the determinant is positive iff $(1 - \beta)\chi_z + \lambda(\chi_\pi - 1) > 0$ and this also suffices to make the trace negative. As for the matrix (74), we note that it is block diagonal so that its eigenvalues are those of B_ρ and B_μ and since the latter two matrices are, respectively, symmetric in ρ, μ it suffices to look only at B_ρ for stability. The trace and determinant of (75) are respectively

$$-[1+\rho(\phi\chi_z+\lambda\chi_\pi)]^{-1}[2-\rho(1+\beta)+\rho\phi\{(2-\beta\rho)\chi_z+\lambda(2\chi_\pi-1)\}],$$
$$[1+\rho(\phi\chi_z+\lambda\phi\chi_\pi)]^{-1}[(1-\rho)(1-\beta\rho)+\rho\phi\{(1-\beta\rho)\chi_z+\lambda(\chi_\pi-1)\}].$$

It is easy to check that $(1-\beta)\chi_z+\lambda(\chi_\pi-1)>0$ implies that the trace above is negative and determinant positive for (75). This proves the following corollary.

Corollary 14 *Assume that for the model (72), the private sector is learning via RLS while the central bank always has RE. The dynamics of the economy is then stable if and only if* $(1-\beta)\chi_z+\lambda(\chi_\pi-1)>0$.[21]

Condition $(1-\beta)\chi_z+\lambda(\chi_\pi-1)>0$ is precisely the Taylor principle that characterized learnability in (Bullard and Mitra 2002), where both the central bank and the private sector were learning via RLS with identical learning rules. The same principle determines stability also when only the private sector is learning as above. Since the central bank has now so much more information than the private sector, it is able to neutralize the destabilizing influence of the latter (which arises since A^P has an eigenvalue more than 1) by subscribing to the Taylor principle in its interest rule. See (Honkapohja and Mitra 2002b) for more on the intuition behind the stabilizing influence from the central bank and the de-stabilizing effect arising from the behavior of the private agents.

7.2.2 Central Bank is Learning While Private Sector Has RE

Consider now the situation when the central bank is learning while the private agents always have RE in the sense defined above. In this case we have

$$A^{CB}+A^P(I-A^P)^{-1}A^{CB}-I=\begin{pmatrix}\lambda^{-1}(1-\beta)\chi_z-1 & \lambda^{-1}(1-\beta)\chi_\pi\\ \chi_z & \chi_\pi-1\end{pmatrix}$$

The determinant and trace of the above matrix equal, respectively,

$$-\lambda^{-1}[(1-\beta)\chi_z+\lambda(\chi_\pi-1)],$$
$$\lambda^{-1}[(1-\beta)\chi_z+\lambda(\chi_\pi-1)]-1.$$

The determinant is positive if and only if $(1-\beta)\chi_z+\lambda(\chi_\pi-1)<0$ and this also makes the trace negative. Therefore, a necessary condition for the equilibrium to be stable is that the Taylor principle be *violated*.

As before, it can be shown that the matrix corresponding to (74) (after inter-changing the roles of A^{CB} and A^P there) is block diagonal with the diagonal matrices being

[21] This and the next proposition are not stated as global results, since the model of monetary policy is often thought to be a linearization of a nonlinear model.

symmetric in μ and ρ. The eigenvalues of the diagonal matrix corresponding to μ are given by -1 and

$$-[(1-\mu)(1-\beta\mu)-\mu\lambda\phi]^{-1}[(1-\mu)(1-\beta\mu)+\mu\phi\{(1-\beta\mu)\chi_z+\lambda(\chi_\pi-1)\}]>0$$

provided $(1-\mu)(1-\beta\mu)-\mu\lambda\phi\neq 0$. This enables us to prove the following corollary.

Corollary 15 *Assume that for the model (72), the private sector always has RE while the central bank is learning via RLS. The necessary and sufficient conditions for the symmetric equilibrium to be stable are*

$$\begin{aligned}(1-\beta)\chi_z+\lambda(\chi_\pi-1)&<0,\\ [(1-\mu)(1-\beta\mu)-\mu\lambda\phi]^{-1}[(1-\mu)(1-\beta\mu)+\mu\phi\{(1-\beta\mu)\chi_z+\lambda(\chi_\pi-1)\}]&>0,\\ [(1-\rho)(1-\beta\rho)-\rho\lambda\phi]^{-1}[(1-\rho)(1-\beta\rho)+\rho\phi\{(1-\beta\rho)\chi_z+\lambda(\chi_\pi-1)\}]&>0.\end{aligned}$$

The result in (Bullard and Mitra 2002) and in the previous section has now been turned on its head by this extreme assumption of rationality of the private sector *vis-a-vis* the central bank. We note that, in general, violation of the Taylor principle is not sufficient for stability of the equilibrium. This is because the latter two conditions in Corollary 15 depend also on μ and ρ. In fact it can be checked numerically for plausible values of parameters used in (Woodford 1999) that equilibrium may be either stable or unstable even when $(1-\beta)\chi_z+\lambda(\chi_\pi-1)<0$.

A case of stability arises when the policy does not react at all to forecasts, i.e. $\chi_z=\chi_\pi=0$. This is natural, since by assumption the private economy has already converged to the MSV REE and so the choice of the interest rate instrument rule need not then be based on considerations of stability under learning. However, we note that interest rate rules that react only to exogenous observables are problematic, as they lead to indeterminacy (and also instability under learning if in fact private agents do not have RE).

Generally, the results in this section show that the conclusions on stability under learning in the model of monetary policy are quite sensitive to the degree of rationality in the forecasting by private agents and the central bank. The assumption that one party has RE is often not an innocent simplification.

8 Concluding Remarks

Most macroeconomic models are based on the assumption of structural homogeneity, i.e. of the representative agent, and in the literature on learning this assumption is usually extended to include the learning rules of the agents. In this paper we have considered the significance of this assumption for stability of learning dynamics by studying the implications of structural heterogeneity, which is captured by the differential effect of the expectations of the different agents on the economy. The class of models we consider includes forward looking models with or without lags. Several cases of structural and expectational heterogeneity were analyzed.

We started by showing that introducing heterogeneity only in beliefs but not in learning rules has no significant consequences, as the convergence conditions are the same as in the corresponding model with homogenous expectations. This result was then reconsidered by analyzing the implications of heterogeneity in learning rules (and not only forecasts) when agents are boundedly rational and are learning about key parameters of the economy. We also briefly considered the case, where some agents have RE continuously while other agents are learning.

In general, the stability conditions for learning are affected by this kind of heterogeneity, but this is not always the case. Some standard models, which have been found to converge to REE under homogenous expectations and learning, continue to do so in the presence of heterogenous expectations and learning rules. This shows that the assumption of homogenous expectations and learning rules is not always as restrictive as it may seem at first sight.

There are, of course, models for which heterogenous learning affects the conditions for convergence of learning. An important case is the basic forward looking model of monetary policy commonly considered in the New Keynesian literature. In this paper we considered this model for two cases. The first illustration focused on the significance of inertia in output or inflation and the second considered the situation in which one class of agents has RE while the other is learning. The companion paper (Honkapohja and Mitra 2002b) provides a thorough analysis of the purely forward looking version of the model and examines to what extent heterogeneity can affect the desirability of different interest rate rules advocated in the literature.

The analysis and the results in this paper are based on the assumption of symmetric information, so that agents observe and make forecasts on the same set of "macro" variables in the economy. This setting is natural in many models, but extensions to our analysis are going to be needed for some specific settings. For example, we have not considered the learnability of non-MSV REE. Perhaps more importantly, we stress that adaptive learning in economies with asymmetric information or when different agents are concerned with different local variables should be considered further as the existing literature is far from comprehensive.

References

Arifovic, J. (1998): "Stability of Equilibria under Genetic Algorithm Adaption: an Analysis," *Macroeconomic Dynamics*, 2, 1–21.

Arrow, K. J., and M. McManus (1958): "A Note on Dynamic Stability," *Econometrica*, 26, 448–454.

Barnett, W., B. Cornet, C. D'Aspremont, J. Gabszewicz, and A. Mas-Colell (eds.) (1991): *Equilibrium Theory and Applications, Proceedings of the Sixth International Symposium in Economic Theory and Econometrics*. Cambridge University Press, Cambridge.

Bomfim, A. (2001): "Heterogenous Forecasts and Aggregate Dynamics," *Journal of Monetary Economics*, 47, 145–161.

Bray, M., and N. Savin (1986): "Rational Expectations Equilibria, Learning, and Model Specification," *Econometrica*, 54, 1129–1160.

Brock, W. A., and P. de Fontnouvelle (2000): "Expectational Diversity in Monetary Economies," *Journal of Economic Dynamics and Control*, 24, 725–759.

Brock, W. A., and C. H. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, 65, 1059–1095.

Bullard, J., and K. Mitra (2002): "Learning About Monetary Policy Rules," *Journal of Monetary Economics, forthcoming.*

Carlstrom, C. T., and T. S. Fuerst (2001): "Learning and the Central Bank," mimeo, Bowling Green State University.

Cho, I.-K., N. Williams, and T. J. Sargent (2002): "Escaping Nash Inflation," *Review of Economic Studies*, 69, 1–40.

Clarida, R., J. Gali, and M. Gertler (1999): "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37, 1661–1707.

Enthoven, A. C., and K. J. Arrow (1956): "A Theorem on Expectations and the Stability of Equilibrium," *Econometrica*, 24, 288–293.

Evans, G. W., and R. Guesnerie (1993): "Rationalizability, Strong Rationality, and Expectational Stability," *Games and Economic Behaviour*, 5, 632–646.

Evans, G. W., and S. Honkapohja (1995): "Local Convergence of Recursive Learning to Steady States and Cycles in Stochastic Nonlinear Models," *Econometrica*, 63, 195–206.

——— (1997): "Least Squares Learning with Heterogeneous Expectations," *Economic Letters*, 52, 197–201.

——— (1998): "Economic Dynamics with Learning: New Stability Results," *Review of Economic Studies*, 65, 23–44.

——— (1999): "Learning Dynamics," in (Taylor and Woodford 1999), chap. 7, pp. 449–542.

——— (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, New Jersey.

——— (2002): "Expectations and the Stability Problem for Optimal Monetary Policies," *Review of Economic Studies, forthcoming.*

Evans, G. W., S. Honkapohja, and R. Marimon (2001): "Convergence in Monetary Inflation Models with Heterogeneous Learning Rules," *Macroeconomic Dynamics*, 5, 1–31.

Evans, G. W., S. Honkapohja, and T. J. Sargent (1993): "On the Preservation of Deterministic Cycles When Some Agents Perceive Them to be Random Fluctuations," *Journal of Economic Dynamics and Control*, 17, 705–721.

Fourgeaud, C., C. Gourieroux, and J. Pradel (1986): "Learning Procedures and Convergence to Rationality," *Econometrica*, 54, 845–868.

Giannitsarou, C. (2001): "Stability Analysis of Heterogenous Learning in Self-Referential Linear Stochastic Models," mimeo, London Business School.

Grandmont, J.-M. (1998): "Expectations Formation and Stability of Large Socioeconomic Systems," *Econometrica*, 66, 741–781.

Grandmont, J.-M., and G. Laroque (1991): "Economic Dynamics with Learning: Some Instability Examples," in (Barnett, Cornet, D'Aspremont, Gabszewicz, and Mas-Colell 1991), chap. 11, pp. 247–273.

Guesnerie, R. (2002): "Anchoring Economic Predictions in Common Knowledge," *Econometrica*, 70, 439–480.

Honkapohja, S., and K. Mitra (2002a): "Learning in Nonlinear Stochastic Economies: A New Stability Result," manuscript.

——— (2002b): "Performance of Monetary Policy with Internal Central Bank Forecasting," manuscript, www.valt.helsinki.fi/RAKA/seppo.hmtl.

Johnson, C. R. (1974): "Sufficient Conditions for D-Stability," *Journal of Economic Theory*, 9, 53–62.

KREPS, D., AND K. WALLIS (eds.) (1997): *Advances in Economics and Econometrics: Theory and Applications, Volume I.* Cambridge University Press, Cambridge.

LEBARON, B. (2001): “Evolution and Time Horizons in an Agent-Based Model,” *Macroeconomic Dynamics*, 5, 225–254.

LJUNG, L., AND T. SÖDERSTRÖM (1983): *Theory and Practice of Recursive Identification.* MIT Press, Cambridge Mass.

LUCAS, JR., R. E. (1973): “Some International Evidence on Output-Inflation Tradeoffs,” *American Economic Review*, 63, 326–334.

MARCET, A., AND T. J. SARGENT (1989a): “Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information,” *Journal of Political Economy*, 97, 1306–1322.

——— (1989b): “Convergence of Least-Squares Learning Mechanisms in Self-Referential Linear Stochastic Models,” *Journal of Economic Theory*, 48, 337–368.

MARIMON, R. (1997): “Learning from Learning in Economics,” in (Kreps and Wallis 1997), chap. 9, pp. 278–315.

MARIMON, R., AND S. SUNDER (1993): “Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence,” *Econometrica*, 61, 1073–1107.

MITRA, K. (2002): “Desirability of Nominal GDP Targeting Under Adaptive Learning,” *Journal of Money, Credit, and Banking, forthcoming.*

MORENO, D., AND M. WALKER (1994): “Two Problems in Applying Ljung’s "Projection Algorithms" to the Analysis of Decentralized Learning,” *Journal of Economic Theory*, 62, 420–427.

MUTH, J. F. (1961): “Rational Expectations and the Theory of Price Movements,” *Econometrica*, 29, 315–335.

QUIRK, J., AND R. SAPOSNIK (1968): *Introduction to General Equilibrium Theory and Welfare Economics.* McGraw-Hill, New York.

SARGENT, T. J. (1993): *Bounded Rationality in Macroeconomics.* Oxford University Press, Oxford.

——— (1999): *The Conquest of American Inflation.* Princeton University Press, Princeton NJ.

TAYLOR, J., AND M. WOODFORD (eds.) (1999): *Handbook of Macroeconomics, Volume 1.* Elsevier, Amsterdam.

WOODFORD, M. (1999): “Optimal Monetary Policy Inertia,” Working paper, NBER No. 7261.

A Appendix

A.1 Proofs and Technical Details

We provide here some proofs of results in the main text.

Proof of Proposition 1: It is clear that equations (19)-(20) are symmetric in $(a_1,\, a_2)$ and equations (21)-(22) are symmetric in $(b_1,\, b_2)$, respectively. Thus $a_1 = a_2 = a$ and $b_1 = b_2 = b$ provided there exists a unique solution. For the a_1, a_2 system the solution is evidently unique if $I - A_1 - A_2$ is invertible.

The b_1, b_2 system needs to be vectorized[22]

$$
\begin{aligned}
vecb_1 &= (F' \otimes A_1)vecb_1 + (F' \otimes A_2)vecb_2 + vecB, \\
vecb_2 &= (F' \otimes A_1)vecb_1 + (F' \otimes A_2)vecb_2 + vecB.
\end{aligned}
$$

The vectorized system can be rewritten as

$$
\begin{pmatrix} I_{nk} - F' \otimes A_1 & -F' \otimes A_2 \\ -F' \otimes A_1 & I_{nk} - F' \otimes A_2 \end{pmatrix} \begin{pmatrix} vecb_1 \\ vecb_2 \end{pmatrix} = \begin{pmatrix} vecB \\ vecB \end{pmatrix},
$$

which has a unique solution provided the left hand matrix is invertible. The determinant of this matrix is easily seen to be non-zero if and only if the matrix $I - F' \otimes (A_1 + A_2)$ is invertible. Q.E.D.

Proof of Proposition 2: The differential equations defining E-stability have the explicit form:

$$
\begin{aligned}
\dot{a}_1 &= \alpha + (A_1 - I)a_1 + A_2 a_2, &\quad (76)\\
\dot{b}_1 &= A_1 b_1 F - b_1 + A_2 b_2 F + B, &\quad (77)\\
\dot{a}_2 &= \alpha + A_1 a_1 + (A_2 - I)a_2, &\quad (78)\\
\dot{b}_2 &= A_1 b_1 F + A_2 b_2 F - b_2 + B. &\quad (79)
\end{aligned}
$$

(76)-(79) are locally stable at the symmetric equilibrium if and only if the eigenvalues of the matrices on the right hand sides of (24) and (25) have negative real parts. To shorten notation, define

$$
A \equiv \begin{pmatrix} A_1 - I_n & A_2 \\ A_1 & A_2 - I_n \end{pmatrix}, \quad (80)
$$

$$
F_1 \equiv \begin{pmatrix} F' \otimes A_1 - I_{nk} & F' \otimes A_2 \\ F' \otimes A_1 & F' \otimes A_2 - I_{nk} \end{pmatrix}. \quad (81)
$$

[22] Here F' denotes the transpose. As F is assumed to be diagonal, this notation is not really necessary. We have kept the transposes as the same formulae then hold for a nonsymmetric F matrix as well.

The determinant for computing the eigenvalues of (80), $|A - mI_{2n}|$, may be simplified as follows:

$$\begin{aligned}
|A - mI_{2n}| &= \begin{vmatrix} A_1 - I_n(1+m) & A_2 \\ A_1 & A_2 - I_n(1+m) \end{vmatrix} \\
&= \begin{vmatrix} -I_n(1+m) & I_n(1+m) \\ A_1 & A_2 - I_n(1+m) \end{vmatrix} \\
&= \begin{vmatrix} -I_n(1+m) & 0 \\ A_1 & A_1 + A_2 - I_n(1+m) \end{vmatrix} \\
&= (-(1+m))^n \left|A_1 + A_2 - I_n(1+m)\right|.
\end{aligned}$$

The computation shows that A has n eigenvalues equal to -1 and the remaining eigenvalues are those of $A_1 + A_2 - I_n$. Hence, A has eigenvalues with negative real parts if and only if $A_1 + A_2 - I_n$ has the same property.

Analogously, the determinant for computing the eigenvalue of the coefficient matrix F_1 in (81) can be written as (after subtracting the second row from the first)

$$\begin{aligned}
&\begin{vmatrix} -(1+m)I_{nk} & (1+m)I_{nk} \\ F' \otimes A_1 & F' \otimes A_2 - (1+m)I_{nk} \end{vmatrix} \\
= \ &\{-(1+m)\}^{nk} \left|F' \otimes (A_1 + A_2) - (1+m)I_{nk}\right|.
\end{aligned}$$

so that F_1 has nk eigenvalues equal to -1 and the rest are the eigenvalues of $F' \otimes (A_1 + A_2) - I_{nk}$. Consequently, F_1 will have eigenvalues with negative real parts if and only if $F' \otimes (A_1 + A_2) - I_{nk}$ has so.

Finally, the result follows since when $\hat{E}_t^1 y_{t+1} = \hat{E}_t^2 y_{t+1} = \hat{E}_t y_{t+1}$, the matrix in front of the common expectations $\hat{E}_t y_{t+1}$ in (1) becomes $A_1 + A_2$, which is the homogenous case. Q.E.D.

We next develop the technical details concerning convergence of the RLS and SG algorithms of Sections 4.1 and 4.2. We will work out the details in the case where one agents uses RLS and the other SG learning and then indicate the necessary modifications for the case of different gain sequences.

Details for Theorem 6: We begin by rewriting (41), (42), and (43) as a stochastic recursive algorithm after making a timing change in (41) and (42) by defining $S_{t-1} = R_t$.[23] These algorithms start from the general form

$$\theta_t = \theta_{t-1} + \gamma_t H(\theta_{t-1}, X_t) + \gamma_t^2 \rho_t(\theta_{t-1}, X_t) \tag{82}$$

where θ_t is a vector of parameter estimates and X_t is the state vector. In our case we have $\theta_t' = (\varphi_{1,t}', \varphi_{2,t}', vec(S_t))$ and $X_t' = (1, w_t', w_{t-1}')$.

Since the $T-$map continues to be given by (18), we substitute (30) into (41) and get

$$\begin{aligned}
\varphi_{1,t} &= \varphi_{1,t-1} + \gamma_t S_{t-1}^{-1} z_{t-1}(T(\varphi_{1,t-1}', \varphi_{2,t-1}')z_{t-1} - \varphi_{1,t-1}' z_{t-1})' \\
&\quad + (\gamma_{1,t} - \gamma_t) S_{t-1}^{-1} z_{t-1}(T(\varphi_{1,t-1}', \varphi_{2,t-1}')z_{t-1} - \varphi_{1,t-1}' z_{t-1})'.
\end{aligned}$$

[23] See Chapters 7 and 8 of (Evans and Honkapohja 2001) for an exposition of the technique.

This gives us the φ_1 components of the function $H(\theta_{t-1}, X_t)$ in (82), which we denote by $H_1(z_{t-1}, \varphi_{1,t-1}, \varphi_{2,t-1}, S_{t-1})$. In other words,

$$H_1(z_{t-1}, \varphi_{1,t-1}, \varphi_{2,t-1}, S_{t-1}) = S_{t-1}^{-1} z_{t-1} z_{t-1}' (T(\varphi_{1,t-1}', \varphi_{2,t-1}')' - \varphi_{1,t-1}). \tag{83}$$

Regarding the second order in γ_t term in (82), we have

$$\rho_{\varphi,t}(\theta_{t-1}, X_t) = \frac{\gamma_{1,t} - \gamma_t}{\gamma_t^2} S_{t-1}^{-1} z_{t-1} (T(\varphi_{1,t-1}', \varphi_{2,t-1}') z_{t-1} - \varphi_{1,t-1}' z_{t-1})',$$

and the validity of the method requires that this be bounded in t. This is easily established as by Assumption A (with $K_i \leq 1$ without loss of generality) we have $\gamma_{1,t}/\gamma_t \leq 1$ $\Rightarrow \gamma_{1,t}/\gamma_t \leq 1 + K\gamma_t$ for any $K > 0 \Rightarrow \frac{\gamma_{1,t}-\gamma_t}{\gamma_t^2} \leq 1$.

For (42) we can write

$$\begin{aligned} S_t &= S_{t-1} + \gamma_t (z_t z_t' - S_{t-1}) + (\gamma_{1,t+1} - \gamma_t)(z_t z_t' - S_{t-1}) \\ &= S_{t-1} + \gamma_t (z_t z_t' - S_{t-1}) + \gamma_t^2 (\frac{\gamma_{1,t+1} - \gamma_t}{\gamma_t^2})(z_t z_t' - S_{t-1}). \end{aligned}$$

Thus the S components of the function $H(\theta_{t-1}, X_t)$ are given by

$$H_S(z_t, S_{t-1}) \equiv z_t z_t' - S_{t-1} \tag{84}$$

while the second order in γ_t term

$$\rho_{S,t}(\theta_{t-1}, X_t) = (\frac{\gamma_{1,t+1} - \gamma_t}{\gamma_t^2})(z_t z_t' - S_{t-1})$$

is bounded in t since

$$\frac{\gamma_{1,t+1} - \gamma_t}{\gamma_t^2} = \frac{\gamma_{1,t+1}}{\gamma_{t+1}} \left(\frac{\gamma_{t+1}}{\gamma_t} \right)^2 \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \leq \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t}$$

by Assumption A.

Finally, in a similar manner we get the φ_2 components of the function $H(\theta_{t-1}, X_t)$ in (43), which for future use we denote by $H_2(t, z_{t-1}, \varphi_{1,t-1}, \varphi_{2,t-1}, S_{t-1})$.

Now

$$\lim_{t \to \infty} EH_1(z_{t-1}, \varphi_1, \varphi_2, S) = S^{-1} M_z (T(\varphi_1', \varphi_2')' - \varphi_1).$$

where M_z is defined in (36). Similarly

$$\lim_{t \to \infty} EH_S(z_{t-1}, S, t) = M_z - S.$$

and

$$\lim_{t \to \infty} EH_2(z_{t-1}, \varphi_1, \varphi_2, S, t) = M_z (T(\varphi_1', \varphi_2')' - \varphi_2).$$

The associated differential equation is then defined by

$$d\theta/d\tau = h(\theta) = \lim_{t\rightarrow\infty} EH(t, \theta_{t-1}, X_t)$$

and in our case it boils down to

$$\begin{aligned} d\varphi_1/d\tau &= S^{-1}M_z(T(\varphi_1', \varphi_2')' - \varphi_1), \\ dS/d\tau &= M_z - S, \\ d\varphi_2/d\tau &= M_z(T(\varphi_1', \varphi_2')' - \varphi_2). \end{aligned}$$

Since the second set of equations is globally stable with $S \rightarrow M_z$ from any starting point, stability is determined entirely by the smaller dimensional system

$$\begin{aligned} d\varphi_1/d\tau &= (T(\varphi_1', \varphi_2')' - \varphi_1), \\ d\varphi_2/d\tau &= M_z(T(\varphi_1', \varphi_2')' - \varphi_2). \end{aligned}$$

The analysis of the stability conditions for this ODE is given in the main text, from equations (44). To prove that convergence is in fact global and takes place almost surely, we first note that the associated ODE is linear and globally stable.[24] Second, it is easy to verify that the conditions of Theorem 6.10 in (Evans and Honkapohja 2001) or Theorem 2 in (Evans and Honkapohja 1998) are satisfied, so that almost sure global convergence obtains.

Details for Theorem 4: In this case, one proceeds for both agents as above for agent 1, but one can write the gain sequence as $\gamma_{1,t} = \gamma_t(\xi_{i,t}\hat{\gamma}_{1,t}\gamma_t^{-1})$ and treat $\xi_{i,t}\hat{\gamma}_{1,t}\gamma_t^{-1}$ as an additional state variable that evolves exogenously from the rest of the system. With random gain sequences in (83) and (84) we get for agent 1

$$\lim_{t\rightarrow\infty} EH_1(\xi_{i,t}\gamma_{1,t}\gamma_t^{-1}, z_{t-1}, \varphi_1, \varphi_2, S) = \delta_1 S^{-1} M_z(T(\varphi_1', \varphi_2')' - \varphi_1)$$

and

$$\lim_{t\rightarrow\infty} EH_S(\xi_{i,t}\gamma_{1,t}\gamma_t^{-1}, z_{t-1}, S, t) = \delta_1(M_z - S)$$

since $\lim_{t\rightarrow\infty} E(\gamma_{1,t}\gamma_t^{-1}) = \delta_1$. Doing the same for agent 2 we arrive at the associated ODE for this, given as (35) in the main text. The rest of the argument is the same as in the preceding proof.

[24] The algorithm must be modified if the value of some moment matrix is singular, see (Evans and Honkapohja 2001), Chapter 6, Sections 6 and 7 for a discussion. The treatment there is for the multivariate Muth model but it can be applied also to the present context.

CESifo Working Paper Series

(for full list see www.cesifo.de.)

---

702 Ralph Chami and Gregory D. Hess, For Better or For Worse? State-Level Marital Formation and Risk Sharing, April 2002

703 Fredrik Andersson and Kai A. Konrad, Human Capital Investment and Globalization in Extortionary States, April 2002

704 Antonis Adam and Thomas Moutos, The Political Economy of EU Enlargement: Or, Why Japan is not a Candidate Country?, April 2002

705 Daniel Gros and Carsten Hefeker, Common Monetary Policy with Asymmetric Shocks, April 2002

706 Dirk Kiesewetter and Rainer Niemann, Neutral and Equitable Taxation of Pensions as Capital Income, April 2002

707 Robert S. Chirinko, Corporate Taxation, Capital Formation, and the Substitution Elasticity between Labor and Capital, April 2002

708 Frode Meland and Gaute Torsvik, Structural Adjustment and Endogenous Worker Recall Probabilities, April 2002

709 Rainer Niemann and Caren Sureth, Taxation under Uncertainty – Problems of Dynamic Programming and Contingent Claims Analysis in Real Option Theory, April 2002

710 Thomas Moutos and William Scarth, Technical Change and Unemployment: Policy Responses and Distributional Considerations, April 2002

711 Günther Rehme, (Re-)Distribution of Personal Incomes, Education and Economic Performance Across Countries, April 2002

712 Thorvaldur Gylfason and Gylfi Zoega, Inequality and Economic Growth: Do Natural Resources Matter?, April 2002

713 Wolfgang Leininger, Contests over Public Goods: Evolutionary Stability and the Free-Rider Problem, April 2002

714 Ernst Fehr and Armin Falk, Psychological Foundations of Incentives, April 2002

715 Giorgio Brunello, Maria Laura Parisi, and Daniela Sonedda, Labor Taxes and Wages: Evidence from Italy, May 2002

716 Marta Aloi and Huw Dixon, Entry Dynamics, Capacity Utilisation and Productivity in a Dynamic Open Economy, May 2002

717 Paolo M. Panteghini, Asymmetric Taxation under Incremental and Sequential Investment, May 2002

718 Ben J. Heijdra, Christian Keuschnigg, and Wilhelm Kohler, Eastern Enlargement of the EU: Jobs, Investment and Welfare in Present Member Countries, May 2002

719 Tapio Palokangas, The Political Economy of Collective Bargaining, May 2002

720 Gilles Saint-Paul, Some Evolutionary Foundations for Price Level Rigidity, May 2002

721 Giorgio Brunello and Daniela Sonedda, Labor Tax Progressivity, Wage Determination, and the Relative Wage Effect, May 2002

722 Eric van Damme, The Dutch UMTS-Auction, May 2002

723 Paolo M. Panteghini, Endogenous Timing and the Taxation of Discrete Investment Choices, May 2002

724 Achim Wambach, Collusion in Beauty Contests, May 2002

725 Dominique Demougin and Claude Fluet, Preponderance of Evidence, May 2002

726 Gilles Saint-Paul, Growth Effects of Non Proprietary Innovation, May 2002

727 Subir Bose, Gerhard O. Orosel, and Lise Vesterlund, Optimal Pricing and Endogenous Herding, May 2002

728 Erik Leertouwer and Jakob de Haan, How to Use Indicators for 'Corporatism' in Empirical Applications, May 2002

729 Matthias Wrede, Small States, Large Unitary States and Federations, May 2002

730 Christian Schultz, Transparency and Tacit Collusion in a Differentiated Market, May 2002

731 Volker Grossmann, Income Inequality, Voting Over the Size of Public Consumption, and Growth, May 2002

732 Yu-Fu Chen and Michael Funke, Working Time and Employment under Uncertainty, May 2002

733 Kjell Erik Lommerud, Odd Rune Straume, and Lars Sørgard, Downstream Merger with Oligopolistic Input Suppliers, May 2002

734 Saku Aura, Does the Balance of Power Within a Family Matter? The Case of the Retirement Equity Act, May 2002

735 Sandro Brusco and Fausto Panunzi, Reallocation of Corporate Resources and Managerial Incentives in Internal Capital Markets, May 2002

736 Stefan Napel and Mika Widgrén, Strategic Power Revisited, May 2002

737 Martin W. Cripps, Godfrey Keller, and Sven Rady, Strategic Experimentation: The Case of Poisson Bandits, May 2002

738 Pierre André Chiappori and Bernard Salanié, Testing Contract Theory: A Survey of Some Recent Work, June 2002

739 Robert J. Gary-Bobo and Sophie Larribeau, A Structural Econometric Model of Price Discrimination in the Mortgage Lending Industry, June 2002

740 Laurent Linnemer, When Backward Integration by a Dominant Firm Improves Welfare, June 2002

741 Gebhard Kirchgässner and Friedrich Schneider, On the Political Economy of Environmental Policy, June 2002

742 Christian Keuschnigg and Soren Bo Nielsen, Start-ups, Venture Capitalits, and the Capital Gains Tax, June 2002

743 Robert Fenge, Silke Uebelmesser, and Martin Werding, Second-best Properties of Implicit Social Security Taxes: Theory and Evidence, June 2002

744 Wendell Fleming and Jerome Stein, Stochastic Optimal Control, International Finance and Debt, June 2002

745 Gene M. Grossman, The Distribution of Talent and the Pattern and Consequences of International Trade, June 2002

746 Oleksiy Ivaschenko, Growth and Inequality: Evidence from Transitional Economies, June 2002

747 Burkhard Heer, Should Unemployment Benefits be Related to Previous Earnings?, July 2002

748 Bas van Aarle, Giovanni Di Bartolomeo, Jacob Engwerda, and Joseph Plasmans, Staying Together or Breaking Apart: Policy-makers' Endogenous Coalitions Formation in the European Economic and Monetary Union, July 2002

749 Hans Gersbach, Democratic Mechanisms: Double Majority Rules and Flexible Agenda Costs, July 2002

750 Bruno S. Frey and Stephan Meier, Pro-Social Behavior, Reciprocity or Both?, July 2002

751 Jonas Agell and Helge Bennmarker, Wage Policy and Endogenous Wage Rigidity: A Representative View From the Inside, July 2002

752 Edward Castronova, On Virtual Economies, July 2002

753 Rebecca M. Blank, U.S. Welfare Reform: What's Relevant for Europe?, July 2002

754 Ruslan Lukach and Joseph Plasmans, Measuring Knowledge Spillovers Using Patent Citations: Evidence from the Belgian Firm's Data, July 2002

755 Aaron Tornell and Frank Westermann, Boom-Bust Cycles in Middle Income Countries: Facts and Explanation, July 2002

756 Jan K. Brueckner, Internalization of Airport Congestion: A Network Analysis, July 2002

757 Lawrence M. Kahn, The Impact of Wage-Setting Institutions on the Incidence of Public Employment in the OECD: 1960-98, July 2002

758 Sijbren Cnossen, Tax Policy in the European Union, August 2002

759 Chandima Mendis, External Shocks and Banking Crises in Developing Countries: Does the Exchange Rate Regime Matter?, August 2002

760 Bruno S. Frey and Lars P. Feld, Deterrence and Morale in Taxation: An Empirical Analysis, August 2002

761 Lars Calmfors and Åsa Johansson, Nominal Wage Flexibility, Wage Indexation and Monetary Union, August 2002

762 Alexander R. W. Robson and Stergios Skaperdas, Costly Enforcement of Property Rights and the Coase Theorem, August 2002

763 Horst Raff, Preferential Trade Agreements and Tax Competition for Foreign Direct Investment, August 2002

764 Alex Cukierman and V. Anton Muscatelli, Do Central Banks have Precautionary Demands for Expansions and for Price Stability? – Theory and Evidence, August 2002

765 Giovanni Peri, Knowledge Flows and Knowledge Externalities, August 2002

766 Daniel Friedman and Nirvikar Singh, Equilibrium Vengeance, August 2002

767 Sam Bucovetsky and Michael Smart, The Efficiency Consequences of Local Revenue Equalization: Tax Competition and Tax Distortions, August 2002

768 Tapio Palokangas, International Labour Market Regulation and Economic Growth with Creative Destruction, August 2002

769 Rudi Dornbusch, The New International Architecture, September 2002

770 Hans-Werner Sinn, Weber's Law and the Biological Evolution of Risk Preferences: The Selective Dominance of the Logarithmic Utility Function, September 2002

771 Thomas Mayer, The Macroeconomic Loss Function: A Critical Note, September 2002

772 Seppo Honkapohja and Kaushik Mitra, Learning Stability in Economies with Heterogenous Agents, September 2002