Journal of Economic Theory **102**, 354–374 (2002)
doi:10.1006/jeth.2000.2768, available online at http://www.idealibrary.com on IDEAL®

Determinacy and Stability under Learning of Rational Expectations Equilibria[1]

Stéphane Gauthier

CREST and ERMES, Université Paris 2, Paris, France
gauthier@ensae.fr

Received December 7, 1998; final version received October 17, 2000

This paper studies relationships between the local determinacy of a stationary equilibrium in the perfect foresight dynamics, and its local stability in dynamics arising from econometric learning procedures. Attention is focused on linear scalar economies where agents forecast only one period ahead, and with an arbitrary, but fixed, number of predetermined variables. In such a framework, it is well known that there are no clear links between the determinacy of the stationary state in the perfect foresight dynamics on the levels of the state variable, and its stability under learning. The paper emphasizes, however, that this is not the right perfect foresight dynamics to look at whenever agents try to learn the coefficients of the perfect foresight dynamics restricted to an eigenspace of lower dimension. Indeed the paper introduces a *growth rate perfect foresight dynamics* on these coefficients and proves equivalence between determinacy in that dynamics and stability under learning provided that a simple sign condition is satisfied. *Journal of Economic Literature* Classification Numbers: E32, D83. © 2002 Elsevier Science

Key Words: rational expectations; determinacy; learning; expectational stability.

1. INTRODUCTION

It is a commonplace knowledge that dynamic models with rational expectations exhibit a multiplicity of equilibrium paths. Several alternative devices have consequently been proposed for the selection of the solutions on which attention should be focused in practice. This paper is an attempt to confront two among these, the determinacy of an equilibrium and its stability in a learning dynamics. This purpose is related to Guesnerie's [9] *dynamic equivalence principle* which claims that local determinacy should be equivalent to local stability under learning provided that agents form

[1] I thank my thesis supervisor Roger Guesnerie for his advice and Jean-Michel Grandmont for the kind interest he paid to my work. I am grateful to Antoine d'Autume, Gabriel Desgranges, George Evans, and Claude Oddou for helpful discussions and to an anonymous referee of this journal for guidance in writing the paper. Of course the usual disclaimers apply.

0022-0531/02 $35.00

"reasonably" their forecasts (see also Lucas [13]), and temporary equilibrium literature provides indeed many examples of reasonable learning rules (Grandmont and Laroque [6], Grandmont and Laroque [7] or Guesnerie and Woodford [10]). Nevertheless, the equivalence principle fails when agents employ recursive (econometric) learning rules (Duffy [1], Evans and Honkapohja [3], Grandmont [5] or Grandmont and Laroque [8]) such as the ordinary least squares algorithm (Marcet and Sargent [14]), the Robbins and Monro [17] scheme (Woodford [18]), or the gradient one (Evans and Honkapohja [4]). Most of these are studied in Ljung and Soderström [12]. Our aim is to show that the determinacy criterion is not applied in a suitable way in these cases.

The intuition is easy to grasp in a linear one step forward looking economy where the current (univariate) state depends on $L=1$ predetermined variable (as, e.g., Reichlin [16]). In such a model, the dynamics with perfect foresight is governed by two (local perfect foresight eigenvalues) growth rates λ_1 and λ_2 (with $|\lambda_1|<|\lambda_2|$) in the immediate vicinity of a stationary state $\bar{x}$ (where the state variable is equal to $\bar{x}$ at all times). Usually the determinacy criterion is applied to $\bar{x}$ in the perfect foresight dynamics on the levels of the state variable, whereas agents try to estimate λ_1 and λ_2 through a standard econometric procedure, the asymptotic behavior of which can be approximated by suitable continuous differential equations involving the expectational stability criterion (Evans [2]). It turns out that agents discover the growth rate of least modulus λ_1 in such a specification, so that the learning dynamics may be stable ($|\lambda_1|<1$) even if $\bar{x}$ is locally indeterminate ($|\lambda_2|<1$). One may wonder, however, whether this is the right perfect foresight dynamics to look at. The main innovation of the paper is indeed to apply the determinacy criterion to the fixed points in the learning dynamics (λ_1 and λ_2) by defining a perfect foresight dynamics of growth rates whose stationary equilibria are the perfect foresight growth rates. The outcome is particularly appealing within this simple economy (with only one predetermined variable) since a perfect foresight growth rate is stable under learning if and only if it is determinate in the perfect foresight dynamics of growth rates so defined.

In the more general framework where the current state depends on expectations of the next state and on an arbitrary, but fixed, number $L\geqslant 0$ of predetermined variables, the dynamics with perfect foresight involves $(L+1)$ local perfect foresight eigenvalues $\lambda_1, ..., \lambda_{L+1}$ (with $|\lambda_1|<\cdots<|\lambda_{L+1}|$) in the neighborhood of $\bar{x}$. In that case, agents are supposed to try to learn the L coefficients of the linear perfect foresight dynamics restricted to a L-dimensional eigenspace spanned by L eigenvectors among those associated with the $(L+1)$ local eigenvalues. When λ_{L+1} and $\sum_{i=1}^{L+1}\lambda_i$ have the same sign (which must be the case if $L\leqslant 1$), the eigenspace that is locally stable in the learning dynamics corresponds to the eigenvalues $\lambda_1, ..., \lambda_L$

of lowest modulus. Hence $\bar{x}$ is locally stable in the learning dynamics if and only if $|\lambda_L| < 1$, which encompasses both the saddle point determinate configuration ($1 < |\lambda_{L+1}|$) and the indeterminate configuration ($\lambda_{L+1}| < 1$), for the perfect foresight dynamics. When λ_{L+1} and $\sum_{i=1}^{L+1} \lambda_i$ have opposite signs, the eigenspace that is locally stable under learning always includes the eigenvector associated with λ_{L+1}, and the learning dynamics will be stable if and only if $|\lambda_{L+1}| < 1$, i.e., in the locally indeterminate configuration for the perfect foresight dynamics. These results imply again that the learning dynamics may be stable even if $\bar{x}$ is locally indeterminate in the perfect foresight dynamics on the levels of the state variable. Nevertheless one should instead look at a new perfect foresight dynamics, the *extended growth rate perfect foresight dynamics*, defined on the L-dimensional vectors of the coefficients that agents try to estimate. The issue is whether there is a neater relation between local stability under the considered class of learning algorithms, and local determinacy in this extended growth rate perfect foresight dynamics, of a particular eigenspace or of the L-dimensional vector of coefficients associated to it. The outcome here is still very simple since the L-dimensional eigenspace corresponding to the L perfect foresight growth rates $\lambda_1, ..., \lambda_L$ of lowest modulus, is the only one to be locally determinate. Therefore, if λ_{L+1} and $\sum_{i=1}^{L+1} \lambda_i$ have the same sign (a condition that may be related to the one sided sign condition for stability of differential equations, and that is always satisfied when $L = 1$), one gets indeed equivalence between the local determinacy of a particular eigenspace, or of the associate vector, and its local stability under learning. However this equivalence fails if λ_{L+1} and $\sum_{i=1}^{L+1} \lambda_i$ have opposite signs.

In the paper, we first present the simple case where $L \leqslant 1$ (Section 2), and then we turn to the more general one where L is arbitrary (Section 3). The conclusion (Section 4) will open a few leads about possible extensions.

2. A PRELIMINARY EXAMPLE

We shall suppose that the current state is a real number x_t linked with the forecast of the next state x^e_{t+1} and with the predetermined state x_{t-1} through the following map,

$$\gamma x^e_{t+1} + x_t + \delta x_{t-1} = 0, \tag{1}$$

where γ (with $\gamma \neq 0$, i.e., expectations matter) and δ represent the relative weights of future and past respectively. This equation stands for a first order approximation of a temporary equilibrium dynamics in a small neighborhood of a locally unique stationary state ($\bar{x} \equiv 0$). We first focus interest on the relationships between the usual concept of the local determinacy of $\bar{x}$, in the perfect foresight dynamics generated by (1) with $x^e_{t+1} = x_{t+1}$, namely,

$$\gamma x_{t+1} + x_t + \delta x_{t-1} = 0 \tag{2}$$

and its stability under learning. This concept of local determinacy is entirely governed by the perfect foresight roots λ_1 and λ_2 of the characteristic polynomial $P(z) = \gamma z^2 + z + \delta$ corresponding to (2). We let λ_1 and λ_2 be real, with $|\lambda_1| < |\lambda_2|$. If $|\lambda_1| > 1$, then $\bar{x}$ is source determinate. If $|\lambda_1| < 1 < |\lambda_2|$, then it is saddle determinate and for every arbitrarily small neighborhood $V(\bar{x})$ of $\bar{x}$, there is a unique equilibrium (x_t) satisfying (2) and staying in $V(\bar{x})$ at all times, for any initial condition x_{-1} close to $\bar{x}$. If $|\lambda_2| < 1$, then $\bar{x}$ is locally indeterminate. In that case, for any arbitrarily small neighborhood $V(\bar{x})$ of $\bar{x}$, and any initial condition x_{-1} close to $\bar{x}$, there are infinitely many perfect foresight equilibria staying in $V(\bar{x})$ at all times.

One can interpret perfect foresight equilibria as a situation where traders believe that the state variable behavior is governed by

$$x_t = \beta x_{t-1} \tag{3}$$

for every $x_{t-1} \in V(\bar{x})$ and every $t \geqslant 0$, and set the growth rate β equal to λ_1 or λ_2. Indeed agents are supposed to form their expectations by iterating twice (3) at time t, i.e., $x^e_{t+1} = \beta^2 x_{t-1}$ (note that they do not condition on x_t), so that the actual dynamics comes by inserting x^e_{t+1} into (1):

$$x_t = -(\gamma\beta^2 + \delta)\, x_{t-1}. \tag{4}$$

Their belief is self-fulfilling for $\beta = \lambda_i$ $(i = 1, 2)$ since then $-(\gamma\beta^2 + \delta)$ equals λ_i by definition. One needs the additional condition $|\lambda_i| < 1$ to get a locally feasible equilibrium that stays near the stationary state $\bar{x}$ at all times. Thus the belief $\beta = \lambda_1$ is the only one to be self-fulfilling and locally feasible when $\bar{x}$ is saddle determinate. Both beliefs $\beta = \lambda_1$ and $\beta = \lambda_2$ are self-fulfilling and locally feasible in the indeterminate case. None of these self-fulfilling beliefs is locally feasible when $\bar{x}$ is source determinate.

According to the above interpretation, perfect foresight requires that agents coordinate their expectations on some self-fulfilling belief $\beta = \lambda_i$ $(i = 1, 2)$. This is a demanding hypothesis in a decentralized framework where agents may even not know the dynamic laws of their environment, summarized by λ_1 and λ_2. It is thus natural to analyze how agents may in fact discover asymptotically λ_1 and λ_2 through some learning process, where they would formulate their expectations at each date t from beliefs (3) with $\beta = \beta(t)$ and would revise them at the beginning of period $(t+1)$ as a function of the actually observed forecasting error $(x_t - \beta(t)\, x_{t-1})$ in period t. Here we shall consider the class of econometric learning algorithms whose recursive form is

$$\beta(t+1) = \beta(t) + \alpha(t)\, h(t)\, x_{t-1}[x_t - \beta(t)\, x_{t-1}], \tag{5}$$

$$x_t = -(\gamma\beta(t)^2 + \delta)\, x_{t-1}, \tag{6}$$

where $\alpha(t) > 0$ tends toward 0 as t becomes large, and $h(t)$ is a function of past history of the state variable. One must impose that $h(t) > 0$ for if agents overestimate the actual growth rate, i.e., $x_t < \beta(t)\, x_{t-1}$, then they set $\beta(t+1) < \beta(t)$ in (5). A particular case of this formulation is the weighted least-squares learning scheme (Marcet and Sargent [14]) where $\alpha(t) = \gamma_t/t$ and $h(t) = ((\gamma_t x_{t-1}^2 + \cdots + \gamma_0 x_{-1}^2)/t)^{-1}$ with a forgetting factor $\gamma_s \geqslant 0$ ($s = 0, ..., t$) that allows us to weight recent observations more heavily ($\gamma_s = 1$ for the ordinary least-squares scheme).

If agents set $\beta(t) = \lambda_i$ in (5), then their belief is self-fulfilling, i.e., $x_t = \lambda_i x_{t-1}$ in (6), and they cease to revise their estimates $\beta(t)$ in (5), i.e., $\beta(t+1) = \beta(t) = \lambda_i$, thus learning the whole trajectory $x_t = \lambda_i x_{t-1}$. Although the dynamics (5)–(6) seems, at first sight, complex to analyze because of the coupling between growth rates and levels of the state variable, it can be shown (Ljung's [11]) that, provided that $|\lambda_i| < 1$ (i.e., the self-fulfilling belief $\beta = \lambda_i$ in (3) is locally feasible) and $\alpha(t)$ goes to 0 but not too fast, local asymptotic stability of $\beta(t) = \lambda_i$ and of $x_t = \lambda_i x_{t-1}$ in (5)–(6) are equivalent to local stability of $\beta(t) = \lambda_i$ in the simpler associated differential equation

$$\frac{d\beta(\tau)}{d\tau} = \bar{\phi}[-(\gamma\beta(\tau)^2 + \delta) - \beta(\tau)], \tag{7}$$

where τ is a fictitious scale of time related to t, and $\bar{\phi}$ assumed to be the limit of some statistics of the process involving $\alpha(t)$, $h(t)$ and x_t.[2] Provided

[2] The sufficient conditions for this result are $|\lambda_i| < 1$, the gain $\alpha(t)$ tends to 0 but the series $\sum_{t=1}^{\infty} \alpha(t)$ diverge to $+\infty$, and a major condition is that the limit $\bar{\phi}$ exists. If $\alpha(t) = 1/t$ (which satisfies both conditions above), then $\bar{\phi}$ is defined as:

$$\bar{\phi} = \lim_{t \to +\infty} \frac{1}{t} \sum_{s=1}^{t} h_s x_{s-1}^2.$$

See Ljung [11] (conditions C3, C5, C6) for more general sequences of gains. Under these assumptions, the weighted least squares dynamics (whose recursive form is given in (2.15) of Ljung and Soderström (12]) is governed in the long run by (see (5) in Marcet and Sargent [14])

$$\frac{d\beta(\tau)}{d\tau} = \frac{M_x}{R(\tau)}[-(\gamma\beta(\tau)^2 + \delta) - \beta(\tau)], \tag{8}$$

$$\text{and} \qquad \frac{dR(\tau)}{d\tau} = M_x - R(\tau), \tag{9}$$

where $R(t) = 1/h(t) > 0$ and $M_x > 0$ (Marcet and Sargent [14], Section 2). Since $R(t)$ tends to M_x in (9), one can rewrite (8) as (7) with $\bar{\phi} = M_x/M_x = 1 > 0$, so that stability of (λ_i, M_x) in (8)–(9) is equivalent to stability of λ_i in (10) for this class of learning schemes, as is asserted in (6) of Marcet and Sargent [14].

that $\bar{\phi} > 0$ (a condition that should be thought as a consistency requirement in the learning process), the root λ_i will be locally stable for (7) if and only if it is locally stable for the following expectational stability differential equation (Evans [2])

$$\frac{d\beta(\tau)}{d\tau} = -(\gamma\beta(\tau)^2 + \delta) - \beta(\tau) \tag{10}$$

which makes the revision of growth rate estimates to depend directly on the discrepancy between the actual growth rate $(-(\gamma\beta(\tau)^2 + \delta))$ and the initial guess $\beta(\tau)$. The next lemma states the properties of (10) in the immediate vicinity of its rest points λ_1 and λ_2.

LEMMA 1. *Let $\gamma \neq 0$. The low perfect foresight root λ_1 is locally stable while the high one λ_2 is locally unstable in the dynamics with learning* (10).

Proof. Local stability of λ_i $(i = 1, 2)$ in (10) is obtained if and only if the first derivative of $(-(\gamma\beta^2 + \delta)) - \beta)$ with respect to β is negative for $\beta = \lambda_i$, namely,

$$\frac{2\lambda_i}{(\lambda_1 + \lambda_2)} - 1 < 0 \tag{11}$$

since $(\lambda_1 + \lambda_2) \equiv -1/\gamma$. If $\lambda_i = \lambda_1$, then $(11) \Leftrightarrow \lambda_1/\lambda_2 < 1$, which is always satisfied. Otherwise, if $\lambda_i = \lambda_2 > 0$, then $(11) \Leftrightarrow \lambda_2 < \lambda_1$, which never holds true. If $\lambda_i = \lambda_2 < 0$, then $(11) \Leftrightarrow \lambda_2 < \lambda_1$, which is impossible too. ∎

Thus the equivalence principle does not hold under the class of learning processes considered here, since agents then learn always the root λ_1 of lowest modulus, and never discover λ_2. Indeed convergence of the trajectories (x_t) to the stationary state value $\bar{x}$ imposes $|\lambda_1| < 1$, but this is compatible with both the saddle determinate configuration $(1 < |\lambda_2|)$ and with the indeterminate case $(|\lambda_2| < 1)$. We argue in this paper that this failure is in large part due to the fact that the usual notion of determinacy, as recalled above, is stated in terms of the perfect foresight dynamics of the *levels* of the state variable x_t, while in fact traders try to discover some *growth rate* $\beta = \lambda_i$ $(i = 1, 2)$. We introduce now a new perfect foresight dynamics of growth rates, which is obtained by assuming that the traders' belief about the law of motion of the state variable fits (3) with $\beta = \beta(t)$,

$$x_t = \beta(t)\, x_{t-1}, \tag{12}$$

for every $x_{t-1} \in V(\bar{x})$ and every $t \geqslant 0$. Perfect foresight of the state variable x induces a dynamics of the growth rate $\beta(t)$ whose fixed points are the roots λ_1 and λ_2. The issue is to study whether there exist links between stability of these roots under learning and their determinacy properties in such a growth rate perfect foresight dynamics. It turns out that, in fact, a perfect foresight growth rate is locally stable under learning if and only if it is locally determinate in the growth rate perfect foresight dynamics.

Since (12) holds for every $t \geqslant 0$, traders' expectations are:

$$x_t^e = \beta(t)\, x_{t-1} \qquad \text{and} \qquad x_{t+1}^e = \beta(t+1)\, x_t^e = \beta(t+1)\, \beta(t)\, x_{t-1}.$$

In the perfect foresight dynamics (2), these forecasts are equal to x_t and x_{t+1}, respectively. One can consequently rewrite (2) as a recursive equation of growth rates only.

Definition 2. Let $\gamma \neq 0$. Assume also that (12) holds for every $x_{t-1} \in V(\bar{x})$ and every $t \geqslant 0$. The growth rate perfect foresight dynamics is a sequence of growth rates $(\beta(t))$ such that:

$$\gamma \beta(t+1)\, \beta(t) + \beta(t) + \delta = 0. \tag{13}$$

Taking $\beta(t) = \beta(t+1)$ in (13) shows that the fixed points of (13) are λ_1 and λ_2. This dynamics is well defined if and only if $\beta(t) \neq 0$ in each period. Therefore it must be the case that $\beta(t) \neq -\delta$ if $\delta \neq 0$ (otherwise $\beta(t+1) = 0$), in which case (13) does not define a global dynamics but is yet well defined around λ_1 and λ_2 (both differ from $-\delta$). If $\delta = 0$, then $\lambda_1 = 0$ (since then $\lambda_1 \lambda_2 \equiv -\delta/\gamma = 0$ and $|\lambda_1| < |\lambda_2|$) so that the dynamics is not well defined locally, but one may say that λ_1 is unstable while λ_2 is stable because either $\beta(t) = \lambda_1$ at all times, or $\beta(t) = -1/\gamma = \lambda_2 \neq 0$ if $\beta(t) \neq 0$ at some date.

The dynamics (13) has the classical one-step forward looking structure without predetermined variables (the current rate $\beta(t)$ is not given at outset of t). So its fixed points are locally determinate if and only if they are locally unstable in (13), which allows us to state our equivalence result.

Proposition 3. *Let $\gamma \neq 0$. The low perfect foresight root λ_1 is locally determinate in the growth rate perfect foresight dynamics* (13) *while the high one λ_2 is locally indeterminate in the same dynamics. Therefore, in view of Lemma* 1, *the root λ_i ($i = 1, 2$) is locally stable in the learning dynamics* (10) *if and only if it is locally determinate in the growth rate perfect foresight dynamics* (13).

Proof. The growth rate perfect foresight dynamics in the neighborhood of λ_i ($i = 1, 2$) is obtained by linearizing (13) at this point:

$$\gamma \lambda_i (\beta(t+1) - \lambda_i) + (\gamma \lambda_i + 1)(\beta(t) - \lambda_i) = 0.$$

The condition for λ_i to be locally unstable in (13) is:

$$\left|\frac{\gamma\lambda_i+1}{\gamma\lambda_i}\right| = \left|\frac{\lambda_i-(\lambda_1+\lambda_2)}{\lambda_i}\right| > 1. \tag{14}$$

If $\lambda_i = \lambda_1$, then (14) $\Leftrightarrow |\lambda_2/\lambda_1| > 1$, which always holds. If $\lambda_i = \lambda_2$, then (14) $\Leftrightarrow |\lambda_1/\lambda_2| > 1$, which never holds. ▮

3. ON THE DYNAMIC EQUIVALENCE PRINCIPLE

We now deal with economies where the current state depends on the forecast of the next state but also on an arbitrary number L of predetermined variables through the following map,

$$\gamma x^e_{t+1} + x_t + \sum_{l=1}^{L} \delta_l x_{t-l} = 0, \tag{15}$$

where the parameter δ_l ($1 \leqslant l \leqslant L$) represents the relative contribution to x_t of the predetermined state x_{t-l} at t. We shall proceed here as in the previous section, namely we shall first state the lack of link between determinacy of the stationary state ($\bar{x} \equiv 0$) in the perfect foresight dynamics on the levels of the state variable x, and its stability under learning when agents try to estimate how the state of period t is related to the L past states $x_{t-1}, ..., x_{t-L}$, i.e., try to discover the L coefficients of the linear perfect foresight dynamics restricted to an L-dimensional eigenspace. Then we shall define the perfect foresight dynamics on the L-dimensional vectors of these coefficients. Thus the issue will be to study whether there is a neater relationship between determinacy and stability under learning with this new perfect foresight dynamics.

3.1. *State Variable Perfect Foresight Dynamics*

The usual concept of local determinacy of the stationary state $\bar{x}$ is defined from the perfect foresight dynamics on the levels of the state variable x obtained by setting $x^e_{t+1} = x_{t+1}$ in (15), namely:

$$\gamma x_{t+1} + x_t + \sum_{l=1}^{L} \delta_l x_{t-l} = 0. \tag{16}$$

Indeed this concept only relies on the $(L+1)$ perfect foresight roots $\lambda_1, ..., \lambda_{L+1}$ of the characteristic polynomial $P(z) = \gamma z^{L+1} + z^L + \sum_{l=1}^{L} \delta_l z^{L-l}$ corresponding to (16). Let λ_i ($i = 1, ..., L+1$) be real, with $|\lambda_1| < |\lambda_2| < \cdots < |\lambda_{L+1}|$. If $|\lambda_L| > 1$, then $\bar{x}$ is source determinate. If $|\lambda_L| < 1 < |\lambda_{L+1}|$, then $\bar{x}$ is saddle determinate and for every arbitrarily small neighborhood $V(\bar{x})$ of $\bar{x}$, and for every initial condition $(x_{-1}, ..., x_{-L})$ close to $(\bar{x}, ..., \bar{x})$, there is a unique perfect foresight equilibrium, i.e., a unique sequence (x_t) that satisfies (16) and stays in $V(\bar{x})$ at all times $t \geqslant 0$. If $|\lambda_{L+1}| < 1$, then $\bar{x}$ is indeterminate, i.e., for every arbitrarily small neighborhood $V(\bar{x})$ of $\bar{x}$, and for every initial condition $(x_{-1}, ..., x_{-L})$ close to $(\bar{x}, ..., \bar{x})$, there are infinitely many perfect foresight equilibria (x_t) that stay in $V(\bar{x})$ at all dates $t \geqslant 0$.

Here again one can interpret perfect foresight as a situation where traders believe that the law of motion of the system is governed by

$$x_t = \sum_{l=1}^{L} \beta_l x_{t-l} \tag{17}$$

for every $x_{t-l} \in V(\bar{x})$ and every $t \geqslant 0$, and where this belief is self-fulfilling. In that case, the expectation is formed at date t by iterating twice (17):

$$x^e_{t+1} = \sum_{l=1}^{L} \beta_l x_{t+1-l} = \sum_{l=1}^{L-1} (\beta_1 \beta_l + \beta_{l+1})\, x_{x-l} + \beta_1 \beta_L x_{t-L}.$$

This forecast, once reintroduced into (15), generates the actual dynamics:

$$x_t = -\sum_{l=1}^{L-1} [\gamma(\beta_1 \beta_l + \beta_{l+1}) + \delta_l]\, x_{t-l} - [\gamma(\beta_1 \beta_L) + \delta_L]\, x_{t-L}. \tag{18}$$

Hence the initial belief (17) is self-fulfilling if and only if it coincides with (18), namely,

$$\beta_l = -[\gamma(\beta_1 \beta_l + \beta_{l+1}) + \delta_l] \qquad \text{for} \quad l = 1, ..., L \tag{19}$$

with $\beta_{L+1} = 0$. Vectors solutions to (19) will be called stationary *extended growth rates* (henceforth $EGR(L)$), and denoted $\bar{\boldsymbol{\beta}} = (\bar{\beta}_1, ..., \bar{\beta}_L)$. Intuitively the self-fulfilling belief (17) with $\boldsymbol{\beta} \equiv (\beta_1, ..., \beta_L) = \bar{\boldsymbol{\beta}}$ should correspond to the perfect foresight dynamics (16) restricted to an invariant subspace W of dimension L, i.e., to an eigenspace spanned by L eigenvectors among the $(L+1)$ eigenvectors $\mathbf{u}_i$ associated with the eigenvalues λ_i $(i = 1, ..., L+1)$. There are clearly $(L+1)$ such invariant L-dimensional eigenspaces W_k (each W_k is spanned by all the eigenvectors but $\mathbf{u}_k$ $(k = 1, ..., L+1)$). Thus one should expect (19) to have $(L+1)$ distinct vector solutions $\bar{\boldsymbol{\beta}}^k$, where $\bar{\boldsymbol{\beta}}^k$ corresponds to the perfect foresight dynamics restricted to W_k. The next lemma makes precise this intuition and gives the expression of $\bar{\boldsymbol{\beta}}^k$ in terms of the perfect foresight roots λ_i $(i = 1, ..., L+1)$.

Lemma 4. *Assume that the characteristic polynomial P corresponding to the* $(L+1)$*-dimensional difference equation* (16) *admits* $(L+1)$ *real and distinct roots* λ_k, *with* $|\lambda_1| < \cdots < |\lambda_{L+1}|$. *Let the* $(L+1)\times 1$ *eigenvector* $\mathbf{u}_k$, $1 \leqslant k \leqslant L+1$, *be associated with* λ_k. *Finally let* $W_k \subseteq IR^L$, $1 \leqslant k \leqslant L+1$, *be the eigensubspace spanned by all the eigenvectors except* $\mathbf{u}_k$. *The perfect foresight dynamics of the state variable restricted to* W_k *writes*

$$x_t = \sum_{l=1}^{L} \bar{\beta}_l^k x_{t-l},$$

where the lth entry $\bar{\beta}_l^k$ *of the stationary EGR*(L) $\bar{\boldsymbol{\beta}}^k$ *is*:

$$\bar{\beta}_l^k = (-1)^{l+1} \sum_{1 \leqslant i_1 < \cdots < i_l} (\lambda_{i_1} ... \lambda_{i_l}) \qquad \textit{for all} \quad i_z \neq k,\ z = 1, ..., l.$$

The $(L+1)$ *stationary EGR*(L) $\bar{\boldsymbol{\beta}}^k$ *are the solutions of the equations* (19).

Proof. See in Appendix 5.1. ∎

Of course, for the self-fulfilling belief (17) with $\bar{\boldsymbol{\beta}}^k$ to be locally feasible, one must also require that the perfect foresight dynamics of the state variable x in W_k be stable (all the perfect foresight roots different from λ_k must be stable, i.e., $|\lambda_i| < 1$ for $i \neq k$). Therefore the self-fulfilling belief $\boldsymbol{\beta} = \bar{\boldsymbol{\beta}}^k$ with $k = L+1$ is the only one to be locally feasible if $\bar{x}$ is saddle determinate in the dynamics (16). Any self-fulfilling belief $\bar{\boldsymbol{\beta}}^k$ with $k = 1, ..., L+1$ is locally feasible if $\bar{x}$ is indeterminate in this dynamics, and none of these self-fulfilling beliefs is locally feasible when $\bar{x}$ is source determinate.

3.2. *Dynamics with Learning*

Even if agents are initially aware of the $(L+1)$ stationary *EGR*(L) to the model (15), they have to coordinate their behavior on one among them, which may be quite demanding in a decentralized framework. In this section we shall instead assume that agents need learning how the state variable behaves in $V(\bar{x})$ through a process where they form at the outset of period t their forecasts from the law (17) with some L-dimensional vector of estimates $\boldsymbol{\beta}(t) = (\beta_1(t), ...\beta_L(t))^{\mathrm{T}}$ and revise these estimates at the outset of period $(t+1)$ once the forecasting error $(x_t - \boldsymbol{\beta}(t)^{\mathrm{T}}\, \mathbf{x}_{t-1})$ is actually observed (here $\mathbf{x}_{t-1}$ stands for the vector of lagged variables $(x_{t-1}, ..., x_{t-L})^{\mathrm{T}}$), according to the class of recursive learning schemes:

$$\boldsymbol{\beta}(t+1) = \boldsymbol{\beta}(t) + \alpha(t)\, \mathbf{H}(t)\, \mathbf{x}_{t-1}(x_t - \boldsymbol{\beta}(t)^{\mathrm{T}}\, \mathbf{x}_{t-1}), \tag{20}$$

$$x_t = -\sum_{l=1}^{L} [\gamma(\beta_1(t)\, \beta_l(t) + \beta_{l+1}(t)) + \delta_l]\, x_{t-l} \equiv \boldsymbol{\Omega}(\boldsymbol{\beta}(t))^{\mathrm{T}}\, \mathbf{x}_{t-1}. \tag{21}$$

Equation (21) is (18) with β_l is replaced by $\beta_l(t)$ (for $l=1, ..., L$), and with the convention that $\beta_{L+1}(t)=0$ (hence the lth component of the $L\times 1$ vector $\mathbf{\Omega}(\boldsymbol{\beta}(t))$ is the actual weight of x_{t-l} in (21)). As in (5), the sequence of scalars $\alpha(t)>0$ in (20) is still assumed to tend toward 0 as time passes, and $\mathbf{H}(t)$ is now an $L\times L$ matrix related to past history of the economic system. The rule (20) is general enough to encompass, e.g., the weighted least-squares schemes where $\mathbf{H}(t)=((\gamma_t \mathbf{x}_{t-1}\mathbf{x}_{t-1}^{\mathrm{T}}+\cdots+\gamma_0 \mathbf{x}_{-1}\mathbf{x}_{-1}^{\mathrm{T}})/t)^{-1}$ and $\alpha(t)=\gamma_t/t$ (γ_t is the weight of period t observations in the least square estimator).

If agents set $\boldsymbol{\beta}(t)=\bar{\boldsymbol{\beta}}^k$ in (20) for some k ($k=1, ..., L+1$) at date t, then $x_t=(\bar{\boldsymbol{\beta}}^k)^{\mathrm{T}}\,\mathbf{x}_{t-1}$ in (21). Therefore they do not revise their estimates, i.e., they set $\boldsymbol{\beta}(t+1)=\boldsymbol{\beta}(t)=\bar{\boldsymbol{\beta}}^k$ in (20), thus learning the law of motion of the state variable restricted to W_k. The dynamics (20)–(21) is, however, complex to analyze because of the coupling between components of $\boldsymbol{\beta}(t)$ and levels of the state variable. Here again, as in Section 2, we may appeal to the existing theory on the convergence of learning algorithms (Ljung's [11]) to assert that local asymptotic stability in (20)–(21) of some particular $\bar{\boldsymbol{\beta}}^k$ corresponding to a locally stable (feasible) dynamics in W_k is equivalent, provided that $\alpha(t)$ goes to 0 but not to fast, to its local stability in the associated ordinary differential equation[3]

$$\dot{\boldsymbol{\beta}}(\tau)=\bar{\mathbf{\Phi}}(\mathbf{\Omega}(\boldsymbol{\beta}(\tau))-\boldsymbol{\beta}(\tau)), \tag{22}$$

where τ is a fictitious continuous scale of time, and where the $L\times L$ matrix $\bar{\mathbf{\Phi}}$ is assumed to be the limit of some statistics involving $\alpha(t)$, $\mathbf{H}(t)$ and the state variable x. We need here additional assumptions to be able to reduce the study of the stability in (22) to the differential expectational stability criterion (Evans [2]):

$$\dot{\boldsymbol{\beta}}(\tau)=\mathbf{\Omega}(\boldsymbol{\beta}(\tau))-\boldsymbol{\beta}(\tau). \tag{23}$$

An intuition for the kind of assumptions needed to ensure local equivalence between (22) and (23) is easy to grasp by considering Jacobian matrices $\bar{\mathbf{\Phi}}(\mathbf{D\Omega}(\bar{\boldsymbol{\beta}}^k)-\mathbf{I}_L)$ and $(\mathbf{D\Omega}(\bar{\boldsymbol{\beta}}^k)-\mathbf{I}_L)$ that govern the dynamics near $\bar{\boldsymbol{\beta}}^k$ in (22) and (23), respectively. One can show that $(\mathbf{D\Omega}(\bar{\boldsymbol{\beta}}^k)-\mathbf{I}_L)$ has distinct real eigenvalues under the assumption of distinct real eigenvalues λ_i ($i=1, ..., L+1$) (see Appendix 5.2). It is therefore diagonalizable. In the

[3] This result applies if $\bar{\boldsymbol{\beta}}^k$ corresponds to a locally feasible equilibrium (i.e., $|\lambda_i|<1$ for $i=1, ..., L+1$ but $i\neq k$), and if $\alpha(t)$ tends to 0 but $\sum_{t=1}^{\infty}\alpha(t)$ diverges to $+\infty$. For $\alpha(t)=1/t$, $\bar{\mathbf{\Phi}}$ is defined as:

$$\bar{\mathbf{\Phi}}\equiv\lim_{t\to+\infty}\frac{1}{t}\sum_{s=1}^{t}\mathbf{H}(s)\,\mathbf{x}_{s-1}\mathbf{x}_{s-1}^{\mathrm{T}}.$$

See again Ljung [11] (conditions C3, C5, C6) for more general sequences of gains.

corresponding basis, one should expect the asymptotic learning dynamics (22) to be consistent with that structure and to involve correcting each "diagonal error" along the direction of the corresponding eigenvector of $(\mathbf{D\Omega}(\bar{\boldsymbol{\beta}}^k) - \mathbf{I}_L)$ with a positive weight. In other words, *both* $\bar{\boldsymbol{\Phi}}$ *and* $(\mathbf{D\Omega}(\bar{\boldsymbol{\beta}}^k) - \mathbf{I}_L)$ *should be diagonalizable in the same basis, and the diagonalized matrix* $\bar{\boldsymbol{\Phi}}$ *should have only positive entries on its diagonal.* Local stability of (22) and (23) are clearly equivalent in such circumstances, which are in principle more general than usual least-squares schemes where $\bar{\boldsymbol{\Phi}}$ is in fact the identity matrix.[4] Of course the dynamics (23) will be locally stable around $\bar{\boldsymbol{\beta}}^k$ if and only if the real part of all the eigenvalues of $(\mathbf{D\Omega}(\bar{\boldsymbol{\beta}}^k) - \mathbf{I}_L)$.

THEOREM 5. *Let $\gamma \neq 0$. Assume that beliefs fit* (17) *for every x_{t-l} in $V(\bar{x})$ and every $t \geqslant 0$. Then there exists a unique stationary EGR(L) $\bar{\boldsymbol{\beta}}^k$ which is locally stable in the learning dynamics* (23). *It governs the behavior of the state variable restricted to the L-dimensional eigensubspace W_k of the state variable perfect foresight dynamics* (16) *spanned by all the eigenvectors except $\mathbf{u}_k$ where $\mathbf{u}_k$ is associated with the perfect foresight root λ_k which satisfies*:

$$\lambda_k \Big/ \sum_{i=1}^{L+1} \lambda_i = \max_i \left\{ \lambda_i \Big/ \sum_{i=1}^{L+1} \lambda_i \right\}.$$

Therefore the stationary EGR(L) $\bar{\boldsymbol{\beta}}^{L+1}$ that governs the perfect foresight dynamics (16) *restricted to the L-dimensional eigenspace associated with the L eigenvalues $\lambda_1, \ldots, \lambda_L$ of lowest modulus, is the unique stationary EGR(L) to be locally stable in the learning dynamics* (23) *if $\lambda_{L+1} \sum_{i=1}^{L+1} \lambda_i$ is positive. Otherwise, i. e., if $\lambda_{L+1} \sum_{i=1}^{L+1} \lambda_i$ is negative, then the stationary EGR(L) that is locally stable in the learning dynamics* (23) *governs the perfect foresight dynamics* (16) *restricted to an L-dimensional eigenspace that contains in particular the eigenvector $\mathbf{u}_{L+1}$ associated to the eigenvalue of largest modulus λ_{L+1}.*

[4] Under the assumptions given in footnote 3, the long run dynamics of weighted least squares is governed by (5) in Marcet and Sargent [14],

$$\frac{d\boldsymbol{\beta}(\tau)}{d\tau} = \mathbf{R}(\tau)^{-1} \mathbf{M}_x[-(\boldsymbol{\beta}(\tau)) - \boldsymbol{\beta}(\tau)], \tag{24}$$

$$\frac{d\mathbf{R}(\tau)}{d\tau} = \mathbf{M}_x - \mathbf{R}(\tau), \tag{25}$$

where $\mathbf{R}(t) = \mathbf{H}(t)^{-1}$ and $\mathbf{M}_x$ is a (positive definite) second moment matrix (Marcet and Sargent [14], Section 2). Note that $\mathbf{R}(t)$ tends to $\mathbf{M}_x$ in (25), so that (24) expresses as (22) in the long run, with $\bar{\boldsymbol{\Phi}} = \mathbf{M}_x^{-1}\mathbf{M}_x = \mathbf{I}_L$ (see also (6) in Marcet and Sargent (14]).

Proof. See in Appendix 5.2. ∎

This result implies again that there are no simple links between the determinacy of $\bar{x}$ in the perfect foresight dynamics (16) and its stability under learning. In the case where λ_{L+1} and $\sum_{i=1}^{L+1} \lambda_i$ have the same sign (the reader will note that this condition is always met in the case of a single predetermined variable $L \leqslant 1$ considered in the previous section), convergence of the state variable x_t to $\bar{x}$ in the learning dynamics requires also that $|\lambda_L| < 1$, but this condition is compatible with both the saddle determinate configuration $1 < |\lambda_{L+1}|$ and the indeterminate case $|\lambda_{L+1}| < 1$. On the contrary, when λ_{L+1} and $\sum_{i=1}^{L+1} \lambda_i$ have not the same sign (which may occur only if $L \geqslant 2$), convergence of x_t to $\bar{x}$ in the learning dynamics requires $|\lambda_{L+1}| < 1$, which corresponds to the indeterminate case.

3.3. *Extended Growth Rate Perfect Foresight Dynamics*

We argue here again that the fact that there are no simple links between the determinacy of the stationary state $\bar{x}$ and its stability under learning, may be due to the fact that determinacy was applied to the perfect foresight dynamics (16) on the levels of the state variable x_t, while agents try to learn extended growth rates. We may thus expect that simpler relationships might arise if one considers instead determinacy in a perfect foresight dynamics on extended growth rates. Such a dynamics is constructed in assuming that the traders' beliefs fit,

$$x_t = \sum_{l=1}^{L} \beta_l(t)\, x_{t-l}, \tag{26}$$

for every x_{t-l} ($l = 1, ..., L$) in $V(\bar{x})$ and every $t \geqslant 0$. The dynamics with perfect foresight (16) of the state variable induces a L-dimensional extended growth rate perfect foresight dynamics on the extended growth rates $\beta(t) = (\beta_1(t), ..., \beta_L(t))$ whose fixed points are the stationary $EGR(L)$. Applying the determinacy criterion to these vectors, in this extended growth rates dynamics, allows us to show by that the stationary $EGR(L)$ that governs the perfect foresight dynamics restricted to the L-dimensional eigenspace associated with the L eigenvalues $\lambda_1, ..., \lambda_L$ of lowest modulus, is actually the only one that is locally determinate. In view of Theorem 5, this fact implies that the equivalence between determinacy and stability under learning will be restored, at least when λ_{L+1} and $\sum_{i=1}^{L+1} \lambda_i$ have the same sign.

Since (26) holds for each t, the corresponding expectations are as follows:

$$x_t^e = \sum_{l=1}^{L} \beta_l(t)\, x_{t-l} \qquad \text{and} \qquad x_{t+1}^e = \beta_1(t+1)\, x_t^e + \sum_{l=2}^{L} \beta_l(t+1)\, x_{t+1-l}.$$

Under the perfect foresight hypothesis x_t^e and x_{t+1}^e are equal to x_t and x_{t+1}, respectively, so that (16) becomes:

$$(1+\gamma\beta_1(t+1))\,x_t = -\sum_{l=1}^{L-1} (\gamma\beta_{l+1}(t+1)+\delta_l)\,x_{t-l} - \delta_L x_{t-L}. \quad (27)$$

Since (26) and (27) holds for every x_{t-l}, the $EGR(L)$ perfect foresight dynamics is obtained whenever the coefficients of (26) and of (27) coincide.

DEFINITION 6. Let $\gamma \neq 0$. Assume also that (26) holds for every x_{t-l} near $\bar{x}$ ($l=1, ..., L$) and every $t \geqslant 0$. The $EGR(L)$ perfect foresight dynamics is a sequence of L-dimensional vectors $(\boldsymbol{\beta}(t))$ such that

$$\beta_l(t) = -\frac{\gamma\beta_{l+1}(t+1)+\delta_l}{1+\gamma\beta_1(t+1)} \quad (28)$$

for each $L=1, ..., L$, and with the convention that $\beta_{L+1}(t+1)=0$.

This dynamics is well defined if and only if $\beta_L(t) \neq 0$ in each period. For in that case (28) for $l=L$ determines $\beta_1(t+1)$ while the other equations determine $\beta_l(t+1)$ for $l=1, ..., L-1$. In the case $L \geqslant 2$ (the case $L=1$ was dealt with Section 2), $\beta_L(t) \neq 0$ implies $\beta_L(t+1) \neq 0$ if and only if $\gamma\beta_L(t+1) = (\beta_{L-1}(t)/\beta_L(t))\,\delta_l - \delta_{l-1} \neq 0$. So, when $\delta_L \neq 0$, global dynamics are not well defined but local dynamics near every stationary $EGR(L)$ are (see appendix 5.3). If $\delta_L = 0$, then $\lambda_1 = 0$ (since $\lambda_1 ... \lambda_{L+1} = -\delta_l/\gamma = 0$ and $|\lambda_1| < \cdots < |\lambda_{L+1}|$), which implies (by using Lemma 4) that $\bar{\beta}_L^1 \neq 0$ but $\bar{\beta}_L^k = 0$ for $k \neq 1$ (i.e., $k=2, ..., L+1$), so that local dynamics are not well defined around $\bar{\boldsymbol{\beta}}^k$ (for $k \neq 1$) but one may say that $\bar{\boldsymbol{\beta}}^k$ is unstable for $k \neq 1$ while $\bar{\boldsymbol{\beta}}^1$ is stable in the sense that either $\boldsymbol{\beta}(t) = \bar{\boldsymbol{\beta}}^k$ ($k \neq 1$) at all times, or $\boldsymbol{\beta}(t) = \bar{\boldsymbol{\beta}}^1$ if $\beta_L(t) \neq 0$ at some date.

Since (28) displays a (multidimensional) one-step forward looking structure without predetermined variables, a stationary $EGR(L)$ is locally determinate if and only if all the eigenvalues which govern (28) close to it have real part of modulus greater than 1.

PROPOSITION 7. *The stationary* $EGR(L)$ $\bar{\boldsymbol{\beta}}^{L+1}$ *that governs the perfect foresight dynamics* (16) *restricted to the L-dimensional eigenspace associated to the L eigenvalues* $\lambda_1, ..., \lambda_L$ *of lowest modulus, is the unique stationary* $EGR(L)$ *to be is locally determinate in the* $EGR(L)$ *perfect foresight dynamics* (28). *Therefore, in view of Theorem* 5, *the dynamic equivalence principle holds true if* λ_L *has the same sign as* $\sum_{i=1}^{L+1} \lambda_i$, *and not otherwise.*

Proof. See in Appendix 5.4. ∎

4. CONCLUSION

The main innovation of the paper is to describe an *extended growth rate perfect foresight dynamics* whose fixed points are the L-dimensional vectors of coefficients which govern the local perfect foresight dynamics restricted to some L-dimensional eigensubspace. Such a dynamics allows us to save the equivalence between determinacy and stability under recursive learning under a simple one sided (sign) condition that may be related to the condition for stability of differential equations. It is clear, however, that this result depends on specific features of the model.

(i) Agents were assumed to condition forecasts at t only on past data up to date $(t-1)$. An interesting topic for further research would be to investigate what would happen if agents may condition also on the current equilibrium state x_t.

(ii) Finally, agents were assumed to care about deviations from the stationary state value $\bar{x}$ so that, in fact, they were *a priori* aware of this value. One may go beyond such an assumption by introducing an estimate ϕ of the value $\bar{x}$ into the perceived laws of motion (17). Namely (17) rewrites:

$$x_t = \sum_{l=1}^{L} \beta_l x_{t-l} + \phi.$$

One can prove that, in this case, saddle determinacy of the stationary state $(\bar{x})$ is equivalent to its stability under learning if $\lambda_{L+1}/\sum_{i=1}^{L+1} \lambda_i > 0$ and $\sum_i \lambda_i = -1/\gamma > 0$, i.e., if $\lambda_{L+1} > 0$. Actually learning a stationary extended growth rate does not rely on the fact that agents already know $\bar{x}$. Therefore the former condition ensures that agents discover the law of motion of the state variable restricted to the subspace corresponding to the L roots of lowest modulus $\lambda_1, ..., \lambda_L$. The remaining one $(\gamma < 0)$ ensures that, in the model where $L = 0$, agents discover the value $\bar{x}$ $(\gamma > -1)$ if and only if the stationary state $(\bar{x})$ is locally determinate $(|\gamma| < 1)$.

5. PROOFS

5.1. *Proof of Lemma* 4

We first transform the dynamics (16) into a vector first order difference equation

$$\mathbf{x}_{t+1} = \mathbf{T}\mathbf{x}_t,$$

where

$$\mathbf{x}_t \equiv \begin{pmatrix} x_t \\ \vdots \\ x_{t-L} \end{pmatrix} \quad \text{and} \quad \mathbf{T} \equiv \begin{pmatrix} -1/\gamma & -\delta_1/\gamma & \cdots & -\delta_L/\gamma \\ & & & 0 \\ & \mathbf{I}_L & & \vdots \\ & & & 0 \end{pmatrix}.$$

The proof proceeds from the fact that $\mathbf{x}_t$ belongs to a L-dimensional eigen-subspace W_k ($k=1, ..., L+1$) of (16) if and only it is a linear combination of the k eigenvectors that span W_k. Of course the $(L+1)$ eigenvalues of the $(L+1)\times(L+1)$ matrix $\mathbf{T}$ are the perfect foresight roots λ_i ($i=1, ..., L+1$) of the characteristic polynomial P. A convenient form for the $(L+1)\times 1$ eigenvector $\mathbf{u}_i$ ($i=1, ..., L+1$) associated with λ_i is obtained by using the relations between the coefficients and the roots of P. Namely:

$$P(\lambda)=0 \Leftrightarrow \prod_{i=1}^{L+1} (\lambda-\lambda_i)=0$$

$$\Leftrightarrow \lambda^{L+1} - \left(\sum_{i=1}^{L+1} \lambda_i\right) \lambda^L + \cdots + (-1)^{L+1} \prod_{i=1}^{L+1} \lambda_i = 0.$$

Let us identify the coefficients in the expression above and those of P. One gets

$$(-1)^{l+1}\,\sigma_{l+1}(\mathscr{L})=\delta_l/\gamma \qquad \text{for} \quad l=0, ..., L+1, \text{ with } \delta_0 \equiv 1, \tag{29}$$

where $\mathscr{L}$ is the set of all the perfect foresight roots, and $\sigma_l(\mathscr{L})$ is the lth ($l=1, ..., L+1$) elementary symmetric polynomial of $\mathscr{L}$, i.e., the sum on all the possible products over l different elements of $\mathscr{L}$:

$$\sigma_1(\mathscr{L}) \overset{\text{def}}{=} \sum_{1 \leqslant i_1 < \cdots < i_l} (\lambda_{i_1} \cdots \lambda_{i_l}).$$

By definition $\mathbf{u}_i$ is such that $\mathbf{T}\mathbf{u}_i=\lambda_i\mathbf{u}_i$ ($i=1, ..., L+1$). It can be shown that

$$\mathbf{u}_i=(\lambda_i^L, \lambda_i^{L-1}, ..., 1)',$$

where the symbol $'$ denotes the (vector) transpose. The perfect foresight trajectory that is restricted to W_k is such that $\mathbf{x}_t$ is a linear combination of all the $\mathbf{u}_i$ but $\mathbf{u}_k$. This condition writes $\det(\mathbf{x}_t, \mathbf{U}_k)=0$ where $\mathbf{U}_k$ is the $(L+1)\times L$ matrix whose columns are all the $\mathbf{u}_i$ but $\mathbf{u}_k$. Developing $\det(\mathbf{x}_t, \mathbf{U}_k)$ with respect to its first column $\mathbf{x}_t$ leads to

$$\Leftrightarrow x_t = \sum_{l=1}^{L} a_l x_{t-l} = \sum_{l=1}^{L} ((-1)^{l+1}\,\Delta_l/\Delta_0)\, x_{t-l},$$

where Δ_0 is the determinant of the $(L-1)\times(L-1)$ Vandermonde matrix and $\Delta_l = \sigma_l(\mathcal{L}_k)\,\Delta_0$ (see respectively Ch. 10 and Ex. 10.12 in Ramis, Deschamps and Odoux [15]) and $\mathcal{L}_k$ is the set of all the perfect foresight roots but λ_k. Thus a_l is as stated in Lemma 1, namely:

$$a_l^k = (-1)^{l+1}\,\sigma_l(\mathcal{L}_k). \tag{30}$$

It remains consequently to prove that vectors $(a_1^k, ..., a_L^k)$ are the solutions to (19), i.e., $a_l^k = \bar{\beta}_l^k$ ($l = 1, ..., L$ and $k = 1, ..., L+1$). We show this directly by replacing β_l in (19) by the coefficient a_l^k given (30) for any k given. Observe first that, for $\gamma \neq 0$, the lth equation of (19) rewrites:

$$(1/\gamma + \beta_1)\,\beta_l = -\beta_{l+1} - \delta_l/\gamma. \tag{31}$$

Since $a_1^k = \sigma_1(\mathcal{L}_k)$ and $(-1/\gamma) = \sigma_1(\mathcal{L})$, we have $(1/\gamma + \beta_1^k) = \lambda_k$. Hence the Lth equation of (19) becomes (recall that $\beta_{L+1} = 0$)

$$\beta_L = \frac{\delta_L/\gamma}{\lambda_k} = \frac{1}{\lambda_k}\,(-1)^{L+1}\,\sigma_{L+1}(\mathcal{L}), \tag{32}$$

where we used (29) with $l = L$. By definition, $\sigma_{L+1}(\mathcal{L})$ is the product over all the perfect foresight roots. Therefore (32) rewrites:

$$\beta_L = (-1)^{L+1} \prod_{\substack{i=1\\ i\neq k}}^{L+1} \lambda_i \stackrel{\text{def}}{=} (-1)^{L+1}\,\sigma_L(\mathcal{L}_k).$$

As (30) shows, β_L is equal a_L^k, and so $a_L^k = \bar{\beta}_L^k$. Assume now that $\beta_{l+1} = a_{l+1}^k$ ($l < L$). We are going to prove that $\beta_l = a_l^k$ ($l < L$). Indeed, by using (29) and (30) in (31) for some l, one obtains:

$$\begin{aligned}\beta_l &= \frac{1}{\lambda_k}((-1)^{l+2}\,\sigma_{l+1}(\mathcal{L}_k) + (-1)^{l+1}\,\sigma_{l+1}(\mathcal{L}))\\ &= \frac{1}{\lambda_k}(-1)^{l+1}(-\sigma_{l+1}(\mathcal{L}_k) + \sigma_{l+1}(\mathcal{L})).\end{aligned}$$

But, by definition, $\sigma_{l+1}(\mathcal{L}_k)$ is the sum on all the products over $(l+1)$ different elements of $\mathcal{L}_k$ while $\sigma_{l+1}(\mathcal{L})$ is the sum on all the products over $(l+1)$ different elements of $\mathcal{L}$. The difference between these sums is just the sum on all the products over $(l+1)$ different perfect foresight roots, provided that each of the remaining products includes the root λ_k. Hence:

$$-\sigma_{l+1}(\mathcal{L}_k) + \sigma_{l+1}(\mathcal{L}) = \lambda_k \sigma_l(\mathcal{L}_k) \Rightarrow \beta_l = (-1)^{l+1}\,\sigma_l(\mathcal{L}_k) = a_l^k.$$

Since $\beta_l = a_l^k$, we have $\bar{\beta}_l^k = a_l^k$ for any l ($l = 1, ..., L$) and for any given k, which completes the proof of Lemma 1.

5.2. *Proof of Theorem* 5

The dynamics (23) in the neighborhood of $\bar{\boldsymbol{\beta}}^k$ ($k = 1, ..., L+1$) is governed by the L eigenvalues μ_j^k ($j = 1, ..., L$) of the $L \times L$ Jacobian matrix $\boldsymbol{\Psi}_k$ of (23) calculated at $\bar{\boldsymbol{\beta}}^k$:

$$\boldsymbol{\Psi}_k = \begin{pmatrix} -2\gamma\bar{\beta}_1^k - 1 & -\gamma & 0 & \cdots & 0 \\ -\gamma\bar{\beta}_2^k & -\gamma\bar{\beta}_1^k - 1 & -\gamma & \cdots & 0 \\ -\gamma\bar{\beta}_3^k & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\gamma \\ -\gamma\bar{\beta}_L^k & 0 & \cdots & 0 & -\gamma\bar{\beta}_1^k - 1 \end{pmatrix}.$$

These eigenvalues are the L solutions of the equation $\det(\boldsymbol{\Psi}_k - \mu \mathbf{I}_L) = 0$, where $\mathbf{I}_L$ denotes the $L \times L$ identity matrix. Let develop $\det(\boldsymbol{\Psi}_k - \mu \mathbf{I}_L)$ with respect to its first column:

$$(-1)^2 (-2\gamma\bar{\beta}_1^k - 1 - \mu) \begin{vmatrix} -\gamma\bar{\beta}_1^k - 1 - \mu & -\gamma & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\gamma \\ 0 & \cdots & \cdots & -\gamma\bar{\beta}_1^k - 1 - \mu \end{vmatrix}$$

$$+ (-1)^3 (-\gamma\bar{\beta}_2^k) \begin{vmatrix} -\gamma & 0 & \cdots & 0 \\ 0 & -\gamma\bar{\beta}_1^k - 1 - \mu & -\gamma \cdots & 0 \\ \vdots & \vdots & \ddots & -\gamma \\ 0 & 0 & \cdots & -\gamma\bar{\beta}_1^k - 1 - \mu \end{vmatrix} + \cdots$$

$$+ (-1)^{L+1} (-\gamma\bar{\beta}_L^k) \begin{vmatrix} -\gamma & 0 & \cdots & 0 \\ -\gamma\bar{\beta}_1^k - 1 - \mu & -\gamma & 0 \cdots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & -\gamma \end{vmatrix}.$$

Therefore,

$$\det(\boldsymbol{\Psi}_k - \mu \mathbf{I}_L) = 0$$
$$\Leftrightarrow Q_k(z) = (1 + \gamma\bar{\beta}_1^k + z)^L - \sum_{l=1}^{L} (-1)^l \gamma^k \beta_l^k (1 + \gamma\bar{\beta}_1^k + z)^{L-l} = 0. \quad (33)$$

The expression of $\bar{\beta}_l^k$ given in (30) lead to

$$(-1)^l \gamma^l \bar{\beta}_l^k = (-\gamma)^l (-1)^{l+1} \sigma_l^k(\mathscr{L}_k) = (-1)^{l+1} \sigma_l^k(\{-\gamma \mathscr{L}_k\}),$$

where the L elements of the set $\{-\gamma \mathscr{L}_k\}$ are $(-\gamma\lambda_j)$ for $j=1, ..., L+1$ and $j \neq k$. We reintroduce now the expression of $(-1/\gamma)$ given in (29):

$$(-1)^l \gamma^l \bar{\beta}_l^k = (-1)^{l+1} \sigma_l^k \left(\left\{ \frac{1}{\sum_{i=1}^{L+1} \lambda_i} \mathscr{L}_k \right\} \right). \tag{34}$$

Let now $y \equiv 1 + \gamma \bar{\beta}_1^k + z$ so that (33) rewrites (with (34)):

$$Q_k(y) = y^L - \sum_{l=1}^{L} (-1)^{l+1} \sigma_l^k \left(\left\{ \frac{1}{\sum_{i=1}^{L+1} \lambda_i} \mathscr{L}_k \right\} \right) y^{L-l} = 0.$$

This shows that the L roots of Q_k are $\lambda_j / \sum_{i=1}^{L+1} \lambda_i$ for $j=1, ..., L+1$ and $j \neq k$. Namely:

$$1 + \gamma \bar{\beta}_1^k + \mu_j^k = \frac{\lambda_j}{\sum_{i=1}^{L+1} \lambda_i}.$$

Since $\bar{\beta}_1^k = \sigma_1(\mathscr{L}_k) = \sigma_1(\mathscr{L}) - \lambda_k$, one finally gets (by using again the expression of $(-1/\gamma)$ given in (29):

$$\mu_j^k = \frac{\lambda_j}{\sum_{i=1}^{L+1} \lambda_i} - 1 - \gamma \bar{\beta}_1^k = \frac{\lambda_j}{\sum_{i=1}^{L+1} \lambda_i} - \frac{\sum_{i=1}^{L+1} \lambda_i}{\sum_{i=1}^{L+1} \lambda_i} + \frac{\sum_{i=1}^{L+1} \lambda_i - \lambda_k}{\sum_{i=1}^{L+1} \lambda_i}$$

$$\Leftrightarrow \mu_j^k = \frac{\lambda_j - \lambda_k}{\sum_{i=1}^{L+1} \lambda_j} \qquad \text{for} \quad j \neq k.$$

A stationary $EGR(L)$ $\bar{\boldsymbol{\beta}}^k$ is locally stable under learning if and only if $\mu_j^k < 0$ for each $j=1, ..., L+1$ but $j \neq k$ $(k=1, ..., L+1)$. The result follows.

5.3. *Local Extended Growth Rate Dynamics*

Recall that, from (29), we have:

$$\delta_L/\gamma = (-1)^{L+1} \sigma_{L+1}(\mathscr{L}) \qquad \text{and} \qquad \delta_{L-1}/\gamma = (-1)^L \sigma_L(\mathscr{L}). \tag{35}$$

If $\delta_L \neq 0$, then $\sigma_{L+1}(\mathscr{L}) = \prod_{i=1}^{L+1} \lambda_i \neq 0$, so that $\lambda_i \neq 0$ whatever i is $(i=1, ..., L+1)$. Now, it follows from (30) that:

$$\bar{\beta}_L^k = (-1)^{L+1} \sigma_L(\mathscr{L}_k) \qquad \text{and} \qquad \bar{\beta}_{L-1}^k = (-1)^L \sigma_{L-1}(\mathscr{L}_k) \tag{36}$$

Since $\sigma_L(\mathscr{L}_k) = \prod_{i=1, i \neq k}^{L+1} \lambda_i$ and since $\lambda_i \neq 0$ when $\delta_L \neq 0$, it follows that $\bar{\beta}_L^k \neq 0$ whatever k is $(k=1, ..., L+1)$. It remains consequently to show that, when $\delta_L \neq 0$, we have:

$$\frac{\bar{\beta}_{L-1}^k}{\bar{\beta}_L^k} \neq \frac{\delta_{L-1}}{\delta_L} = \frac{\delta_{L-1}/\gamma}{\delta_L/\gamma}. \tag{37}$$

Using definitions (35) and (36), this condition is equivalent to:

$$\frac{\sigma_L(\mathscr{L})}{\sigma_{L-1}(\mathscr{L}_k)} \neq \frac{\sigma_{L+1}(\mathscr{L})}{\sigma_L(\mathscr{L}_k)}.$$

Observe now that $\sigma_{L+1}(\mathscr{L})/\sigma_L(\mathscr{L}_k) = \lambda_k$, so that (37) rewrites

$$\lambda_k \sigma_{L-1}(\mathscr{L}_k) - \sigma_L(\mathscr{L}) \neq 0 \Leftrightarrow \sigma_L(\mathscr{L}_k) \neq 0$$

which is true when $\delta_L \neq 0$. This proves the claim.

5.4. *Proof of Proposition* 7

The dynamics (28) in the neighborhood of $\bar{\boldsymbol{\beta}}^k$ is obtained by linearizing (28) at $\bar{\boldsymbol{\beta}}^k$:

$$\gamma \mathbf{B}_k(\boldsymbol{\beta}^k(t+1) - \bar{\boldsymbol{\beta}}^k) = -(1+\gamma\bar{\beta}_1^k)\,\mathbf{I}_L(\boldsymbol{\beta}^k(t) - \bar{\boldsymbol{\beta}}^k),$$

where

$$\boldsymbol{\beta}^k(t) \equiv \begin{pmatrix} \beta_1^k(t) \\ \vdots \\ \vdots \\ \beta_{L-1}^k(t) \\ \beta_L^k(t) \end{pmatrix}, \qquad \mathbf{B}_k \equiv \begin{pmatrix} \bar{\beta}_1^k & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ \bar{\beta}_{L-1}^k & 0 & \cdots & 0 & 1 \\ \bar{\beta}_L^k & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Assume that det $\mathbf{B}_k = \bar{\beta}_L^k \neq 0$, i.e., $\delta_L \neq 0$ (see Section 5.3). One can express $(\boldsymbol{\beta}^k(t+1) - \bar{\boldsymbol{\beta}}^k)$ as a function of $(\boldsymbol{\beta}(t) - \bar{\boldsymbol{\beta}}^k)$ for $\boldsymbol{\beta}^k(t)$ near $\bar{\boldsymbol{\beta}}^k$. Namely:

$$(\boldsymbol{\beta}^k(t+1) - \bar{\boldsymbol{\beta}}^k) = \mathbf{F}_k(\boldsymbol{\beta}^k(t) - \bar{\boldsymbol{\beta}}^k) \qquad \text{where} \quad \mathbf{F}_k \equiv -\frac{1+\gamma\bar{\beta}_1^k}{\gamma}\,\mathbf{B}_k^{-1}.$$

Let θ_j^k and b_j^k $(j=1, ..., L)$ denote the L eigenvalues of $\mathbf{F}_k$ and $\mathbf{B}_k$ $(k=1, ..., L+1)$ respectively. Then we have: $\theta_j^k = -(1+\gamma\bar{\beta}_1^k)/(\gamma b_j^k)$. Remark now that $\boldsymbol{\Psi}_k = -\gamma\mathbf{B}_k - (1+\gamma\bar{\beta}_1^k)\,\mathbf{I}_L$, so that: $\mu_j^k = -\gamma b_j^k - \gamma\bar{\beta}_1^k - 1$. Hence,

$$\theta_j^k = \frac{1+\gamma\bar{\beta}_1^k}{(1+\gamma\bar{\beta}_1^k) + \mu_j^k} = \frac{\lambda_k}{\lambda_j},$$

where the last equality comes by replacing $\bar{\beta}_1^k$ by $\sigma_{L-1}(\mathscr{L}_k)$ (see (30)). A given stationary $EGR(L)$ $\bar{\boldsymbol{\beta}}^k$ $(k=1, ..., L+1)$ is locally determinate in (28) if and only if $|\theta_j^k| > 1$ $(j=1, ..., L+1$ and $j \neq k)$, which holds true if and only if $k = L+1$. The result follows.

REFERENCES

1. J. Duffy, On learning and the non-uniqueness of equilibrium in an overlapping generations model with fiat money, *J. Econ. Theory* **64** (1994), 541–553.
2. G. Evans, The fragility of sunspots and bubbles, *J. Monet. Econ.* **23** (1989), 297–317.
3. G. Evans and S. Honkapohja, On the robustness of bubbles in linear rational expectations models, *Int. Econ. Rev.* **33** (1992), 1–14.
4. G. Evans and S. Honkapohja, Stochastic gradient learning in the cobweb model, *Econ. Lett.* **61** (1998), 333–337.
5. J. M. Grandmont, Expectations formation and stability of large socioeconomic systems, *Econometrica* **66** (1998), 741–783.
6. J. M. Grandmont and G. Laroque, Stability of cycles and expectations, *J. Econ. Theory* **40** (1986), 138–151.
7. J. M. Grandmont and G. Laroque, Stability, expectations and predetermined variables, *in* "Essays in Honor of Edmond Malinvaud" (P. Champsaur *et al.*, Eds.), Vol. 1, MIT Press, Cambridge, MA, 1990.
8. J. M. Grandmont and G. Laroque, Economic dynamics with learning: some instability examples, *in* "Equilibrium Theory and Applications: Proceedings of the Sixth International Symposium in Economic Theory and Econometrics" (W. A. Barnett *et al.*, Eds.), Cambridge University Press, Cambridge, UK, 1991.
9. R. Guesnerie, Successes and failures in coordinating expectations, *Europ. Econ. Rev.* **37** (1992), 243–268.
10. R. Guesnerie and M. Woodford, Stability of cycles with adaptive learning rules, *in* "Equilibrium Theory and Applications: Proceedings of the Sixth International Symposium in Economic Theory and Econometrics" (W. A. Barnett *et al.*, Eds.), Cambridge University Press, Cambridge, UK, 1991.
11. L. Ljung, Analysis of recursive stochastic algorithms, *IEEE Trans. Auto. Control* **22** (1977), 551–575.
12. L. Ljung and T. Soderström, "Theory and Practice of Recursive Identification," MIT Press, Cambridge, MA, 1983.
13. R. Lucas, Adaptive behavior and economic theory, *J. Business* **59** (1986), 401–426.
14. A. Marcet and T. Sargent, Convergence of least-squares learning mechanisms in self-referential stochastic models, *J. Econ. Theory* **48** (1989), 337–368.
15. E. Ramis, C. Deschamps, and J. Odoux, "Cours de Mathématiques Spéciales," Tome 1, Algèbre, Paris, Masson, 1974.
16. P. Reichlin, Equilibrium cycles in an overlapping generations economy with production, *J. Econ. Theory* **40** (1986), 89–102.
17. H. Robins and S. Monro, A stochastic approximation method, *Ann. Math. Statist.* **22** (1951), 400–407.
18. M. Woodford, Learning to believe in sunspots, *Econometrica* **58** (1990), 277–307.