Adaptive learning in the Cobweb with an endogenous gain sequence.

D. Colucci, V. Valori*

September 2004

Abstract

We develop a learning rule that generalises the well known fading memory learning in the sense that the weights attached to the available time series data are not constant and are updated in light of the forecast error(s). The underlying idea is that confidence in the available data will be low when large errors have been realized (e.g. in times of higher volatility) and vice versa. A class of functional forms compatible with this idea is analysed in the context of a standard Cobweb model with boundedly rational agents. We study the problem of convergence to the perfect foresight equilibrium and give conditions that ensure the coexistence of different attractors. We refer to experimental and numerical evidence to establish the possible range of application of the generalised fading memory learning.

Journal of Economic Literature Classification Numbers: C91, D83, E32.

Key Words: learning; convergence to rational expectations; cobweb dynamics.

1 Introduction

The issues of expectation formation and learning have attracted a lot of scientific work in the last few decades both in the economic-theoretic and in the applied/experimental literature. It is widely recognized that such issues touch crucial aspects of many dynamic economic models in which, beside

*DiMaD - Dipartimento di Matematica per le Decisioni - Università degli Studi di Firenze - Via C. Lombroso, 6/17 - 50134 Firenze - Italy. E-mail: domenico.colucci@dmd.unifi.it, vincenzo.valori@dmd.unifi.it

the usual expectations feedback, the beavioural aspects underlying agents' choices are explicitly considered. However there is no general agreement on the mechanism of expectation formation. The wide use of strong forms of individual rationality and optimising behaviour in mainstream economic models is probably the result of the confidence in a process of convergence towards rational expectations seen as steady states of some adaptive process (see Lucas [19]). Nonetheless, the issue is quite complex and the growing evidence coming from experiments has not entirely clarified the matter: quoting Camerer [8]

> An important source of disagreement between psychologists and economists concerns learning. Psychologists often suspect that the immediate, frequent, exogenous feedback subjects receive in economic experiments overstates how well people learn in natural economic settings. Economists, in contrast, think that experiments understate the rate of natural learning because context, access to advice, higher incentives, and added time to reflect or calculate are absent from experiments, and probably improve performance in natural settings.

Several papers have been dedicated to the process of expectations formation of real economic agents. Among these a good number indicate the use of various forms of adaptive expectations, whereas rational expectations are, in most cases, not supported by the data, basically on the grounds that forecast errors exhibit autocorrelation and correlation with observables. For example Schmalensee [22], Smith et al. [24] and Williams [25] conduct experiments with human subjects, whereas Figlewski and Wachtel [11] and Lovell [18] are econometric studies of survey data on expectations. In particular both Schmalensee [22] and Figlewski and Wachtel [11] suggest that agents' expectations are best described as adaptive, although with a parameter that is nonconstant across agents and through time following the rate of "uncertainty" perceived by the agents. Our working assumption in this paper is that, in specific settings, economic agents are bound to behave according to a mechanism of expectation formation in which uncertainty is measured in terms of forecast errors and the reaction to higher uncertainty is to put more weight on the last available observation (therefore approaching a form of myopic expectation). We model this by way of a rather general form of "fading memory" process, which we study chiefly in the context of a standard Cobweb setting. In the broader family of adaptive approaches to learning, fading memory learning has been studied in Bischi-Gardini [4, 3], Bischi-Naimzada [5], Barucci [1, 2], Pötzelberger and Sögner [21] and Chiarella *et al.* [9].

In this paper we investigate a learning mechanism that generalises the fading memory process. The rationale of the generalisation consists in the fact that the weights used by the agents to extract information (and in fact, expectations) from past data are not considered constant but instead, on the basis of the error in their predictions (once it is known), they are updated according to the idea mentioned above: the confidence in the reliability of the available observations as instruments to form expectations lowers when the forecast error is significantly high and vice versa. This fact couples with the decrease in the weight attached to older data, which is a known feature of fading memory, to build the learning device which is the object of this paper. The spirit of the Generalised Fading Memory (GFM) is rather similar to the learning mechanism studied in the paper by Marcet and Nicolini [20] in which agents use either a simple average or a form of constant gain learning, at each period endogenously selecting one rule: in fact the GFM can be thought of as a sort of 'smooth' version of the Marcet and Nicolini learning mechanism.

The paper is organised as follows: in the next section we introduce fading memory (FM). Section 3 defines its generalisation and in Subsection 3.1 we derive analytic results concerning conditions to have convergence towards perfect foresight equilibria, learnability of cycles and emergence of multiple attractors. In Section 4 we work with a specific functional form to argue the utility of GFM on experimental grounds and to show numerically the consequences of having multiple attractors when exogenous shocks are allowed for.

2 Fading memory learning

We briefly present the fading memory learning. Expectations for the future are a weighted average of available data. The general form of a model with fading memory can thus be written as

$$\begin{cases} x_{t+1}^e = \sum_{k=0}^n a_k x_k \\ \sum_{k=0}^n a_k = 1, \ \ a_k \geq 0 \end{cases} \tag{1}$$

where typically the evolution of the state variable, x, depends on traders expectations about it, x^e, through a map F describing the expectations feedback. Possibly there can be a time gap (when $n \neq t$) that can imply either a forward-looking or a backward-looking feature. Here we shall stick to the simple case $x_t = F(x_t^e)$, allowing for standard regularity conditions on F. The weights used are a normalized geometric progression:

$$\begin{cases} a_k = \frac{\rho^{t-k}}{W_n} & \rho \in (0,1) \\ W_t = \sum_{i=0}^t \rho^i \end{cases} \tag{2}$$

Therefore (1) can be transformed into:

$$\begin{cases} x^e_{t+1} = \frac{1}{W_t} x_t + \rho \frac{W_{t-1}}{Wt} x^e_t \\ W_{t+1} = 1 + \rho W_t \qquad\qquad \rho \in (0,1) \end{cases} \tag{3}$$

with the initial condition $W_0 = 1$. Allowing t to tend to infinity W_t tends to $\frac{1}{1-\rho}$ so that, defining $\alpha = 1 - \rho$, (3) is approximated by the *limiting map*

$$x^e_{t+1} = x^e_t + \alpha(x_t - x^e_t) \tag{4}$$

which is a standard model with adaptive expectations. There are two extreme cases: when $\rho = 1$ all the past observations receive the same weight, so that we are left with

$$x^e_t = \frac{1}{t} \sum_{k=0}^{t-1} x_k$$

that is, with a simple average (see Bray [7]). Conversely when $\rho = 0$ expectations are myopic (static)

$$x^e_t = x_{t-1} .$$

Now, going back to (3), with the substitution $\alpha_t = \frac{1}{W_n}$, we get

$$\begin{cases} x^e_{t+1} = x^e_t + \alpha_t(F(x^e_t) - x^e_t) \\ \alpha_{t+1} = \frac{\alpha_t}{\alpha_t + \rho} \qquad\qquad 0 < \alpha_0 < 1 \end{cases} \tag{5}$$

One way to deal with the dynamic properties of (3) is by studying its limiting map (4). This approach, taken for example in Barucci [1, 2], makes sense thanks to fact that several facts of the (local) dynamics of (3) carry through to its limiting map. More precisely, locally attracting fixed points and periodic orbits of the limiting map (4) correspond to identical objects for the original system (3) (see Bischi-Gardini [3] for details and proofs). These results though, hold for the local analysis of steady states and cycles only, whereas the limiting map is uninformative about the global dynamic behaviour.

Fading memory learning, be it expressed as in (3) or as in (5), is clearly a very basic form of expectations updating. In particular, in (5), the internal variable α_t that determines the correction on the previous expectation in the direction of the last error, is itself only dependent on its own path and on a parameter. The way α is updated is therefore completely independent from external signals. Indeed, α_t converges to $1 - \rho$ regardless of the dynamics of the state varaiable x. Our aim in the sequel is to generalise this mechanism to include past prediction performances in the determinants of the dynamics of α_t by endogenising the parameter ρ.

3 A generalisation of fading memory

We shall assume that the individuals, on the basis of a law that uses the most recent forecasting error as a benchmark, revise the weight attributed to the past observations. At a generic time t agents weigh the most recent observation against their relevant forecast: on the basis of the error $(x_t - x_t^e)$ they check the significance of the past data to trace the recent evolution of the state variable and set their weight, $\rho_t = H\left(x_t - x_t^e\right)$, on the calculation of the new expectations. Clearly, the mechanism is not fully specified until we impose some assumptions on the function H. Our basic idea regarding the behavioural rationale behind H is the following. Expectations in this framework can be interpreted as a weighted mean of the available data; therefore a large forecast error is interpreted as the failure of the data to be informative about the present tendency of the state variable, e.g. as in case of structural breaks or important exogenous shocks. Hence the most recent observation gets to play a dominant role in shaping the forecast after a significant error, whereas a low weight is attributed to the bulk of older observations (by the choice of a small ρ_t): we assume that the function H incorporates this feature. Remark that the variable ρ_t can be interpreted as a voluntary choice of how much it is worth recalling, the choice being the result of a simple form of assessment of the significance of the available information in terms of predicting the future. The recursive form of the system writes

$$\begin{cases} x_{t+1}^e = x_t^e + \alpha_t(F(x_t^e) - x_t^e) \\ \alpha_{t+1} = \frac{\alpha_t}{\alpha_t + \rho_{t+1}} \\ \rho_{t+1} = H(x_{t+1} - x_{t+1}^e) \end{cases} \tag{6}$$

This can be seen as a natural generalisation of the mechanism of fading memory (compare with (5)): therefore we shall label it Generalised Fading Memory (GFM).

A functional form for H which fits this description is a bell-shaped, symmetric, H function: indeed in the last section we shall develop on a particular gaussian functional form. Obviously the basic fading memory case corresponds to the degenerate choice of a constant H between 0 and 1. Therefore, the way GFM produces expectations lies, at each step, between the two extreme cases of the FM (myopic expectations and simple average): in fact when the forecast error is low (assuming $H\left(0\right) = 1$) the GFM is very close to a simple average, whereas with large errors we can expect to have α close to 1 (as with mypoic expectations). This feature is very much in the spirit of the learning rule used in Marcet and Nicolini [20], in the sense that GFM can be thought of as a smooth version of that rule.

It is probably useful to describe in detail the sequence of "moves" at time t which are recorded by the system (6):

$$x_t^e \rightarrow x_t \rightarrow \rho_t \rightarrow \alpha_t \rightarrow x_{t+1}^e \rightarrow x_{t+1} \cdots$$

Therefore the system's recursive structure is fully compatible with off-equilibrium dynamics in the sense that it is not affected by any contemporaneity puzzles in the relevant variables which are sometimes encountered in equilibrium models. Indeed this allows us to numerically simulate the dynamics and to test the practical descriptive power of the system in lab experiments.

Notice that, despite its appearance, (6) is in fact a standard two-dimensional dynamical system. In fact the ρ variable can be easily eliminated from the system bringing us to the equivalent (though hardly more suggestive) form

$$\begin{cases} x_{t+1}^e = x_t^e + \alpha_t \left(F(x_t^e) - x_t^e\right) = E\left(x_t^e, \alpha_t\right) \\ \alpha_{t+1} = G\left(x_t^e, \alpha_t\right) \end{cases} \tag{7}$$

with the definition $G\left(x_t^e, \alpha_t\right) = \frac{\alpha_t}{\alpha_t + H(F(x_t^e + \alpha_t(F(x_t^e) - x_t^e)) - x_t^e - \alpha_t(F(x_t^e) - x_t^e))}$.

3.1 Some results

Some salient properties of GFM regard its steady states. Remark that a steady state for system (6) necessarily is of the form (x^*, α^*) where x^* is a fixed point of the expectations feedback map F, and $\alpha^* = 1 - H(0)$.[1] The steady state can be analysed locally in the usual way by means of the Hartmann-Grobmann theorem. We state the following proposition focusing on the relation between $F'(x^*)$ and $H(0)$.

Proposition 1 *A steady state (x^*, α^*) of (7) is locally stable and hyperbolic if:*

$$-\frac{1 + H(0)}{1 - H(0)} < F'(x^*) < 1$$

and

$$-1 < H(0) < 1$$

Proof. In the Appendix.

Unsurprisingly, none of the additional features of the GFM appear in the conditions for local stability with respect to the baseline FM case (with

[1] In fact from a purely mathematical perspective $(x, 0)$ are also steady states for any x: but these points are meaningless unless $x = x^*$ because when $\alpha = 0$ anything goes for the first equation (i.e. expectations are constant regardless of the observable x). The following analysis assumes that $\alpha = 0$ can be ruled out.

$\rho = H(0)$). Our generalisation indeed matters when the forecast error is not zero: otherwise the two (GFM and FM) are in fact the same.

Another feature worth attention of this family of learning rules is the fact that cycles of period $p \geq 2$ cannot be detected.

Proposition 2 *Let $(x^1, \alpha^1), \ldots, (x^p, \alpha^p)$ be a cycle of period $p \geq 2$ for the map (7) with $H(\cdot) > 0$, then $x_t^e - x^i = 0$ implies $x_{t+1}^e - x^{i+1} \neq 0$.*

Proof. In the Appendix.

In other words cycles of any period ≥ 2 are not learnable under the GFM. This fact can help us locate the GFM in a sort of "sophistication" ranking of learning rules in terms of how far they can go in detecting dynamic patterns (in the spirit of Grandmont [13]): it is clearly less sophisticated than SAC learning (see Hommes and Sorger [16]) or recursive least squares (provided a constant term is included in the estimation). On the other hand the possibilities of "learning" a steady state are enlarged with respect to simple adaptive expectations. To show this precisely, let us now underline a global dynamic property of system (7). It is well known that in the case of a fixed α, i.e. with simple adaptive expectations, a steady state $x^* = F(x^*)$ with $F'(x^*) < -1$ is locally stable[2] whenever $0 < \alpha < \hat{\alpha} < 1$ for a suitable $\hat{\alpha}$. When α is updated using GFM instead, convergence to a steady state can happen starting from whatever initial value for α. This is described in the following Proposition.

Proposition 3 *Let (x^*, α^*) be a locally stable steady state for (7) and assume that $0 \leq H(\cdot) \leq \sigma < 1$. Then for all $\alpha > 0$ there is a neighborhood I_α of x^* such that for all $x \in I_\alpha$ the point (x, α) belongs to the basin of attraction of (x^*, α^*).*

The proof of this Proposition can be found in [10]. The consequence is that the basin of attraction of (x^*, α^*) extends to any choice of positive α provided x is suitably chosen. From a geometric point of view, the basin of attraction of the steady state (x^*, α^*), contains a region that covers points (x, α) which would not be compatible with stability (would not converge to (x^*, α^*)) in the 1-dimensional model in which α is taken as a fixed parameter. A rather natural question is to ask whether such region, as a subset of the basin of attraction, is in some sense minimal (the basin is actually much bigger) or whether it can be considered a good approximation. A partial answer is given by the following Proposition.

[2] The case $F'(x^*) \geq -1$ is not interesting because if $|F'(x^*)| < 1$ the steady state x^* is always locally stable for the map $x_{t+1}^e = x_t^e + \alpha (F(x_t^e) - x_t^e)$, whereas if $F'(x^*) > 1$ the steady state x^* is always locally unstable. The case $F'(x^*) = 1$ implies non-hyperbolicity.

Proposition 4 *Suppose F is decreasing and x^* is a steady state with $F'(x^*) < -1$. Then there is a class of bell-shaped $H(\cdot)$ functions for which the system with GFM learning shows multiplicity of attractors.*

Proof. In the Appendix.

An interesting consequence of this fact is that, in the presence of shocks, the system can actually switch among various dynamical regimes. Indeed this phenomenon appears quite clearly in our simulations in the last section.

4 The GFM in practice

We now consider a particular specification to the function H through which agents adapt to their forecast error. In fact it is the following class of functions:

$$H(x_t - x_t^e) = ke^{-[h(x_t - x_t^e)]^2} + d \tag{8}$$

with $d, k \geq 0$ and $k + d \leq 1$. An example of this type of function (basically a gaussian curve) is depicted in Figure 1.

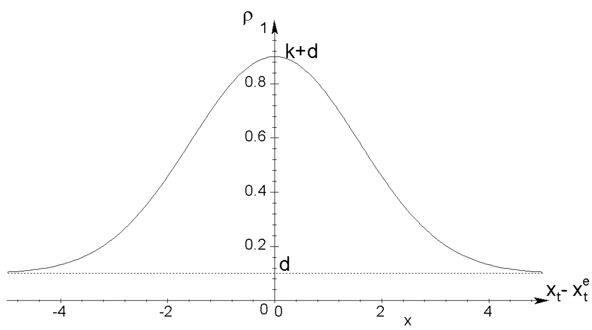

Figure 1: Function $H(\cdot)$ with $d = 0.1$ and $k = 0.8$

It is a class of functions in which height, base value and dispersion depend on the parameter values $d + k$, k and h; given the constraints on the parameters the entire class satisfies our interpretation of a rule that should wipe out most of the weight attributed to old data as a consequence of a significant prediction error and vice versa. It is straightforward to check that the (perfect foresight) steady state (p^*, α^*) for the equation

$$\begin{cases} x_{t+1}^e = x_t^e + \alpha_t(F(x_t^e) - x_t^e) \\ \alpha_{t+1} = \frac{\alpha_t}{\alpha_t + ke^{-\left[h\left(x_{t+1} - x_{t+1}^e\right)\right]^2} + d} \end{cases} \tag{9}$$

has $\alpha^* = 1 - k - d$.

We now introduce the cobweb model studied by Hommes in several papers: this model underlies what we do in the sequel. Consider a single market for a perishable good that requires a time period to be produced. Demand q_t^d for the good depends on its current price p_t. Due to the production lag, decisions on the supply side depend on the price expected by producers. Assume a linearly decreasing demand

$$D(p_t) = a - bp_t \tag{10}$$

and an S-shaped supply curve, with a unique inflection point $\bar{p}$ which we specify as

$$S(p_t^e) = \tanh(\lambda(p_t^e - c)) + 1 \tag{11}$$

as in Hommes et al. [17] and elsewhere. Finally, the price is determined by market clearing:

$$p_t = \frac{a - \tanh\lambda(p_t^e - 6) - 1}{b} \tag{12}$$

As usual there is a unique fixed point.

4.1 Experimental data

We try to understand whether the GFM, beside our interpretation, incorporates any of the features of real agents' behaviour. A simple but potentially rich setting for running experiments is the well known cobweb model. Many papers have dealt with the dynamic properties and the implications of this model with bounded rationality under various specifications. Experiments simulating human behaviour in cobweb-type situations though, are not very frequent in the literature (the most recent experimental paper concerning the cobweb that we are aware of is Sonnemans et al. [23]). One possible reason is that the original hog-cycle idea motivating the model is not very easily replicated in the laboratory[3]. Still the evidence coming from such experiments is interesting because it can tell something about the way people go about forecasting the future when they have very little or no prior information about the object of their predictions and there is a strong feedback from their predictions or actions on the observed state. Hommes et al. [17] have conducted one person experiments in which agents' predictions are used

[3] A rather brutal question to explain this difficulty is the following: can a lab experiment in which subjects are required to produce a prediction in about a minute replicate the type of reasoning and foresight individuals would put into action in a yearly decision problem? It must be stressed though, that this potential problem arises for many other experimental settings.

to generate the price dynamics of a cobweb model. The model underlying the experiment is basically described by equation (12) to which a random shock is to be added: the shocks are either drawn from the uniform or from the normal distribution. We have analysed the data of that experiment to understand whether any inference can be drawn on the subjects' behaviour. In particular we try and answer the questions: does the GFM produce a real advantage with respect to FM in terms of modelling the subjects' behaviour? How does it fare compared with other simple learning rules? To do so we assume that agents form expectations using a predictor chosen among a given set and, once chosen, they stick to it. Furthermore, we suppose that agents misreport their expected price[4]. This error can be the result of a lack of full attention to the data due to the short time agents have to form expectations in an experimental environment.

In detail, we suppose that agents form expectations according to a law of the form

$$x_{t+1}^e = G_i\left(x_t, \ldots, x_{0,} x_t^e\right) + \varepsilon_{t,i}$$

where G_i is chosen by each agent, once and for all, in the set of predictors $\mathcal{P}$. For each agent and each available predictor, we generate artificial time series of expectations x_t^i where

$$x_{t+1}^i = G_i\left(x_t, \ldots, x_{0,} x_t^e\right)$$

and x, x^e are taken from the experimental data. Finally we choose the predictor which minimizes the mean square difference between artificial and experimental forecasts[5]. More specifically, for each agent, we select in $\mathcal{P}$ the predictor G_i solving the problem

$$\min_i \sum_{t=1}^{T} \left(x_t^i - x_t^e\right)^2$$

We have considered a set of three simple expectation functions: adaptive expectations with constant gain (AD), $x_{t+1}^e = x_t^e + \alpha\left(x_t - x_t^e\right)$, fading memory learning (FM) and its generalisation previously discussed (GFM)[6]. Both

[4]The same assumption can be found in Branch [6] where a similar test on agents behaviour is done using data from the Survey of Consumer Attitudes and Behavior of the University of Michigan.

[5]For each predictor and each agent we have numerically selected the best parametrisation in terms of implied mean squared error.

[6]Indeed, in a first attempt to see whether and which among common (and simple) learning rules could give good result in replicating agents behaviour we have, at first, checked a wider class of predictors including "Myopic", "Bray", SAC learning à la Hommes and Sorger and various formulation of OLS. As the optimal predictor was never found in this set we have decided to focus on a smaller set of rules.

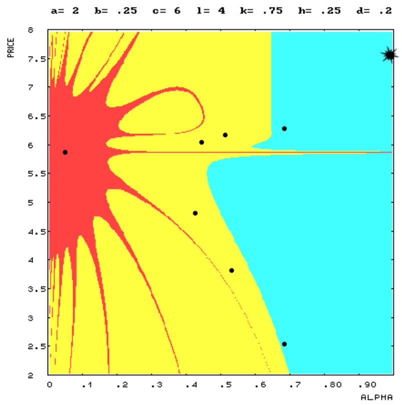

Figure 2: Three attractors and their basins

FM and GFM are used in the recursive formulation of equations (5) and (6) respectively.

The results are summarized in the following table

	AD	FM	GFM
$N.$	29	7	41

On a total of 77 subjects of the experiment GFM works better then the other predictors in more than half of cases. Notice that if we do not consider GFM the situation would be

	AD	FM
$N.$	54	23

so the generalization improves the performance of fading memory which otherwise would be outperformed by the simpler constant gain adaptive learning.

4.2 Simulations

We have run a number of simulations for the model in (12) under GFM learning, that is with expectations defined as in (9). Our aim is to illustrate the phenomenon mentioned in Proposition 4, namely the emergence of multiple attractors. Figure 2 shows the phase space for the couple (α, p): there is a fixed point on the left, a 4-cycle in the middle and a 2-cycle on the right. All these are locally attracting and their basins of attraction are in different colours. It is interesting to see what happens when one perturbes the system depicted in Figure 2 with a stochastic noise. In particular what we did was adding a random disturbance to the expectations feedback map

$$p_t = \frac{a - \tanh \lambda \left(p_t^e - 6\right) - 1}{b} + \varepsilon_t$$

and we chose $\varepsilon_t \sim N\left(0, \sigma^2\right)$ for all t. The simulations are for various (increasing) values of σ^2. Figure 3 shows three typical cases: in all of them the initial condition is the star on the top right of Figure 2. On top the variance is low ($\sigma^2 = 0.05$): after a transient, in which the dynamics visits the neighbourhood of the two periodic orbits, the system converges to the fixed point. In the middle, with a higher variance ($\sigma^2 = 0.3$), the various types of behaviour are recurrent, in the sense that eventually the dynamics escapes from each of the basins. Finally on the bottom graph the variance is even higher ($\sigma^2 = 0.5$) and the various patterns mix up together. Indeed $\sigma^2 = 0.5$ is also the variance used in Hommes et al. [17].

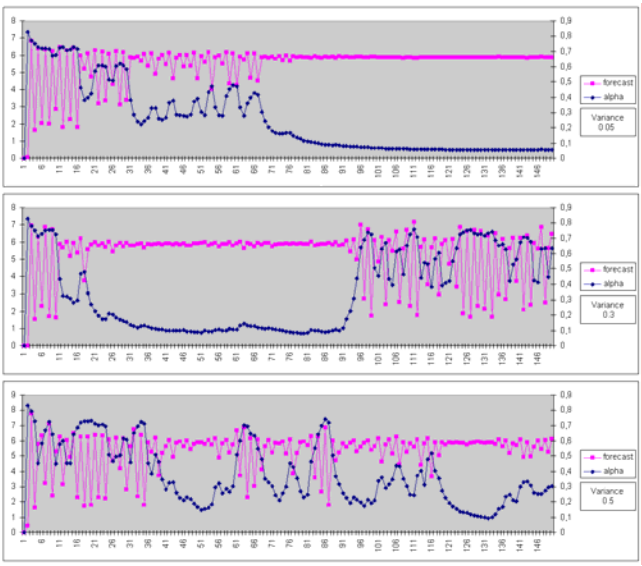

Figure 3: Effects of random disturbance: three typical cases

To show in which sense the graphs in Figure 3 are typical, let us consider Figure 4 that reports a summary of our simulations with noise. We see that with variance up to a certain threshold the system will converge to the fixed point after a sufficient number of iterations; past this threshold though (which can be placed around 0.25-0.30) this phenomenon rapidly desappears as variance is further increased.

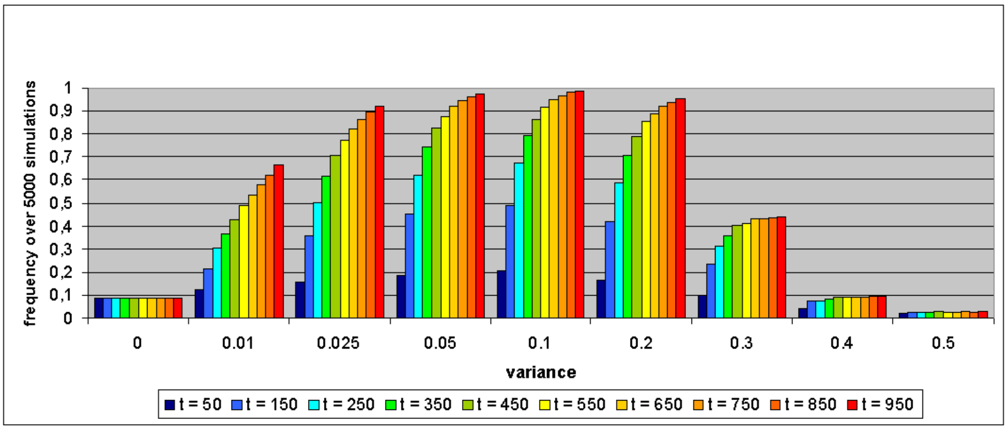

Figure 4: Frequency of convergence at a given time period

5 Appendix

Proof. [of Proposition 1]

The Jacobian of (7) evaluated at (x^*, α^*) is

$$J(x^*, \alpha^*) = \begin{pmatrix} 1 + \alpha^* \left(F'\left(x^*\right) - 1\right) & 0 \\ \frac{\partial G}{\partial x_t^e}(x^*, \alpha^*) & \frac{\partial G}{\partial \alpha_t}(x^*, \alpha^*) \end{pmatrix}$$

Now $\frac{\partial G}{\partial \alpha_t}(x^*, \alpha^*) = 1 - \alpha^* = H\left(0\right)$. The Jacobian is lower triangular so the eigenvalues are $1 + \alpha^*[F'\left(x^*\right) - 1]$ and $H\left(0\right)$. Imposing the usual condition that they be in $(-1, 1)$ returns[7] the claimed result. ■

Proof. [of Proposition 2] Trivially if $x_t^e - x^i = 0$ then $x_{t+1}^e = x_t^e = x^i$. Now, to show that $x^i \neq x^{i+1}$ for all i, observe first that if $x^i = x^{i+1}$ then $x^i = x^j$ for all i, j. On the cycle, which must be of the form $\left(x, \alpha^1\right), \ldots, \left(x, \alpha^p\right)$, the forecast error vanishes and the second equation of (7) reduces to $\alpha_{t+1} = \frac{\alpha_t}{\alpha_t + H(0)}$ which implies monotone convergence to $1 - H\left(0\right)$. Therefore it must be $\alpha^1 = \cdots = \alpha^p$, so we are on a steady state. ■

Proof. [of Proposition 4]

It suffices to have $H\left(0\right)$ large enough to have local stability for x^*. Further, suppose that the sensibility of the function H to the error depends by a multiplicative parameter h

$$\begin{cases} x_{t+1}^e = x_t^e + \alpha_t \left(F(x_t^e) - x_t^e\right) = E\left(x_t^e, \alpha_t\right) \\ \alpha_{t+1} = \frac{\alpha_t}{\alpha_t + H\left(h\left(x_{t+1} - x_{t+1}^e\right)\right)} = G_h\left(x_t^e, \alpha_t\right) \end{cases} \tag{13}$$

[7] Notice that as $H\left(0\right)$ approaches 1 the inequality $-\frac{1+H(0)}{1-H(0)} < F'\left(x^*\right)$ tends to become redundant: in the limit though, because $H\left(0\right) = 1$ implies $\alpha^* = 0$, hyperbolicity is violated.

Given the system (13), assume that $\lim_{|y|\to\infty} H(y) = 0$. Then we prove that there are $\delta, \bar{\alpha}, \bar{h}$ such that, for any initial condition (x_0, α_0) satisfying:

$$\begin{aligned} |x_0 - x^*| &\geq \delta \\ \alpha_0 &\geq \bar{\alpha} \end{aligned}$$

we have that

$$|x_t - x^*| > \delta \text{ and } \alpha_t > \bar{\alpha}$$

for all $t > 0$. This implies there are other attractors beside x^*.

Let a, b be such that $b - x^* = x^* - a$ and $x \in [a, b] \Rightarrow F'(x) < -1$. Clearly $F(a) > b$ and $F(b) < a$. Define $\delta = x^* - a$ and $\hat{F}' = \sup F'(x)$. Choose $\bar{\alpha} \geq \frac{1}{1-\hat{F}'}$ such that

$$\alpha \geq \bar{\alpha} \Rightarrow E(b, \alpha) < a, \ E(a, \alpha) > b. \tag{14}$$

Now it is straightforward to check that $\alpha \geq \bar{\alpha}$ ensures that

$$E_x(x, \alpha) \leq 0. \tag{15}$$

Therefore, using (14) and (15),

$$\left\{ \begin{array}{c} \alpha_t \geq \bar{\alpha} \\ |x_t - x^*| \geq \delta \end{array} \right. \Rightarrow |x_{t+1} - x^*| \geq \delta.$$

Then, let $\bar{h}$ be such that

$$h \geq \bar{h} \Rightarrow H(h\delta) \leq 1 - \bar{\alpha}$$

Finally, noting that $|x^* - x^e| \geq \delta$ implies that $|F(x^e) - x^e| \geq |x^* - x^e|$ we have

$$\left\{ \begin{array}{c} \alpha_t \geq \bar{\alpha} \\ |x_{t+1} - x^*| \geq \delta \end{array} \right. \Rightarrow \alpha_{t+1} = \frac{\alpha_t}{\alpha_t + H\left(h\left(x_{t+1} - x^e_{t+1}\right)\right)} \geq \frac{\alpha_t}{\alpha_t + H(h\delta)} \geq \frac{\bar{\alpha}}{\bar{\alpha} + H(h\delta)} \geq \bar{\alpha}$$

■

Remark 1 *The assumption that* $\lim_{|y|\to\infty} H(y) = 0$ *is in fact rather strong. Indeed our interpretation of this learning mechanism requires that* H *is decreasing with the error but it is not necessary that it tends to zero to the limit. More important, if* $\sup F'(x) = 0$ *we have* $\bar{\alpha} = 1$ *which is very restrictive. In any case, versions of Proposition 4 could be proved under milder* *assumptions. For example if* $\sup F'(x) = 0$ *but* $F(x)$ *is bounded (which is the case for the models of section 4)* $\bar{\alpha}$ *can be chosen to be strictly smaller than* 1*. It is also possible to relax the assumptions on* H *linking it to further properties of the function* F*. For example assuming* $H(F(x_t^e) - x_t^e) \leq \min\{-\bar{\alpha}F'(x_t^e), 1 - \bar{\alpha}\}$ *and* $\frac{\partial H}{\partial |F(x_t^e) - x_t^e|} \leq 0$ *(where the value of* $\bar{\alpha}$ *is determined, and smaller than* 1*, once* F *is known) is sufficient to prove the Proposition for any strictly decreasing* F*. A detailed proof of this is available from the authors.*

References

[1] E. Barucci: Exponentially fading memory learning in forward looking economic models. *Journal of Economic Dynamics and Control*, 24(5-7), 1027-46, (2000).

[2] E. Barucci: Fading memory learning in a class of forward looking models with an application to the hyperinflation dynamics. *Economic modelling.* 18(2), 233-52, (2001).

[3] G. I. Bischi, L. Gardini: Mann iterations reducible to plane endomorphisms. *Quaderni di Economia, Matematica e Statistica*, **36**, Università di Urbino, (1995).

[4] G. I. Bischi, L. Gardini: Basin fractalization due to focal points in a class of triangular maps. *International Journal of Bifurcation and Chaos*, **7**(7), (1997).

[5] G. I. Bischi, A. Naimzada: Global analysis of a nonlinear model with learning. *Economic Notes*, **26**(3), 445-476 (1997).

[6] W. Branch: The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations. *The Economic Journal*, **114**, 592-621 (2004).

[7] M. Bray: Convergence to rational expectations equilibrium. In R. Friedman and E.S. Phelps (eds.), *Individual forecasting and aggregate outcomes*, Cambridge University Press, (1983).

[8] C. Camerer: Individual decision making. In J. Hagel and A. Roth (eds.), *Handbook of experimental economics*, Princeton University Press (1995).

[9] C. Chiarella, X. He, P. Zhu: Fading memory learning in the Cobweb model with risk averse heterogeneous producers. Working Paper - University of Technology Sidney (2003).

[10] D. Colucci, V. Valori: Error learning behaviour and stability revisited. DiMaD Working Paper, 3/01, (2001). Forthcoming on the *Journal of Economic Dynamics and Control*.

[11] S. Figlewski, P. Wachtel: The formation of inflationary expectations. *Review of Economics and Statistics*, **63(1)**, 1-10 (1981).

[12] H. Kelley, D. Friedman: Learning to Forecast Prices. *Economic Inquiry*, **40**, 556-573 (2002).

[13] J. M. Grandmont: Expectations formation and the stability of large socioeconomic systems. *Econometrica*, **66**, 741-781 (1998).

[14] J. D. Hey: Expectations formation: Rational or adaptive or...? *Journal of Economic Behavior and Organization* **25**, 329-349 (1994).

[15] C. H. Hommes: Dynamics of the cobweb model with adaptive expectations and nonlinear supply and demand. *Journal of Economic Behavior and Organization* **24**, 315-335 (1994).

[16] C. H. Hommes, G. Sorger: Consistent Expectations Equilibria. *Macroeconomic Dynamics*, **2(3)** 287-321, (1998).

[17] C. H. Hommes, J. Sonnemans, H. van de Velden: Expectations formation in a Cobweb Economy: Some One Person Experiments. In D. Delli Gatti, M. Gallegati and A. Kirman (eds.) *Interaction and market structure Essays on heterogeneity in economics*. Lecture Notes in Economics and Mathematical Systems - Springer Verlag, 253-266 (2000).

[18] M. Lovell: Tests of the Rational Expectations Hypothesis. *American Economic Review*, **76(1)**, 110-24 (1986).

[19] R. Lucas: Adaptive behaviour and economic theory. *Journal of Business*, **59(4)**, 401-425 (1986).

[20] A. Marcet, J. Nicolini: Recurrent hyperinflations and learning. *American Economic Review*, **93(5)**, 1476-1498 (2003).

[21] K. Pötzelberger, L. Sögner: Stochastic equilibrium: learning by exponential smoothing. *Journal of Economic Dynamics and Control*, **27(10)**, 1743-1770 (2003).

[22] R. Schmalensee: An experimental study of expectation formation. *Econometrica*, **44**, 17-41 (1976).

[23] J. Sonnemans, C. H. Hommes, J. Tuinstra, H. van de Velden: The instability of a heterogeneous cobweb economy: a strategy experiment on expectation formation. *Journal of Economic Behavior and Organization* **54**, 453-481 (2004).

[24] V. Smith, G. Suchanek, A. Williams: Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, **56**, 1119-1151 (1988).

[25] A. Williams: The Formation of Price Forecasts in Experimental Markets. *Journal of Money, Credit and Banking*, **19(1)**, 1-18 (1987).