Learning and Emergent Coordination in Speculative Markets: Some Properties of "Minority Game" Dynamics

G. Bottazzi*
G. Devetag*
G. Dosi *

First version: August 10, 1999; Third draft: December 10, 1999

Abstract

The work studies the properties of "Minority Games" - i.e. a stylized representation of arbitraging in speculative markets - with heterogeneous learning agents. The probabilistic learning model belongs to a well-known class of learning models developed in evolutionary game theory and experimental economics, which have been widely applied to describe human behavior in experimental games.

We test the aggregate properties of this population of agents (i.e., presence of emergent cooperation, asymptotic stability, speed of convergence to equilibrium) as a function of the behavioral rules of the agents and we analyze which properties of the system are sensitive to the precise characteristics of the learning rules and which ones on the contrary can be considered as "generic" features of the game.

Our results indicate that the system dynamics are indeed dependent on the ecology of learning algorithms: in particular, when the degree of "inertia" of the learning rule increases, the market reaches a higher level of allocative and informational efficiency, although on a longer time scale.

*S. Anna School of Advanced Studies, Pisa, Italy. An earlier draft of this work has circulated under the title of "Probabilistic learning and emergent coordination in a non-cooperative game with heterogeneous agents". Financial support from the Italian Ministry of University and Scientific and Technological Research (MURST) and the St. Anna School of Advanced Studies is gratefully acknowledged.

1 Introduction

This work investigates the collective outcomes of different learning procedures by heterogeneous agents who play "minority games"- as such, a highly stylized metaphor of speculative market interactions.

In the original version of this game, first introduced by Challet and Zhang (1997, 1998), a population of N artificial agents must each simultaneously and independently choose between two sides, say 0 and 1. The side chosen by the minority of the agents is the winning one, and agents who choose it are awarded one point each, while those who choose the majority side win nothing. Each agent is initially endowed with a fixed number of strategies (which will be defined more in detail later), and updates them throughout the game according to a deterministic algorithm.

The game, notwithstanding its extreme simplicity, does capture some basic features of speculation in financial markets, whereby agents try to infer some - actual or imagined - structure in the history of collective interactions and, through that, "beat the market" - that is, "beat the majority view" - by arbitraging against it. In the literature on "chartist rules", this roughly corresponds to *negative feedback* or *contrarian strategies*(cf., among others, De Bondt and Thaler 1995).

¿From a formal point of view, the model may be understood as an N-person game with multiple asymmetric equilibria in pure strategies and a unique symmetric mixed-strategy equilibrium. In the following, we assume a large population, whose members cannot communicate with each other - except through the very process of buying and selling -, and we investigate the conditions under which repeated interactions amongst players causes some form of aggregate self-organization to spontaneously emerge.

In its substance, the work that follows links with diverse but complementary streams of economic analysis. First, it straightforwardly relates with that growing literature trying to account for a few "stylized facts" characteristic of financial markets (cf., for example, Brock 1997 and Guillaume et al. 1997) as collective properties of explicit interactive processes among agents who differ at least in terms of "rationally derived beliefs" - such as in Brock and Hommes (1998) -, and possibly also in terms of representations and mental models of the environmental structure in which the agents operate - such as in Marengo and Tordjman (1996) and Arthur et al. (1997) -(see also the remarks in Chiaromonte and Dosi, 1998).

Moreover, the latter models, as well as the work presented here, share a partial or total delinking of trading behaviors from the dynamics of the purported "fundamentals" of the economy whose assets the market under consideration is meant to trade.

Indeed, in order to focus on genuinely speculative behaviors, the implicit assumption here - in tune with preceding Minority Game models and with cognitively richer ones such as Marengo and Tordjman (1996) - is that "fundamentals" remain constant throughout the market history: The whole dynamics is driven by an everlasting sequence of individual attempts of "outsmarting" collective behaviors.

A classic reference in this respect is to Keynes' "Beauty Contest" metaphor, whereby the payoff to an individual player does not stem from the accuracy of the appreciation of the "intrinsic" beauty of various contest candidates but rather from guessing the guesses of the ensemble of the other evaluators..

Strictly speaking, the beauty contest metaphor corresponds to *positive feedback* investment strategies - in the language of technical trading in finance - since the payoff is based on a majority rule [1]

A fulfledged account of speculative phenomena most likely ought to incorporate both "beauty-contest" and "minority game" features - matching, together, herd-like phenomena, bubbles, crashes, etc. Short of that very ambitious task, this work focuses on those aspects of speculation dynamics which essentially involve activities of arbitrage against average market behaviors.

It is an old conjecture from economic theory (cf., for example, Friedman, 1953) that these very activities a) have a stabilizing effect on prices around their "fundamental value"; and b) yield market efficiency by washing away arbitrage opportunities themselves.

Here, the utmost simplicity of our "fundamentals-free" model prevent us from addressing the former issue. However, we do address the latter and ask under what circumstances such a proposition holds. Is it a generic property of minority-type market processes, independently of any further specification of microeconomic decision processes? Or, conversely, do finer details of the "ecology" of agents populations (such as their sheer size), and their decision processes (e.g., the number and sophistication of their strategies, their memory, etc.) affect collective adjustment processes and long term outcomes?

In order to tackle the problem, in the following we shall study the variation of the asymptotic properties and dynamics of a population playing a minority game when the learning rule of the agents is modified, and in particular when agents display varying degrees of "inertia" in adjusting to novel market information. Here we adopt a probabilistic learning algorithm for the agents and leave any other parameter of the original game unmodified (see Cavagna et al. (1999) for a similar investigation and section 4 for a comparison with our results). Note that the chosen probabilistic learning rule is broadly supported by the available evidence on human learning in both experimental and "natural" conditions, notwithstanding significant differences amongst scholars on the precise nature of learning processes (For different views, which however do not contradict the probabilistic representation adopted here, see, among others, Bush and Mosteller 1955, Erev and Roth 1998, Camerer and Ho 1997).

In section 2 we briefly illustrate the basic features of the game, describe the major results presented in the literature on the system asymptotic properties, and introduce the probabilistic learning rules adopted for our simulations.

In section 3 the length of the "training phase" is analyzed. In fact, together with the long run properties of the system, another significant object of inquiry concerns the duration of the adjustment phase. The "transient length" issue bears important theoretical implications, in that it highlights a tradeoff between longer times and better performances, common to probabilistic search

[1] Note incidentally that this process is likely to display significant differences , as compared to Keynes' original metaphor, from the "p-beauty contest" game, first studied in experimental economics by Nagel (1995). In the latter, n players have to each pick a number in the [0,100] interval, and the player who picks the number closest to p times the average of all numbers chosen wins a sum of money. For $p < 1$, the game has a unique Nash equilibrium in which all players pick 0. The solution can be easily found through the application of iterated dominance. Conversely, the original spirit of Keynes' suggestion intuitively hints at multiple dynamics with varying fractions of the population able to guess "where the wind is blowing" and "go with the wind" - as one of the chartist rules is indeed called.

algorithms.

Next, we analyze the system aggregate performance (section 4). In order to do that, we define some complementary measures of efficiency. In particular, we make use of a notion of *allocative efficiency* in tune with Challet and Zhang (1997, 1998), and Manuca et al. (1998), strictly connected to the size of the minority: in fact, the smaller is the winning minority the more points are, so to speak, "left on the table" instead of being distributed over the population. The influence of learning "inertia" on the degree of the system allocative efficiency is analyzed, by comparison with the level of efficiency notionally attainable by perfectly rational and perfectly informed players who "solve" the game analytically.

Section 5 analyzes the effect of learning inertia on the degree of "informational efficiency", connected to the existence of arbitrage opportunities (Appendix A discusses the impact on our results of adding a time decay factor to the learning algorithm).

2 Minority Game: Structure and Learning Algorithm

Let us start by briefly recalling the features of the Minority Game (Challet and Zhang 1997, 1998).

As already mentioned, the Game is played by a collection of N agents. At each time step each agent choose independently between two sides, say 0 and 1. The side chosen by the minority of the agents, i.e. the "minority" side, is the winner, and agents who choose it are awarded one point each, while those who choose the majority side win nothing [2]

The only information available to the players after each round is which side (0 or 1) was the winning one. The market information is represented by the (history) H of the game, that is a binary string specifying which side has won at every stage.

The "rational sophistication" of the agents is determined once for all by two parameters homogeneous over all the population. The first parameter is the amount of "memory"of the past that agents are able to retain, corresponding to the last m bits h_m of the game history H. The second parameter is the number of strategies s assigned to each agent.

A strategy is defined as a prescription on the action to take on the next round of play (i.e. to choose 0 or 1) after a particular history (that is, a particular sequence of m bits) has been observed up to that point. For example, in the case in which $m = 3$, a strategy is defined in Table 1.

The "history" columns specify all the possible histories of the game in the last m periods; the "action" columns specify which action to choose on the next trial in correspondence to each particular history observed. The strategies are randomly drawn from a common pool consisting of the 2^{2^m} ways of assigning all the 2^m possible strings of length m to an action. Note that even if m and s are the same for all the population, heterogeneity follows from the random initial

[2] In literature various modifications of this payoff function have been proposed. (see for instance Challet and Zhang, 1997, 1998 or De Cara et al. (1998)). We stick to the original one for its simplicity and because the essential feature of the model are highly insensitive to the proposed variations.

history	action	history	action
000	1	100	1
001	0	101	0
010	0	110	1
011	1	111	0

Table 1: An example of strategy with $m = 3$.

strategy assignment. Each active strategy will be characterized by a value $q_i(t)$, which indicates the total number of points accumulated by strategy i at time t. Indeed, after each period of the game, all the strategies that have "predicted" correctly in that period (that is, all strategies prescribing the side resulting ex post the winning side) are assigned one point each.

Given strategies and updating rules, in the original game behavior at each stage is completely deterministic, in that each agent at each period plays, among the strategies she possesses, the one with the highest number of accumulated points.

In order to judge the system's performance, it is necessary to introduce a measure of *allocative efficiency*. A natural candidate is provided by the average number of players belonging to the winning party i.e. the average number of points awarded to the whole population at each round. Such quantity measures also the degree to which the system is near a Nash-type notion of equilibrium. In fact, when the winning party is equal to $N/2$ the system finds itself in equilibrium in the sense that no player can do better by unilaterally deviating (note that the size of the winning party is not an information available to the players). That quantity is indeed also a measure of the degree of self-organization of the system.

In accordance to the previous literature (Challet and Zhang 1997, 1998), as a measure of allocative efficiency we compute a quantity associated to the foregoing one, namely the mean squared deviation from the half population σ. Let N be the number of agents and $N_0(t)$ the number of agents attaining side 0 at time step t, then in a given simulation of length T the mean squared deviation is computed as

$$\sigma = \frac{1}{T} \sum_{\tau=0}^{T} (N_0(\tau) - \frac{N}{2})^2 \,. \tag{1}$$

Note that σ is also a measure of the fluctuations around the $N!/(((N-1)/2)!)^2$ game Nash equilibria in which exactly $(N-1)/2$ players form the winning minority.

The dynamic process can be expected to depend on the parameters N, m, and s and also on the initial distribution of strategies among agents and on the initial history, both generated randomly. Therefore, if not stated otherwise, all the quantities shown in the simulations that follow are obtained via an averaging procedure over 50 independent sample paths with randomly generated initial histories and strategy distributions. This averaging procedure is performed in order to produce asymptotically stable quantities, i.e. a different resampling with different initial histories and strategies will produce an equal "asymptotic" state for the system [3].

Moreover at the beginning of each simulation the system is left evolving for a "training phase" of length T_0 in order to wash away any possible transient effect on the subsequent averaging procedure. The quantities so obtained can be considered "asymptotic" properties of the system as long as T_0 and T are chosen high enough as to provide a good approximation of the limit $T \to \infty$.

The dependence of the volatility σ on N, m and s for the original minority game has been studied in many works (Challet and Zhang 1997, 1998, Savit et al. 1998, Manuca et al. 1998) and is summarized in Fig. (1) for $s = 2$.

As noticed by Savit et al. (1998), and Manuca et al. (1998) the type of market regime is determined, at least in first approximation, by the ratio $z = 2^m/N$: hence the curves for various N collapse if plotted in this variable. In this respect notice that even if the actual number of possible strategies is 2^{2^m}, their relative strengths are completely defined in term of the frequency $P(0|h_m)$ with which, over a history, a 0 follows a given $m-$length string h_m. And there are 2^m of such variables. So, z can be interpreted as the density of agents in the strategy space degrees-of-freedom.

As shown in Fig. (1) tree different "regimes" of the system can be identified. First, a "random regime" occurs when z is large (the agent are sparse in the strategy space). The system can hardly organize and its behavior can be described as a collection of random agents that choose their side with a coin toss. In fact suppose the past history to be a given h_m and suppose there are $N_d(h_m)$ agents whose strategies prescribe different actions based on that history while there are $N_0(h_m)$ and $N_1(h_m)$ agents whose strategies prescribe the same party (we restrict ourselves to the $s = 2$ case), respectively 0 and 1. If the agent in N_d choose randomly, the variance is $\sigma(h_m) = N_d(h_m)/4 + (N_0(h_m) - N_1(h_m))^2/4$. The average over the possible h_m will then give $\sigma = N/4$. Notice that σ is shaped by two factors, namely a fluctuation in the choices of agents able to choose and a fluctuation in the initial distribution of strategies.

The second regime could be called the "inefficient regime" for $z << 1$. Here the agents densely populate the strategy space and they in fact "coordinate" in the sense that their actions are strongly correlated. This coordination however leads to a worsening of the overall performance due to a "crowd" effect (Johnson et al. 1998): the agents in fact are too similar to each other and they tend all to choose the same party based on the available information. Loosely speaking, "overcrowding" tends to produce, even in this arbitrage game, "beauty-contest" bandwagon phenomena.

The third regime for $z \sim 1$ is where coordination produces a better-than-random performance. Here the agents are differentiated enough not to yield "crowd effects" but sufficiently distributed over the strategy space not to produce a random-like behavior.

The literature on Minority Game in fact, has focused on the "criticality" of the value z_c where σ is minimum, suggesting that a major change in the system behavior happens when this point is crossed (see Challet and Marsili, 1999).

[3] Below we will restrict our analysis to the case $N = 101$ and $s = 2$ and we will speak of the "optimal" value for memory length m_o referring to the value of m which minimize σ with this parameter choice. Note that the values chosen for m and N are justified on the grounds of the findings of Challet and Zhang (1997, 1998), Savit et al. (1997), and Manuca et al. (1998); the choice to set s=2 is justified by the fact that the system exhibits the same qualitative properties for any $s \geq 2$, while reducing to a trivial case for s=1.

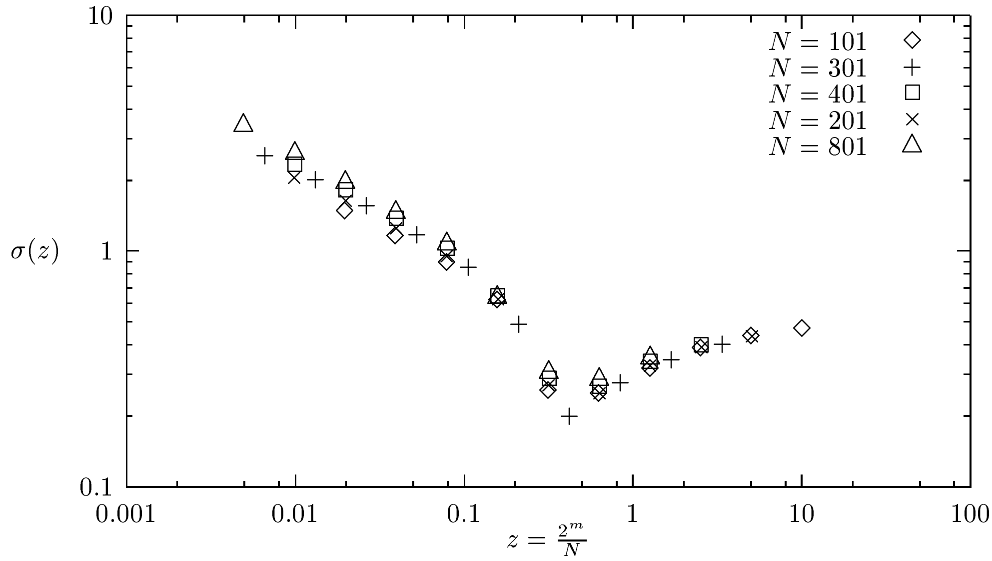

Figure 1: The volatility $\sigma(z)$ for $s = 2$ and different values for N and m

As we will see in the following sections, this "criticality" clearly appears also with respect to changes in the parameter characterizing the learning rules.

In order to introduce them, however, few words on the notion of strategy in this game are required.

In the view suggested here, each strategy may been seen as a particular "mental model" or "hypothesis" about the world (the "world" may include values of the fundamentals of the market in question, or, in the simple Minority Game set-up, the beliefs and behavior of the other players). Each general hypothesis then translates into specific predictions on which will be the winning action for each particular history observed so far. In this respect, a strategy in this game vaguely resembles the notion of "repeated game strategy" in standard game theory (see, e.g., Osborne and Rubinstein 1994, for an introduction).

Each agent initially has a number of different strategies, that is a number of different (and competing) hypotheses. After each period, more evidence is collected and all the hypotheses consistent with the evidence are revised through a process similar in spirit to Bayesian updating.

¿From a cognitive point of view, both the definition of strategy and the choice of the updating rule are particularly demanding in terms of the degree of rationality of the agents. In fact, players are not only supposed to form several hypotheses about the game, but also to consistently perform sophisticated counterfactual reasoning in their updating. For continuity with previous analyses of Minority Games we maintain this set-up, even if we consider it a sort of "upper bound" to the cognitive abilities one ought to realistically attribute to the players.

On the other hand, a rather robust evidence (cf., for example, Camerer and Ho 1997, Erev and Roth, 1998) supports the idea that learning dynamics follow some stochastic processes in the choice across strategies (or actions), driven by experience (either actual experience in simple reinforcement dynamics or also notional experience when allowing for counterfactual reasoning). Hence, in the model below, even if we leave unaltered the updating mechanism as compared to the existing formal literature on Minority Game (that is, all winning strategies are updated regardless of whether they were actually played or not), we assume that the choice between strategies in each period is probabilistic instead of deterministic.

Recall the definition of $q_i(t)$ as the total number of points strategy i would have won if played until time t. Then each agent chooses among her strategies following the probability distribution:

$$p_i(t) = \frac{e^{\beta q_i(t)}}{\sum_j e^{\beta q_j(t)}} \,. \tag{2}$$

where the sum on j is over all the strategies possessed by the player [4]. Note that in general, different players will assign different probabilities to the same strategy due to different strategy endowments.

The model bears similarities with a discrete time replicator dynamics (Weibull 1995). The parameter β can be considered as a sort of "sensitivity of choice to marginal information ": when it is high, the agents are sensitive even to little differences in the notional score of their strategies. In the limit for $\beta \to \infty$ the usual minority game rule is recovered. On the contrary for low values of β a great difference in strategy strengths is necessary in order to obtain significant differences in probabilities.

The connection of (2) with the replicator dynamics is straightforward if one looks at the probability updating equation associated with it:

$$p_i(t+1) = p_i(t) \frac{e^{\beta\, \delta q_i(t)}}{\sum_j p_j(t) e^{\beta\, \delta q_j(t)}} \,. \tag{3}$$

where $\delta q_i(t) = q_i(t+1) - q_i(t)$ are the points won by strategy i at time t. If one thinks of a continuous process $\delta q_i(t) = \dot{q}_i(t)\delta t$, where $\dot{q}_i(t)$ is the instantaneous "fitness" of strategy i, then the continuous time replicator dynamics equation is recovered keeping only the first terms in δt expansion.

3 Transient length

Let us consider the problem of defining the correct values for T_0 and T in (1). The central question is: How long must the system be left evolving before it reaches the asymptotically stable dynamics?

Fig. (2) plots the average σ value based on the original minority game (as in Challet and Zhang 1997, 1998, Savit et al. 1997, Manuca et al. 1998) as a function of the time length T over which this average is taken with a transient $T_0 = T$. As it can be seen from the graph, the values used in the literature on the minority game are generally sufficient to obtain a prediction correct to a few percent. However, two things are worth noticing:

- The system approaches the asymptotic value from above, intuitively suggesting that the system "learns" over time to self-organize

[4] For our $s = 2$ case, the summation will obviously contain two terms

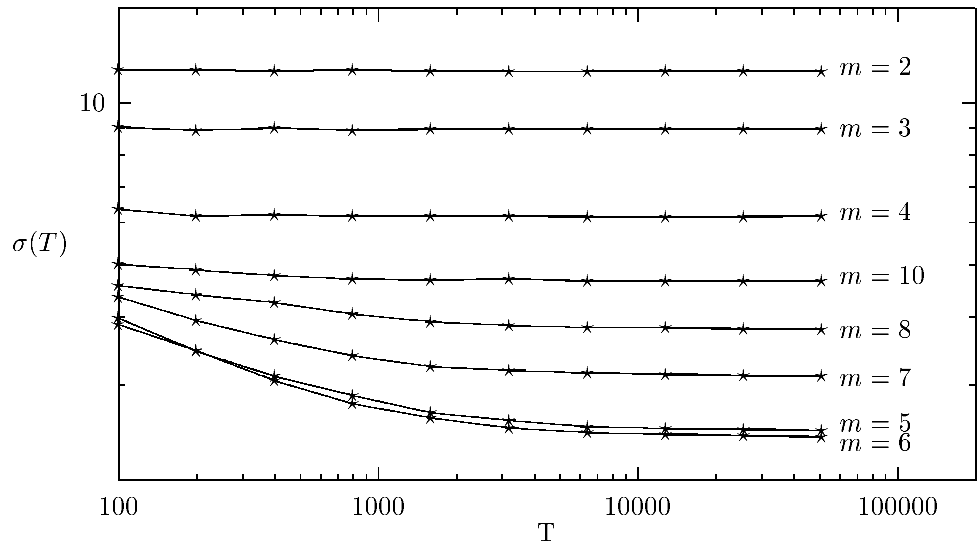

Figure 2: The mean σ as a function of the run length for different m. The points are averages over a sample of 30 independent runs with $N = 101$ and $s = 2$

- For low values of m, in the "inefficient regime", and for high value of m, in the "random regime",the system reaches a stable dynamic quite fast. On the contrary, for values of m near the "optimal" value m_o, the system takes a longer time to self-organize

Consider now the case in which the learning rule is the one described in (2). For high values of β this learning rule approaches the standard one, and accordingly, the transient length is similar to the one found in such cases. However, as β decreases, the length generally increases. The increase is most dramatic for values of m near the optimal value m_o, and it progressively disappears for higher values of m, as can be seen in Fig. (3). The interpretation of such a result stems from the meaning of β in terms of the learning rule. Supposing a non trivial dynamics for m near m_o, the parameter β sets the time scale on which such dynamics is attained.

As an illustration, consider the following example:

Let be $r(t) = p_1(t)/p_2(t)$ the ratio of the probabilities that an agent associates to her two strategies, and $\Delta q(t) = q_1(t) - q_2(t)$ the difference in their respective strengths. From (2) it follows that $r(t) = e^{\beta \Delta q(t)}$. Assuming that the difference in the two strategies performance holds constant over time (an assumption which is generally true in the initial transient regime where agents' behavior is basically random) we obtain $\Delta q(t) \sim t$: hence, from the equality above, a given difference in probability is obtained at a time which is inversely proportional to β.

In order to estimate the time scale over which the system long-run properties are attained, we use the following procedure: Holding all the parameters and the initial conditions constant, the system volatility can be expressed as a function

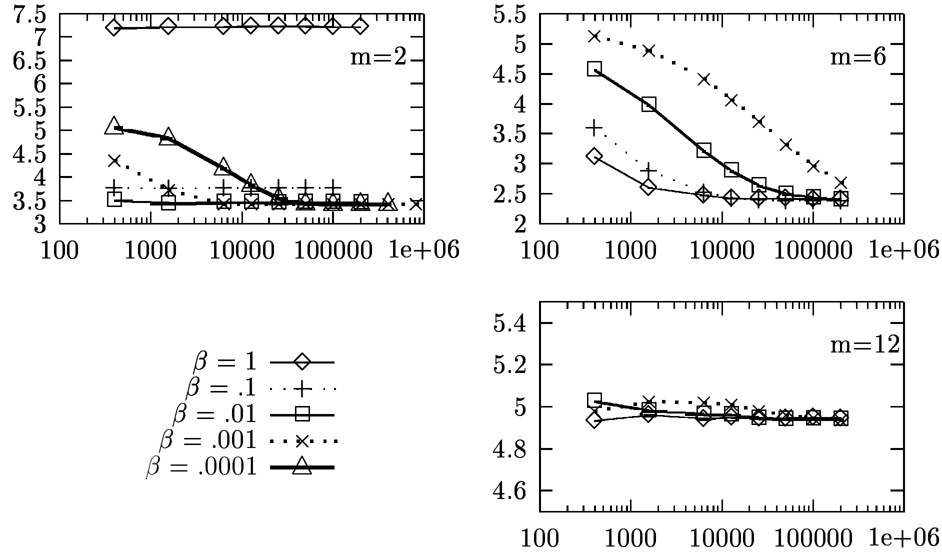

Figure 3: σ as a function of run length T for different β's. The points are average over a 30 runs sample with a transient time $T_0 = T$.

of both the "transient" phase duration, and of the time length over which it is averaged, i.e. $\sigma = \sigma(T, T_0)$.

Starting from a reference time T_r,[5] we compute the mean volatility progressively doubling t and t_0, and thus obtaining a series of values $\sigma_n = \sigma(2^n T_r, 2^n T_r)$.

When the relative variation $|\sigma_n - \sigma_{n-1}|/\sigma_n$ falls below a fixed threshold ϵ, we stop and take the last computed value of σ as an estimate of its asymptotic value. The corresponding time length $\hat{T}(\epsilon)$ will be an estimate of the time implied by the system to reach this asymptotic state.

As can be seen in Fig. (3) the increase in $\hat{T}$ when β is lowered is mainly concentrated around m_o, with shapes that suggest the presence of a "regime" discontinuity.

4 Allocative efficiency

In order to analyze the asymptotic properties of $\sigma(m)$ for different β, we use the procedure just described regarding the calculation of $\hat{T}$, i.e. we leave the system evolve until "stability" is reached[6]. The simulation results are plotted in Fig. (5). Interestingly, when β decreases, the performance level of the system generally increases. Moreover, such increase is larger the lower the value of m,

[5] Note that the chosen value for T_0 is irrelevant as long as it is small compared to the typical time scale.

[6] Let us emphasize that the "stable" state does not imply point convergence to any state but simply long-run stability of the relevant time-averages (e.g., the mean volatility σ) even if the system continues to fluctuate also in its limit state.

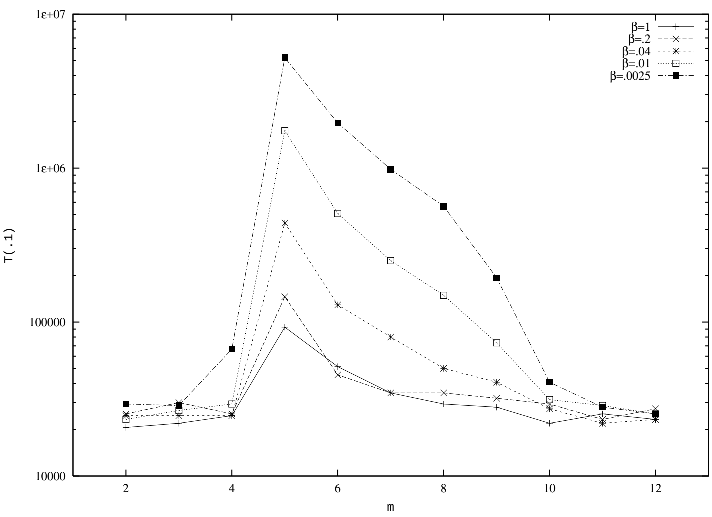

Figure 4: A (rough) estimate of the time $\hat{T}$ needed to reach the stable asymptotic σ value with an error not greater than a given threshold (The plots are made against m for different value of β).

and it becomes negligible for $m \geq m_0$. The observed behavior is consistent with the idea that for high values of m, the system dynamics tends to be determined by the initial distribution of strategies among players, and the players have no opportunities to attain a higher performance by adjusting their behavior. Therefore, the particular learning rule used is largely irrelevant.

Recall that an increasing m means an increasing number of possible strategies over which players may initially draw. For a fixed N, the "ecology" of drawn strategies becomes thinner as compared to the notionally available ones. That phenomenon, it turns out, prevents the system from self-organizing. Conversely, for low values of m, the original learning rule ($\beta = \infty$) produces an "overcrowding effect" (Johnson et al. 1998) (corresponding to large groups of agents choosing the same side) which, due to homogeneity in the initial strategy endowments, prevents the system from attaining a high degree of efficiency.

In some sense, one can interpret the overcrowding effect as a collective form of "overreaction" (Thaler 1993). Of course, introducing a probabilistic learning rule for the strategy choice acts like a brake that dumps the amplitude of such correlated fluctuations. At the individual level, this corresponds to lower β's, i.e. to higher degree of "inertia" as agents update their probabilities more slowly. In other words, as β decreases each agent behaves as if she was applying a sort of stochastic "fictitious play" approximation (cf., for example, Fudenberg and Levine 1998) [7], with an implicit assumption of stationarity on the distribution of other agents' choices. If the whole population shares the same β - as in the present model - the assumption is, in a way, self-fulfilling: a decrease in β makes the behavior of the population as a whole change at a slower pace.

A slower probability updating at the individual level and the resulting more

[7] Note that fictitious play implies that a player best responds to the observed frequency of the opponent's play

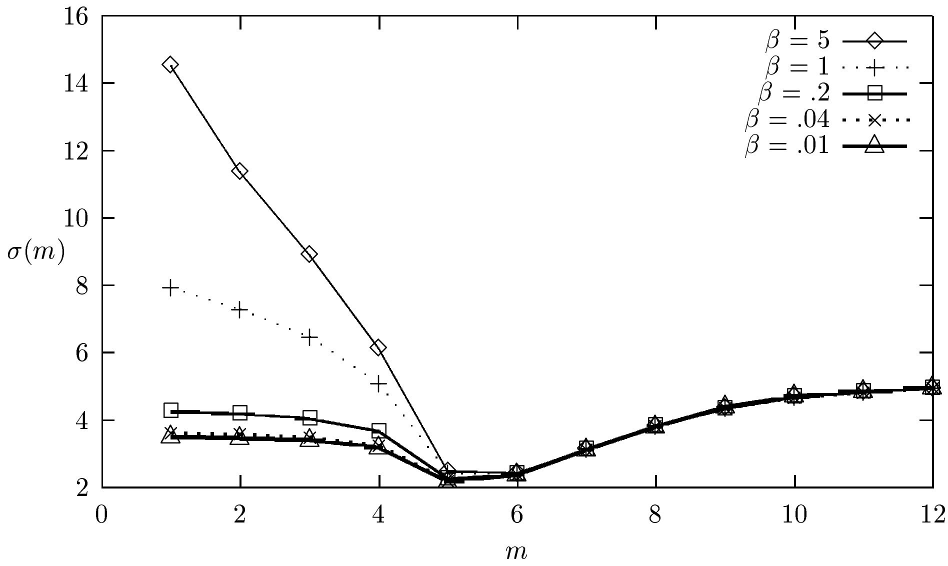

Figure 5: Volatility σ for $s = 2$, $N = 101$ and various m and β. The run are performed doubling the time length T until the last two value are different by less then 1%.

stable collective behavior, together, imply that σ is a non increasing function of β.

In fact, if the system reaches a dynamical stability via an averaging procedure over the past outcomes, increasing the time scale over which averaging occurs cannot rule out previously attainable equilibria. However, note that if one performs the simulations with a fixed time length, when β is small enough, the system behavior resembles the behavior of a random system. This phenomenon is due to both the increase in the transient length and the purely random initial dynamics which occur when β is decreased.

In order to study the asymptotic properties of the system one is facing a double limiting problem: we are interested in the value of volatility in both limits $\beta \to 0$ and $T \to \infty$ and therefore it is necessary to specify which limit is taken first. The results of the fixed time simulations are plotted in Fig. (6) and are in line with Cavagna et al. (1999). As the figure shows, when $\beta \to 0$ the system approaches a collection of randomly choosing agents.

In general, the performance attainable in the minority game via a collective organization of agents with limited information and limited ability to choose is actually surprisingly high, compared to the efficiency attainable with more informed and more "rational" agents endowed with a greater flexibility in choice.

Consider for instance a collection of agents, in line with the original minority game, with the following characteristics: each agent is assigned $S = 2$ strategies, and a vector of length 2^m containing the probability $p(h_m)$ of playing according to the first strategy after the occurrence of h_m. Moreover, for each h_m, each

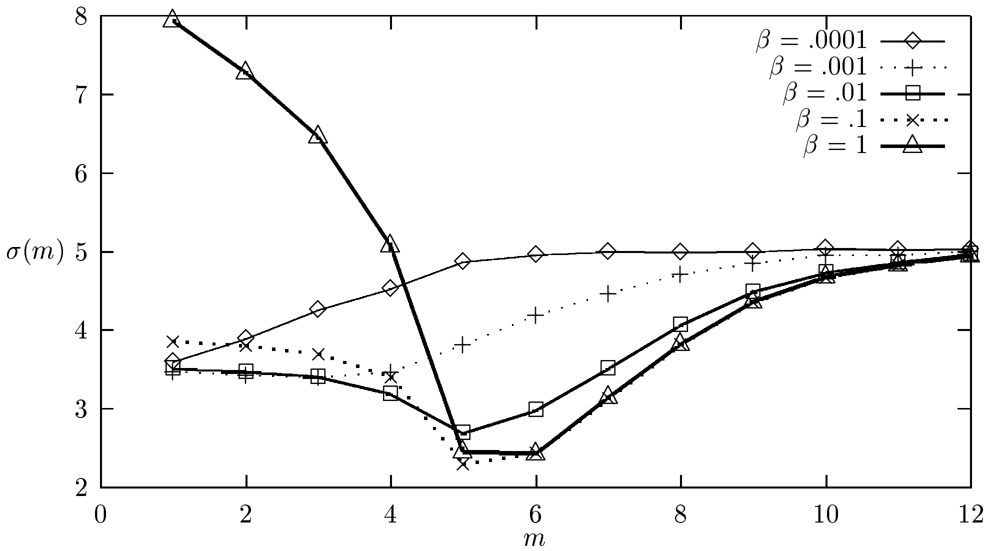

Figure 6: The volatility σ for $s = 2$, $N = 101$ and various m and β. The runs are performed with a fixed time length $T = T_0 = 10000$. When $\beta \to 0$ the system approaches a collection of randomly choosing agents.

agent knows the values of $N_0(h_m)$, $N_1(h_m)$ and $N_d(h_m)$ indicating respectively the number of agents for which their strategies both prescribe to play 0, both to play 1, or to play differently.

Assuming that the game structure and the amount of information available to agents is common knowledge and that the agents are perfectly rational, the problem completely factorizes and, for each h_m, every agent in $N_d(h_m)$ will solve the game analytically choosing $p(h_m)$ in order to minimize

$$\frac{(N_1(h_m) - N_0(h_m))}{2} - p(h_m)N_d(h_m) \tag{4}$$

i.e. making the average fraction of the population choosing a given side as near to $N/2$ as possible. This choice will produce a volatility $\sigma \sim N_d/4 = N/8$ which is roughly similar to what obtained in simulation Fig. (5) in low m low β region [8].

A final remark concerns the variance of the distribution of σ as a function β for various m, as plotted in Fig. (7): when β decreases the variance of σ decreases for any m. However it remains three times greater for $m = m_o$ suggesting a stronger dependence of the asymptotic performance on the initial strategy assignment which the system is not able to wash out. That is, significant *path-dependence effects* are present.

[8]We are assuming $\Delta N = N_1(h_m) - N_0(h_m) < N_d(h_m)$. Notice that for random agents $\Delta N \sim \sqrt(N)$ and $N_d \sim N$ and that one can neglect the $\Delta N/N_d$ terms in the solution of (4)when N is large.

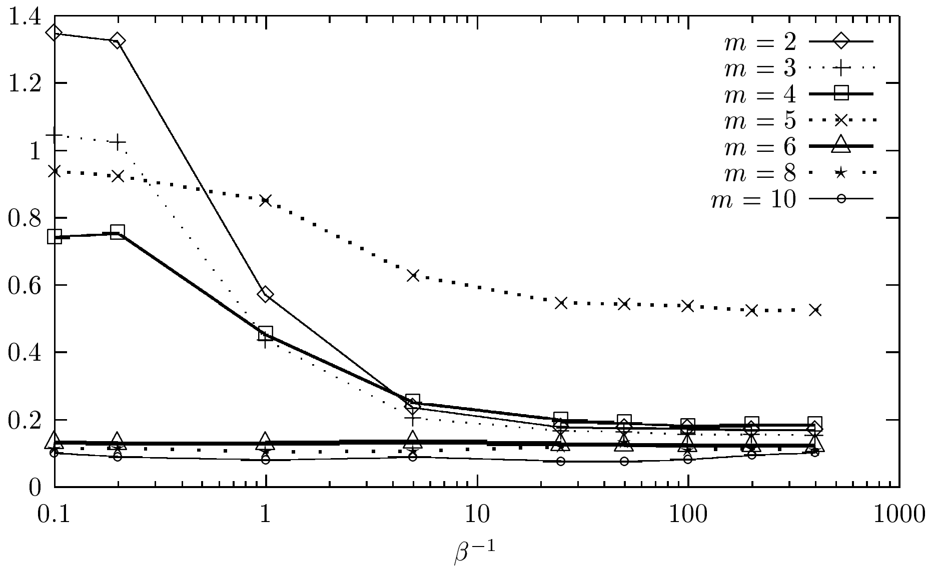

Figure 7: Variance of the distribution of σ over a sample of 50 independent runs. As β becomes small the point $m \sim m_0$ maintains a significantly larger variance.

5 Informational efficiency

What we have been calling allocative efficiency basically highlights the collective ability of capturing the payoffs which the game notionally allows.

A complementary issue regards the *informational efficiency* of the market process, i.e., the extent to which the future system outcomes are unpredictable, or, in other words, the absence of any arbitrage opportunities.

Thus, let us analyze the informational content of H, the binary string of successive winning sides.

Let $p(0|h_l)$ be the probability that a 0 follows a given string h_l of all the possible 2^l strings of length l.

The analysis performed in Manuca et al. (1998) for the original game leads to the identification of two regimes: a "partially efficient" regime for $m < m_o$ in which $p(0|h_l) = .5$, as long as $l \leq m$; thus no informational content is left for strategies with memory less or equal to the one used by the agents. For $m > m_o$ an "inefficient" regime is entered in which the distribution of $p(0|h_l)$ is not flat, even for $l \leq m$, meaning there are "good" strategies that might exploit the market signal to obtain differential profits. For $l > m$ both the regions show a non trivial distribution $p(0|h_l)$ with an increasing degree of "ruggedness" as l increases.

The effect of introducing some degree of behavioral randomness through the parameter β leads to the obvious effect of reducing the "ruggedness" of the distribution of $p(0|h_l)$ (see Fig (8)).

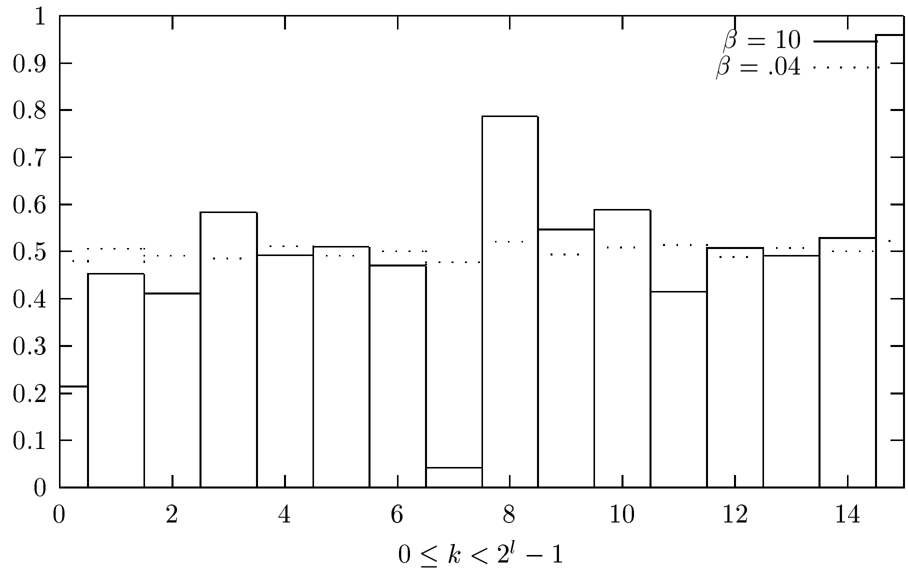

Figure 8: The probability $p(0|h_l)$ of obtaining 0 following a given binary string h_l in system history for $m = 3$ and $l = m + 1 = 4$. When β is reduced the distribution "flatten" and any structure is lost.

In order to study the behavior of the system as β changes we introduce two related quantities which can be used to characterize the informational content of the time series. The first is the conditional entropy $H(l)$ defined as:

$$H(l) = -\sum_{h_l} p(h_l) \sum_{i \in \{0,1\}} p(i|h_l) \log p(i|h_l) \tag{5}$$

where the summation is intended over all the possible string of length l and $p(h_l)$ is the frequency of a given string in the system history H. The maximum value $H(l) = 1$ is reached for a flat distribution $p(0|h_l) = .5$, and it corresponds to the impossibility of forecasting (in probability) the next outcome based on the previous l outcomes. The idea that the information content can be used to "exploit the market" leads to the definition of a second quantity $A(l)$:

$$A(l) = \sum_{h_l} p(h_l) \max \{p(0|h_l), p(1|h_l)\} \tag{6}$$

which is the average fraction of points won by the best strategy of memory l. This is a measure of the reward obtained by the best arbitrageur with memory l (whereby if no arbitrage opportunities are present $A(l)$ is equal to .5.)

Before analyzing the behavior of these quantities when β is varied, let us briefly consider as a sort of ideal benchmark the properties of a population composed of "perfectly rational", perfectly informed, agents with common knowledge of strategy distributions, etc. Not surprisingly, in these circumstances the

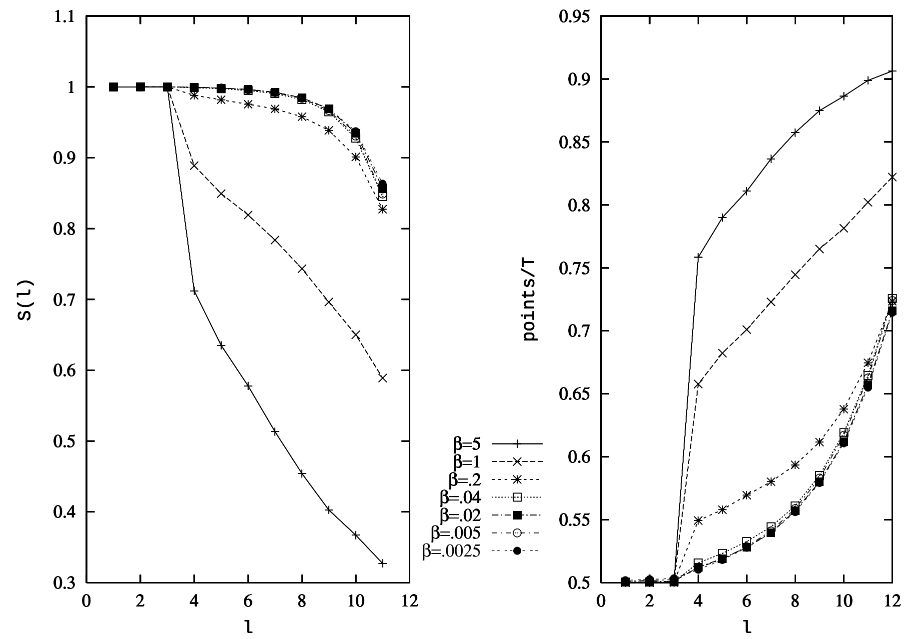

Figure 9: The conditional entropy $S(l)$ (left) and arbitrage opportunity $A(l)$ (right) as a function of time depth l for $m = 3 < m_o$.

problem factorizes for each past history and the dependence on m disappears. The history produced by such a system is a random series of 0 and 1. Indeed the number of agents choosing one side is distributed according to a binomial around $N/2$ with different widths for different h_m. This in particular means that in this extreme case the "memory" loses any predicting power and no arbitrage opportunity is left for agents with longer memory, i.e. no residual information is left in the time series and the behavior of agents makes the market perfectly efficient from an informational point of view. In the last resort, there is nothing to be learned from any history because agents know everything from the start and coordinate their mixed strategies accordingly. Under this assumption we expect $S \sim 1$ and $A \sim .5$.

Short of such an ideal case where the market loses its coordinating role, because agents *ex ante* generate the coordination "in their heads", let us consider, for example, a population of "random agents". Here, due to the unbalance in the initial strategy endowments we expect a non trivial structure to appear for every l; thus $S < 1$ and $A > .5$.

In Fig. (9) we plot $S(l)$ and $A(l)$ for histories generated with a value of $m > m_0$, in the "partially efficient" regime. The effect of decreasing β shows up when $l > m$ but the information content for high l is never completely eliminated. The market becomes less efficient the larger is the time scale l at which it is observed. In fact it can be shown under very general assumptions that certain strings in the history are more abundant than others (Savit et al. 1997) and the long-range correlation that was responsible for the "overcrowding effect" at high β survives as a non trivial structure in $p(0|h_l)$ for high l. All

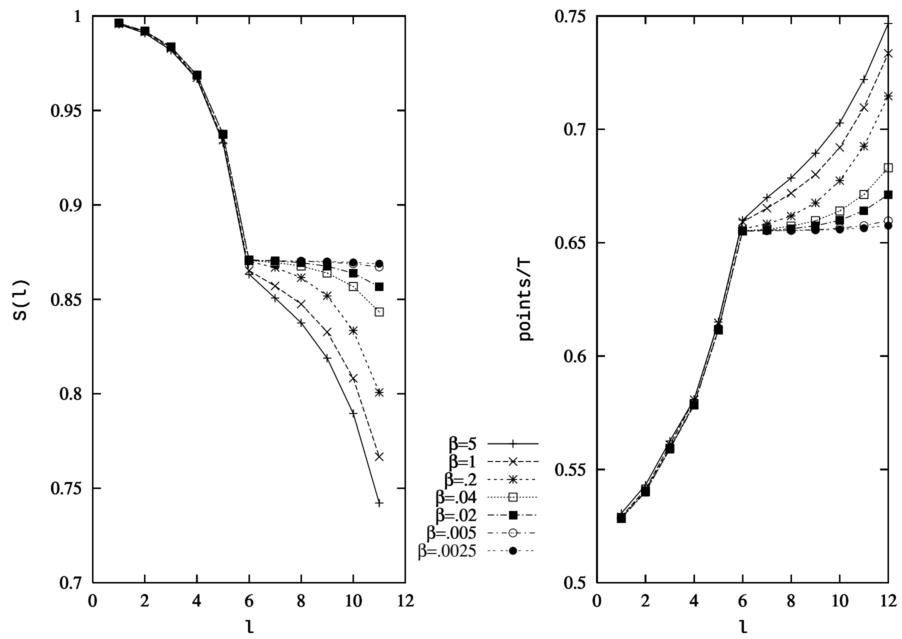

Figure 10: The conditional entropy $S(l)$ (left) and arbitrage opportunity $A(l)$ (right) as a function of time depth l for $m = 6 > m_o$.

this applies despite the fact that to an agent with memory $l \leq m$ the market appears perfectly efficient regardless of the β value.

For values of $m > m_0$ in the "inefficient phase" the effect is in some sense reversed. As can be seen in Fig. (10) the effect of decreasing β is again negligible for $l \leq m$ but in the limit $\beta \to 0$ the curve becomes flat for $l > m$. This last result deserves some comments: the flatness in $l \geq m$ means that no gain is achieved from inspecting the time series with a very long memory $l >> m$ because no more arbitrage opportunities are open for an agent with a longer memory) agent than the best possible agent of memory m. The market can be considered to be, again, "partially efficient" in the sense that it generates an upper bound on the maximal attainable arbitrage capability which does not depend on the arbitrageur memory.

The particular form of the conditional entropy in Fig. (10) suggests that in the limit $\beta \to 0$ the system can be described as a Markov chain of memory m [9]. The result can be explained by noticing that when β is small only great differences in the past performances of strategies are relevant and in the limit $\beta \to 0$ only infinite differences become relevant. Putting it another way, the frequency of victories of the various strategies becomes constant implying the formation of a static hierarchical structure in the strategy space which at the end is responsible of the Markov character of the resulting history. The appearance

[9] Notice that by construction, in the learning rules considered here the past is not discounted, and the agents weigh their strategies on the basis of all the game outcomes starting from the beginning of the simulation (however, see Appendix A for an analysis of the system properties when a time "decaying" factor is introduced).

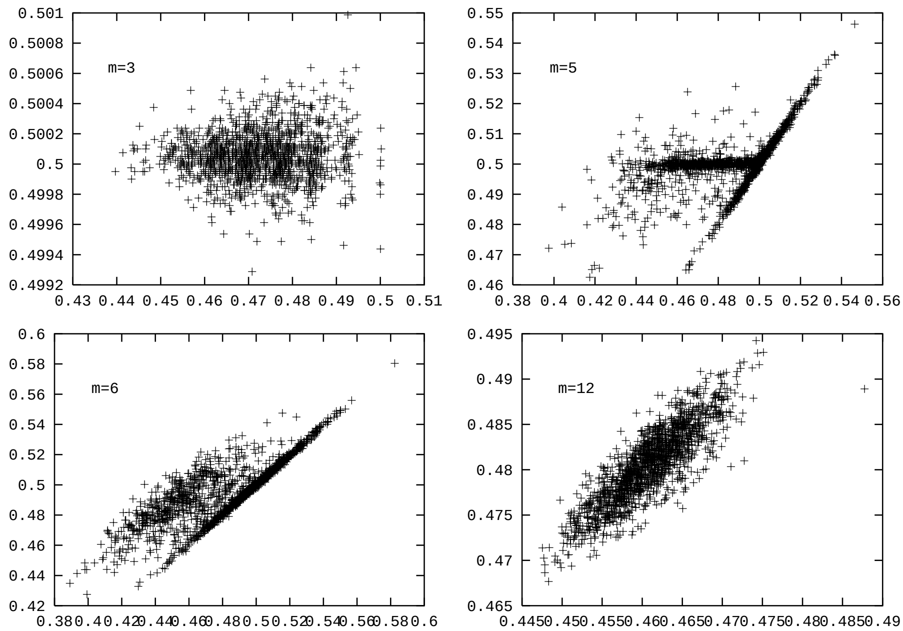

Figure 11: Plot, for each player, of the scoring rate of its best strategy against her own winning rate (population of 101 players over 30 independent runs. $\beta = .04$.

of "best strategies" in $m > m_o$ region is well revealed by the plot of the average score by the best strategy versus the average point scored by the player (see Fig. (11)): a correlation in fact appears between the performance of a player and the performance of its best strategy. Moreover, in the $m \sim m_o$ region a sub-population showing the same kind of high correlation coexists with another group that presents no correlation, composed of agents possessing two equally performing strategies.

Conversely, only in the low m region no strategy ends up being preferable to others and no player is bound to lose due only to his initial strategy endowment.

Notice, however, that equivalence between strategies does not necessarily imply equivalence in agent performances. This is highlighted by the plot of the variances and the supports of the points distribution for different values of β and m in Fig. (12). Interestingly, only for low m and low β does equivalence in strategy performance imply a relatively uniform distribution of points over the population. In the other parameter regions learning does not eliminate performance heterogeneity over the population. Loosely speaking, the market self-organizes over an ecology of "mental models" and players, entailing the long-run coexistence of relative "winners and "suckers".

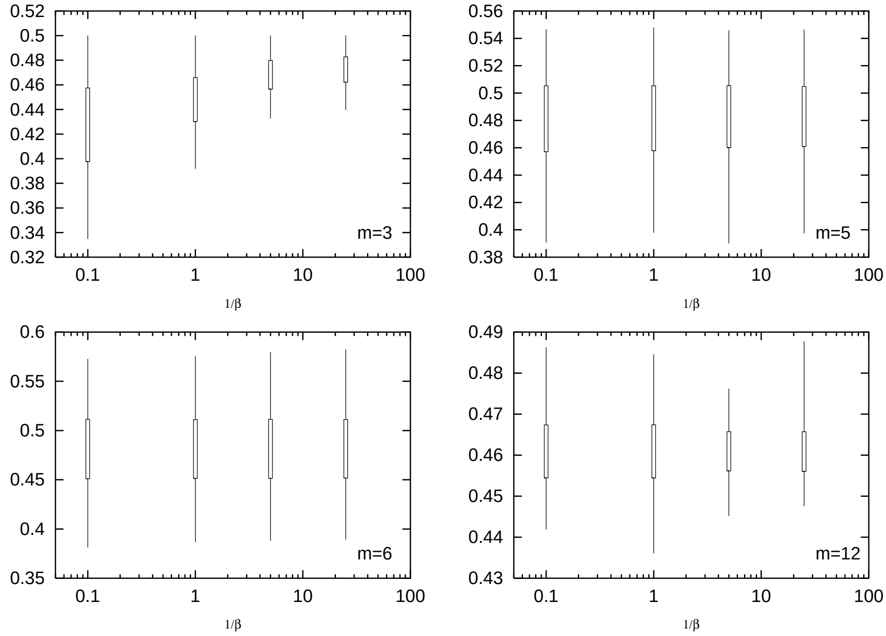

Figure 12: Variance (rectangle) and support (straight line) for the scored points distribution on a population of $N = 101$ with $s = 2$, over 30 independent runs. Notice that while in the high β simulations the distributions are similar in width for any m, when β in reduced the low m region emerges are the "social optimal".

6 Conclusions and Outlook

One of the central questions we addressed in this work is the extent to which market dynamics generated by arbitraging activities, as represented in Minority Games, display "generic" properties, independent from particular behavioral assumptions. Our answer is largely negative: simple variations in the agents' learning algorithms, we show, yield important modifications in the asymptotic properties of the system.

More specifically, we show that *less* sensitivity to marginal information, i.e., *more inertia* in the learning algorithm entails improved long-run collective performances, although at the expense of longer adjustment phases. Together, *performance asymmetries* across agents, as measured by the variance (or analogously) the support of the earnings distribution over the population, fall as inertia in the agents' behavior increases.

In general, some degrees of randomness help in improving allocative and informational efficiencies of the market - as defined above. The major effect of randomness is that it performs like a brake on the system dynamics, thus preventing groups of players who densely populate the strategy space from acting in a strongly correlated way and thus from producing "overcrowding" effects which worsen the system performance.

The introduction of randomness in individual behavior is indeed only one of possible ways to maintain behavioral heterogeneity in the population. For

	low z	high z
low β	high A.E., high I.E.	partial A.E., partial I.E.
high β	low A.E., high I.E.	partial A.E., low I.E.

Table 2: System properties: a summary (A.E. and I.E. stand for Allocative and Informational efficiency, respectively.)

instance, the same effect has been obtained in De Cara et al. (1999) substituting the "global" evaluation of strategies on the system history H with a "personal" evaluation in which each agent uses the binary string made up of her own record of victories. A "diversification" mechanism is again at work breaking the correlation among agents.

Table 2 summarizes the different system properties - i.e. different market "regimes" - conditional on different microeconomic "learning regimes".

Indeed, one of our major conclusions, which refines over already existing results in the Minority Game literature, is that market efficiency - in the complementary definitions proposed above - is only achieved under particular ecologies of "mental models" and learning rules.

The general sensitivity of system dynamics upon particular learning algorithms also indicates a natural way forward, beyond the exercises presented in this work, experimenting with cognitively less demanding learning rules. So, for example, it would be interesting to explore the properties of pure reinforcement learning, based on the prescription of "update only the strategy you play". That would also set a sort of "zero-level" model in terms of degrees of required information and cognitive abilities - somewhat at the extreme opposite to the models studied so far in the Minority Game literature. And, somewhere in between, one might explore models with counterfactual updating but without any sophisticated "if...then" structure such as that from Easley and Rustichini (1995).

Moreover, beyond the strict set-up of Minority Games so far, the impact of phenomena of social imitation still awaits to be studied. [10]. And, more generally, the robustness of the foregoing conclusions ought to be checked in less "reduced form", institutionally richer, models, such as "artificial markets" of the *genre* outlined in Marengo and Tordjman (1996), Arthur et al. (1997) and Chiaromonte and Dosi (1998)(see also Kirman 1999).

Finally, a complementary line of inquiry - still largely neglected - regards the analysis of behaviors and learning of human subjects under experimental settings isomorphic to the market interactions formalized above.

In the last resort, all these latter exercises, together with the results presented here ought to be considered as adding some pieces of evidence to the much broader effort aimed at identifying the variables which determine a "universality class", if any, of speculative market processes, as distinguished from those characteristics which strictly depend upon specific distributions of "mental models" and learning rules.

[10] A germane model of financial dynamics with stochastic choice and social imitation is Kirman (1993).

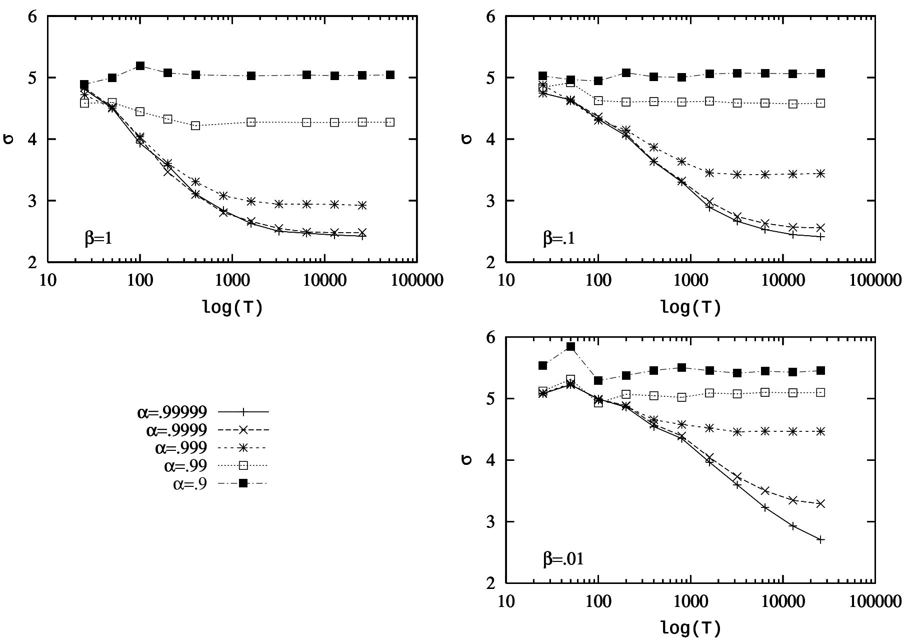

Figure 13: σ as a function of run length T for different values of β and α. The simulations are performed with $m = 6$ where a greater sensitivity of the transient time length toward "learning" parameter β and α is expected, see Sec. (3).

7 Appendix A

Many authors especially in the experimental literature (e.g. Erev and Roth, 1998) introduce one more parameter in the description of learning processes, connected to the idea that agent weigh more the information they received in the recent past as compared to the one coming from far back in past. This parameter takes typically the form of a decay factor. If $\epsilon_i(t)$ are the points scored by strategy i at time t and α $(0 < \alpha \leq 1)$ the information decay factor, the updating rule for the total strength becomes

$$q_i(t+1) = \alpha q_i(t) + \epsilon_i(t) \tag{7}$$

and the associated updating rule for the probabilities:

$$p_i(t+1) = p_i^\alpha(t) \frac{e^{\beta q_i(t)}}{\sum_j p_j^\alpha(t) e^{\beta q_j(t)}} \,. \tag{8}$$

The effect of introducing such a "memory leakage" is twofold: First, it puts an upper limit to the maximal strength any strategy could reach, namely $1/(1-\alpha)$. Second, in presence of no information flux, the equiprobability between strategies is steadily restored. This effect implies that if one takes the limit $\beta \to 0$ keeping the value of α constant, the system will converge to a collection of random agents. In turn, this implies that agents, loosely speaking, have to collect a larger amount of information before they start behaving as a self-organized system.

The effect of introducing "forgetting" in the learning rule is easily understood: if the agents forget more rapidly than they learn they are always bounded to less efficient behavior. Indeed ,as can be seen from Fig. (13), if the value of α is decreased the efficiency of the system is proportionally reduced.

References

[1] Arthur, W.B., J.H. Holland, B. Le Baron, R. Palmer and P. Taylor, (1997), "Asset Pricing under Endogenous Expectations in an Artificial Stock Market", in Arthur, Durlauf and Lane (1997).

[2] Arthur, W.B., S.N. Durlauf, and D. Lane (Eds.), (1997), *The Economy as an Evolving Complex System II*, Reading, Mass., Adison Wesley.

[3] Brock, W.A. and C.H. Hommes, (1998), "Heterogeneous beliefs and routes to chaos in a simple asset pricing model", *Journal of Economic Dynamics and Control*, vol. 22, 1235-1274.

[4] Bush, R. and F. Mosteller, (1955), *Stochastic Models for Learning*, New York, Wiley.

[5] Camerer, C. and T. Ho, (1997), "Experience Weighted Attraction Learning in Games: A Unifying Approach", Social Science Working Paper 1003, California Institute of Technology (forthcoming *Econometrica*, 1999).

[6] Cavagna, A., J. P. Garrahan, I. Giardina and D. Sherrington, (1999), "A thermal model for adaptive competition in a market", preprint cond-mat/9903415.

[7] Challet, D. and M. Marsili, (1999), "Phase Transition and Symmetry Breaking in the Minority Game", preprint cond-mat/9904071.

[8] Challet, D. and Y.-C. Zhang, (1997), "Emergence of cooperation and organization in an evolutionary game", *Physica A* 226, 407.

Challet, D. and Y.-C. Zhang, (1998), "On the minority Game: Analytical and Numerical Studies", *Physica A* 256,514.

[9] Chiaromonte, F. and G. Dosi, (1998), "Modeling a Decentralized Asset Market: An Introduction to the Financial "Toy Room", IIASA, Laxenburg, Austria, Interim Report (Available also as Working Paper, Pisa, St. Anna School of Advanced Studies, 1999).

[10] De Bondt, W.F.M. and R.H. Thaler, (1995), "Financial Decision-Making in Markets and Firms: A Behavioral Perspective", in R. Jarrow et al (eds.), *Handbooks in OR & MS*, Amsterdam, Elsevier Science.

[11] de Cara, M.A.R. and O. Pla and F. Guinea, (1998), "Competition, efficiency and collective behaviour in the "El Farol" bar model", preprint cond-mat/9811162v2.

[12] de Cara, M.A.R. and O. Pla and F. Guinea, (1999), "Learning, competition and cooperation in simple game", preprint cond-mat/9904187.

[13] Easley, D. and A. Rustichini, (1995), "Choice without beliefs", Louvain-la-Neuve, University of Louvaine, CORE, Working Paper.

[14] Erev, I. and A. Roth, (1998), "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria", *American Economic Review*, 88, 848-881.

[15] Friedman, M., (1953), *Essays in Positive Economics*, Chicago, The University of Chicago Press.

[16] Fudenberg, D. and D. K. Levine, (1998), *The Theory of Learning in Games*, Cambridge, Mass., The MIT Press.

[17] Guillaume, D.M., M.M. Dacarogna, R.R. Dav, V.A. Muller, R.B. Olsen and O.V. Pictet, (1997), "From the bird's eye to the mycroscope: A survey of new stylized facts of the intra-daily foreign exchange markets", *Finance and Stochastics*, vol.1, 95-129.

[18] Johnson, N.F. and M. Hart and P. M. Hui, "Crowd effects and volatility in a competitive market", cond-mat/9811227.

[19] Kirman, A., (1993), "Ants, Rationality and Recruitement", *Quarterly Journal of Economics*, 108, 137-156.

[20] Kirman, A., (1999), "Aggregate Activity and Economic Organization", *Revue Europeenne des Sciences Sociales*, 38, 189-230.

[21] Manuca, R. and Yi Li and R. Riolo and R. Savit, (1998), "The structure of Adaptive Competition in Minority Games", adap-org/9811005.

[22] Manuca, R., Yi Li, R.Riolo and R.Savit, (1998), "The structure of Adaptive Competition in Minority Game" adap-org/9811005.

[23] Marengo, L. and H. Tordjman, (1996), "Speculation, Heterogeneity and Learning: A Simulation Model of Exchange Rate Dynamics", *Kyklos*, vol. 49, 407-437.

[24] Nagel, R., (1995), "Unraveling in Guessing Games: An Experimental Study", *American Economic Review*, 1313-1326.

[25] Osborne, M.J. and A. Rubinstein, (1994), *A Course in Game Theory*, Cambridge, Mass., The MIT Press.

[26] Savit, R., R. Manuca and R Riolo, (1997), "Adaptive Competition, Market Efficiency, Phase Transition and Spin Glasses" adap-org/9712006.

[27] Thaler, R.H., (1993), *Advances in Behavioral Finance*, New York, Russel Sage Foundation.

[28] Weibull, J.W., (1995), *Evolutionary Game Theory*, Cambridge, Mass., The MIT Press.