Coordination and self-organization in social systems: experiments and learning models

V. Semeshenko[1], A. Garapin[2], B. Ruffieux[2], and M. B. Gordon[1]

[1] TIMC-IMAG, University of Grenoble I, Domaine de La Merci - Jean Roget, 38706 La Tronche, France
viktoriya.semeshenko@imag.fr, mirta.gordon@imag.fr
[2] GAEL-INRA, University of Grenoble II, BP 47, 38040 Grenoble Cedex 9, France
alexis.garapin@grenoble.inra.fr, bernard.ruffieux@ensgi.inpg.fr

Individual decisions in social situations are frequently influenced by the choices of other individuals. These social interactions lead to coordination problems. In this paper we present experimental results on a stylized coordination game with heterogeneous individuals, obtained under different information conditions. The empirical results are interpreted using reinforcement learning models.

1 Introduction

Individual decisions or behaviors in social situations are frequently influenced by that of other people. The collective outcomes of these interactions may be intuitively inconsistent with the intentions of the individuals who generate them. We consider here the "dying seminar" paradigm of Schelling [5], a situation that illustrates well such behaviors.

Schelling describes a stylized situation in institutions where collective activities are driven by personal motivations. In his model, a group of persons that are eager to meet regularly to discuss some subject of common interest organize a weekly seminar. Each potential participant has a private preference to attend a seminar. Even if the first meeting has a good attendance, it arrives quite often that the number of persons in successive meetings drops down, i.e. the seminar "dies". Nevertheless, most of the participants regret the issue, claiming that they would have attended if enough others had attended regularly. This is a typical situations, that arises in many everyday life circumstances, where social or economic activities are considered worth only if enough individuals participate to them. The outcomes on the long run are collective states that emerge as a consequence of the individual choices. In game theory these situations are represented through coordination games.

Social interactions lead to multiple collective states (equilibria). In the dying seminar example there may be multiple equilibria: one with a low attendance, reflecting the failure of the seminar and another with a high attendance, resulting from a rising dynamics and reflecting a successful seminar. Multiple equilibria bring on coordination dilemmas to the individuals: for a given set of payoffs, individuals may fail to take an action that would be in their collective interest, because they fear that others will not do so. This results in coordination failures. When individuals have heterogeneous preferences, information available to them may help to coordinate their decisions.

In this paper we present experimental results obtained on a dying seminar coordination game played under controlled conditions in the laboratory. The purpose of these experiments was to explore the influence of the information available to the players on the outcome of this game. We propose learning models that allow to interpret the observed behaviors. The paper is organized as follows. We introduce the model in section 2. In section 3, we shortly present the experimental design. In section 4 we report the obtained results. In section 5, we describe the learning models and report the simulations results.

2 Theoretical model with heterogeneous individuals

We consider a group of N potential participants (agents) to the seminar. Following Schelling, we assume that each agent i $(1 \leq i \leq N)$ has a private threshold, which is the minimal number of participants H_i including himself above which he finds it worth to attend the seminar. If agent i attends the seminar and the actual number of participants P turns out to be larger or equal (smaller) to H_i, his payoff is large (small). If he does not participate, he has an intermediate payoff. In the experiments, payoffs are translated into corresponding monetary earnings.

We assume that the thresholds H_i are idiosyncratic characteristics of the agents. They may take any integer value between 0 and N (a threshold N simply means that the individual wants to participate if everyone else participates; a threshold of 0 or 1 means that the corresponding individual is willing to participate unconditionally to the number of participants).

If agent i decides to attend the seminar, his decision is denoted $\omega_i = 1$; otherwise $\omega_i = 0$. Then, the bare sum of these decisions gives the total number of participants,

$$P = \sum_{i=1}^{N} \omega_i. \tag{1}$$

The payoff of an agent i depends on his decision through

$$\pi_i = \begin{cases} U \text{ if } P \geq H_i \quad \text{and} \quad \omega_i = 1 \\ V \text{ if } P < H_i \quad \text{and} \quad \omega_i = 1 \\ W \text{ if } \omega_i = 0 \end{cases} \tag{2}$$

where $U > W > V$, and W is the initial endowment.

Before considering how individuals make their decisions, we analyze the possible equilibria of this model assuming a given distribution of the individual thresholds H_i. To this end we do not need to know who has each threshold, but only the proportion of individuals that have each threshold value, i.e. the probability density function (pdf) $f(H)$. The corresponding cumulative distribution, $F(H)$, represents the fraction of agents that have a threshold H or smaller.

Infinite population limit. We first give a short outline of the equilibrium properties for the case of a very large population, analyzed by Schelling [5]. More precisely, we consider the limit $N \to \infty$ and follow the same lines as in [4]. In this limit, the problem must be formulated in terms of *fractions* of individuals in the population. We introduce thus $h_i \equiv H_i/N$ for the reduced thresholds, which represent the *fraction* of participants below which i is not willing to participate. Then, $f(h)$ is the pdf of the reduced thresholds. The corresponding cumulative distribution is $F(h)$. The actual fraction of individuals that participate to the seminar is denoted η, with $\eta = P/N$. Taking eq. (1) into account, the fraction of participants is

$$\eta = \frac{1}{N} \sum_{i=1}^{N} \omega_i. \tag{3}$$

At equilibrium, assuming perfectly rational players with complete information, the probability that individual i participates is:

$$\mathcal{P}(\omega_i = 1) = f(h_i \leq \eta) = F(\eta), \tag{4}$$

so that we expect a participation

$$\langle \eta \rangle = \frac{1}{N} \sum_{i=1}^{N} \mathcal{P}(\omega_i = 1), \tag{5}$$

which is nothing but the average decision. Since by the law of large numbers, in the $N \to \infty$ limit the variance of averages vanishes (averages are equal to mean values), $\eta = \langle \eta \rangle$. Taking (4) into account we obtain:

$$\eta = \int_{-\infty}^{\eta} f(h) dh = F(\eta). \tag{6}$$

The expected fraction of participants at equilibrium is the fraction of individuals whose reduced threshold is smaller than η. This non-linear equation (6) may have multiple solutions, depending on the (pdf) $f(h)$. They correspond to stable (Nash) or unstable (critical mass) equilibria. A Nash equilibrium is a state of the system where no one can improve his payoff by changing his strategy unilaterally. Thus, it is expected that once in such equilibrium, the system will remain unchanged for ever. On the contrary, in an unstable equilibrium, an individual (generally risky) strategic change may give raise to a chain reaction where others' improve their payoffs through changing their strategies, thus driving the system to one of the possible Nash equilibria. The actually reached equilibrium depends on the shape of the distribution and on the initial decisions. The efficient coordination equilibrium, called Pareto optimal equilibrium in game theory, is the outcome with the highest attendance because, since the individual payoffs of winners are all identical, optimality corresponds to having a majority of winners.

Finite populations. The infinite population limit is pertinent for the analysis of very large systems. The fact that actual systems are not infinite may be taken into account through perturbation theory, which allows to calculate finite size corrections. Since these are of order $O(1/\sqrt{N})$ they may be neglected for N large enough. However, in usual experimental settings, where the number of subjects is seldom larger than some tenths, finite size corrections are too large. In this case we have to analyze the finite systems in each specific situation.

To this end we come back to the original variables N, H_i and P. Here we have to take into account the fact that the pdf $f(H)$ has a discrete support. Correspondingly, the cumulative function $F(H)$ is a monotonic non-decreasing function only defined at integer values of H (see for example figures 1).

Since $F(H)$ is the *fraction* of agents with threshold H, the corresponding *number* of agents is $NF(H)$. Thus, the equilibria are now the solutions to:

$$P = NF(P) = N \sum_{H=H_{min}}^{P} f(H), \tag{7}$$

where H_{min} is the smallest among the agents' thresholds. The intersections of $y = NF(H)$ with the diagonal $y = H$ at integer values of H correspond to the situations that conform the P participants whose thresholds satisfy $H_i \leq P$ (see figures 1). Among these, the Nash equilibria (stable equilibria) are those corresponding to slopes smaller than 1. Equilibria where the slope of $F(H)$ is larger than 1 are unstable: small changes in the attendance will drive the system to one of the neighbouring stable points. In contrast with the case of continuous pdfs ($N \to \infty$), the critical masses are not necessarily unstable equilibria. They may correspond to non-equilibrium configurations. As is explained in the description of the experimental seminars (see section 3.1), these non-equilibrium critical masses may arise when several agents have the same threshold value.

If several agents share the same value of H, pure coordination problems within the sub-group may hinder the convergence to the system's Nash equilibrium, as is discussed in section 3.1. This problem does not exist for generic distributions $f(h)$ in the $N \to \infty$ limit, because the probability of having finite fractions of agents with the same (reduced) threshold h_0 vanishes whenever $f(h)$ is a continuous function. In other words, the fraction of individuals in the population with the same threshold has zero measure. In finite size population this hypothesis is equivalent to assuming that each individual has a different threshold. In the case of infinite populations, when a given value h_0 is shared among a finite fraction of agents, say ϕ, it has to be represented by a Dirac-delta distribution of weight ϕ centered at h_0, i.e. $\phi\delta(h - h_0)$, and has to be carefully handled when solving the equilibrium equation (6).

3 Experimental design

In our settings, $N = 16$ participants are asked to play four different simultaneous games. The subjects are given individual thresholds at the beginning of the experiment, which are in principle different for the 4 different seminars and remain constant during the experiment. These thresholds are drawn from four different distributions $f(H)$, which correspond to different game profiles. In each session, with the same subjects, the 4 games are played for $T = 15$ periods. Earnings of the different periods are cumulated, which in principle should act as an incentive for early convergence to the efficient equilibrium.

3.1 Four seminars

The thresholds distributions $Nf(H)$ of the seminars are presented on figures 1. The corresponding cumulative distributions $NF(H)$ represent the number of agents that have a threshold H or smaller. In our experiments, the distribution $f(H)$ has a central maximum at intermediate values of H in seminars SW1 and SW2. Seminars SM1 and SM2 have a bimodal distribution. As a consequence, there exist Nash equilibria with intermediate attendances in seminars SM1 and SM2, which require coordination of less individuals than in seminars SW1 and SW2. In all four seminars the Pareto optimal Nash equilibrium corresponds to a full attendance, and requires coordination among all the participants. The Nash equilibria with less attendances are called Pareto dominated equilibria. In more details, the main characteristics of the four seminars are (see figure 1):

SW1. There are three stable Nash equilibria, with $P = 3$, $P = 14$, and $P = 16$, the latter being the Pareto-optimal one. $P = 0$ is not an equilibrium, because agents with thresholds $H = 0$ and $H = 1$ have a strictly dominant (maximizing their payoffs) and non-strategic (independent of others' decisions) choice: $\omega = 1$. The outcome $P = 3$ is a risk-dominant equilibrium: it is the rational outcome where only players with non-strategic thresholds decide to participate to the seminar. Equilibrium with $P = 16$ Pareto dominates that with $P = 14$: in the former everyone gets the highest payoff. The critical mass allowing to overcome the attraction toward the risk-dominant equilibrium is $P = 8$. Notice however that there is no intersection between $y = NF(H)$ and $y = H$ at $H = 8$. An attendance of $P = 8$ subjects would require participation of only two of the

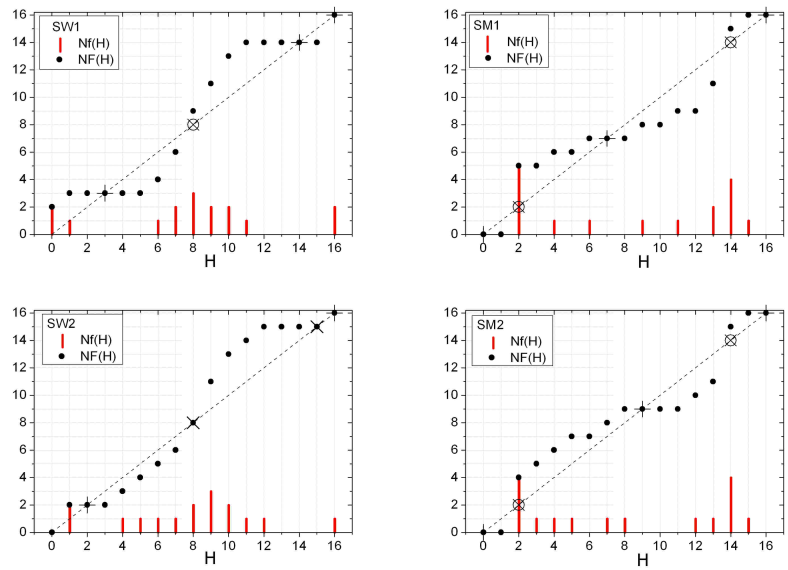

Fig. 1. Stable (Nash) equilibria (+) of the experimental seminars. Critical masses that are unstable equilibria (×) and those that are not equilibria (⊗) are also represented.

three subjects having a threshold $H = 8$. This situation, where the critical mass is *not* an (unstable) equilibrium cannot arise in infinite populations with smooth pdfs $f(h)$. Besides the coordination necessary to reach the Pareto optimal equilibrium in this game, (pure) coordination problems may arise within the subpopulations that share the same threshold value. Lack of coordination (i.e. only one of the individuals with $H = 8$ attends) may hinder the system to reach the equilibrium with higher attendance. If two individuals participate, then the third one does not afford any risk upon participating, and allows the system to overcome the critical mass and reach the $P = 14$ Nash equilibrium.

SW2. There are two stable Nash equilibria: the risk-dominant one with $P = 2$ and the Pareto optimal equilibrium with $P = 16$. Notice that $P = 15$ is an unstable equilibrium: if there are $P = 15$ participants already, then the individual with $H_i = 16$ can improve his payoff through participation without altering the others' payoffs, inducing the stable equilibrium $P = 16$. The critical masses here are $P = 8$ and $P = 15$. $P = 1$ is not a critical mass: decisions of individuals with $H_i = 1$ are non-strategic, and their participation drives the system to the Nash equilibrium with $P = 2$.

SM1. There are three stable Nash equilibria: a risk-dominant one with $P = 0$, one with $P = 7$ and the Pareto optimal equilibrium with $P = 16$. There are two critical masses: one with $P = 2$ allowing to overcome the zero outcome and reach the equilibrium $P = 7$, and one with $P = 14$ that allows to reach the Pareto-optimal equilibrium. None of the critical masses correspond to (unstable) equilibria, because they correspond to an attendance P with many individuals with the same threshold $H_i = P$, like in seminar SW1 for $H = 8$.

SM2. There are three stable Nash equilibria: a risk-dominant one with $P = 0$, an equilibrium with $P = 9$, and the Pareto-optimal one with $P = 16$. The critical masses here are $P = 2$ and $P = 14$; none of them are (unstable) equilibria.

Seminars SM1 and SM2 are particularly interesting: in order to reach the Pareto optimal equilibria, the system needs risk-prone individuals.

3.2 Experimental treatments

We experiment four information treatments. In all the settings, subjects know their own thresholds and payoffs earned at all the previously played periods. Besides, supplementary informations are given, which are different in each of the tested treatment:

On Line treatment (OL). Each individual is given the values of all the other players' thresholds. Within each period, the on-line (i. e. without waiting the end of the period) decisions of all the participants, as well as the number of participants to each seminar in all the preceding periods, are displayed. This treatment provides *a priori* maximal chances for an efficient coordination.

Attendance-based treatment (NP). Each individual is given the number of participants to each seminar at all the preceding periods. This information is refreshed after each period. Thus, each subject knows the past attendances and his own payoffs, but not the others' thresholds nor their decisions at the time he must make his own decision.

Threshold-based treatment (H). Each individual, independently of his decision, is informed about whether the number of participants in the last period has reached his threshold or not.

Earning-based treatment (E). For each individual, only his successive payoffs at the different periods and his cumulated gains are displayed, without any additional information. As a consequence, individuals not attending the seminar do not have any information of the outcomes.

3.3 Procedure

In each experiment, subjects decide at each period whether to participate or not in each of the four seminars. The different information treatments have been tested in separate experimental sessions, with different subjects, to guarantee independence of the tests. Upon arriving at the laboratory, the subjects are randomly seated in front of the computers. Communication between them is strictly forbidden. The instructions are read loudly before the experiment is started, to make sure that everybody understands the rules, so that they are common knowledge. Paper copies of these instructions are distributed among the subjects. In our setting, the payoffs defined in (2) are $U = 200\,yen$ (a fictitious currency converted to euros at the end of the experiment at the rate $100\,yen = 1.45\,euro$), $V = 0\,yen$ and $W = 50\,yen$ per period. The subjects are told that they will gain the cumulated payoffs of one of the seminars, which will be selected at random at the end of the experiment. Notice that individuals who decide to never attend the seminar obtain TW, and those who always make the wrong choice do not gain anything. In experimental economics, W is viewed as an initial endowment, that is lost whenever the individual decides to participate and his threshold is not reached, so that his net gain is $V = 0$. Within this point of view, if the threshold is reached, he obtains a supplementary payoff equal to $U - W$.

The information concerning the four seminars is displayed on four distinct squares on the screen in front of the subject. His threshold for each seminar is indicated inside the corresponding square. Once all the players have validated their decisions for all the four seminars, the period is closed. The subject's gains of the period are displayed on the screen inside the square of the corresponding seminar. The treatment-dependent information about the current and previous periods status are displayed and a new period begins. Sessions lasted about 1h 30. At the end of the experiment, subjects are asked to answer anonymously several questions on a paper questionnaire. These answers are used to enrich the interpretations of the experimental results, and to help to identify the components of information that most affect the subjects' decisions. After the completed questionnaires are returned to the experimenter, subjects are paid an amount corresponding to their cumulated yens in one of the four seminars, which is selected at random.

4 Experimental results

In this section we report the experimental results. The aggregate results are represented by the number of participants as a function of the period.

Treatment OL. Subjects reached very early the Pareto optimal equilibrium: in all the seminars the saturation of 16 participants is achieved (see figures 2, upper-left). Allowing the information to be shared among the subjects speeds up the convergence to this equilibrium: the attendance stabilized already at the second period. This kind of dynamics is actually expected, since the full information represents a potentially powerful coordination mechanism, as subjects may decide to participate just when their thresholds are reached, or may participate through "imitation" of the others (who share or not the same thresholds). The drops in the attendances in the middle of the experiment in seminars SW2, SM1 and SM2 are due to what we call non-rational behavior

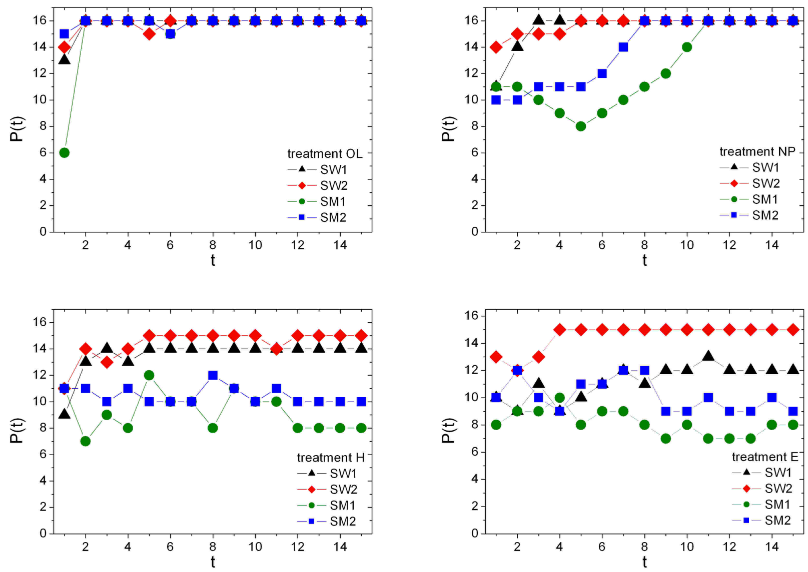

Fig. 2. Aggregate results: the number of participants to the four seminars as a function of the period, for each treatment separately.

of some individuals, who tried to inflect (as comes out from the answers to the questionnaire) the collective dynamics without success.

Treatment NP. In all the seminars, the saturated coordination on the Pareto optimum is reached, although at longer times than in treatment OL (see figure 2, upper-right). The access to the period by period attendance may induce an adaptive behavior: as soon as the attendance reaches the subject's individual threshold, he may decide to participate in the next period. Individuals satisfying $H_i = P(t) + 1$ may maximize their payoffs through participation at the next period (assuming that subjects who participated at the previous period will persist in participating). This is just like a myopic learning: if the number of participants is smaller than the individual threshold the subject does not participate and does otherwise. These behaviors were observed in each of the four seminars, although not among all the subjects.

There is a clear correlation among the seminars' attendances: they converge one after the other. This dependent dynamics indicates that individuals learn to coordinate based on the success or failure in the other seminars.

Treatment H. Different from treatment NP, in the present case individuals do not know how far the attendance $P(t)$ is from their thresholds: all they know is whether their thresholds for participation have or not been reached. Despite this rather succinct (private) information we still observe a good coordination among them (see figure 2, down-left). However, in contrast with treatments

OL and NP, coordination arises, but not on the Pareto optimal equilibrium. Individuals with $H_i = P(t) + 1$, who might improve their earnings through participation in treatment NP, may persist in non participating in the present case. Participating entails a risk for all the non-participants, independently of their thresholds. This explains the relative slow convergence of seminars SM1 and SM2, that reach their equilibria only after 12 periods.

Treatment E. This is the minimum possible information that subjects may possess in this game, since it is strictly private: subjects know their own actions and earnings. Thus, only the subjects that afford the risk of participating obtain some information about the attendance level, the others just earn their endowments. There are more temporal fluctuations in the aggregate attendance. Roughly speaking, due to the lack of information the subjects seem to be groping around, trying to explore the consequences of their decisions to obtain information concerning the attendance. A good coordination is still observed (see figure 2, down-right), although besides seminars SW1 and SW2 the attendances are not yet stabilized after the 15 played periods. Surprisingly, despite the low information level, there is no convergence to the equilibrium with $P = 0$. Even if these seminars did not reach a stable attendance, we expect that allowing playing for more periods should not make the attendances drop significantly. It seems that the willingness to get large payoffs overcomes the individual risk aversion. This may also be a consequence of the small size of the group.

An unexpected result of our experimental setting is the following: convergence to high attendances in difficult seminars (that need coordination of large numbers of individuals to jump from a Nash equilibrium to the Pareto optimal one) may be induced through simultaneous play of other, less risky games.

4.1 Analysis seminar by seminar

Figure 3 shows the results of the four treatments -number of participants at each period- for each seminar separately. There is a clearcut difference on the type of dynamics, depending on the shape of the corresponding thresholds distribution. From this point of view, as discussed in section 3, seminars SW1 and SW2 belong to the same class of distributions, with no intermediate attendance equilibrium, whereas seminars SM1 and SM2 have bimodal distributions and intermediate attendance equilibria. In the discussion we group together seminars with similar behaviors.

First period decisions. The payoff-dominant strategy $\omega = 1$ is generally the majority choice in the first period play, in agreement with the literature [8, 9]. In all but SM1 seminar, the critical masses that allow to overcome the risk-dominant equilibria are overshot already at the first period. This strategy seems a natural way of exploring the possible outcomes when playing with unknown opponents.

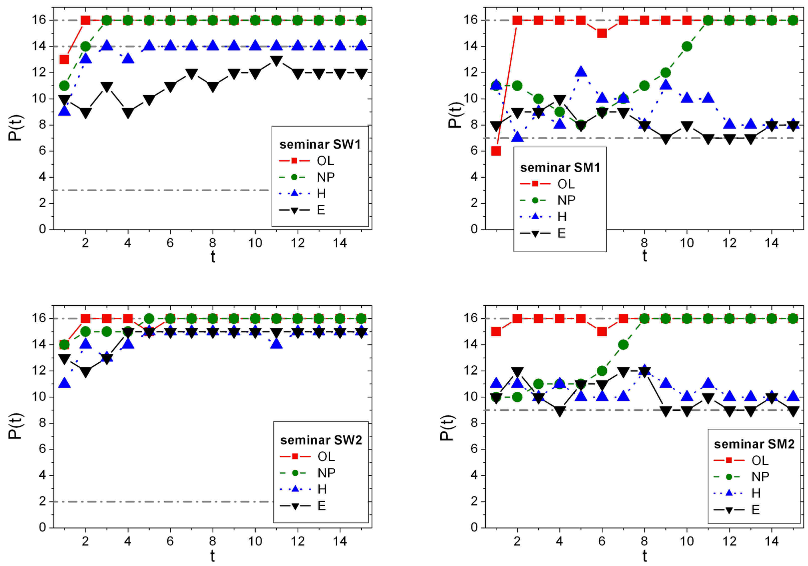

Fig. 3. Aggregate results: the number of participants as a function of the period for each seminar separately, and for four treatments within each seminar. Nash equilibria are represented as dash-dot horizontal lines.

Subsequent periods: seminars SW1 and SW2. Both seminars illustrate a remarkable coordination at the aggregate level toward an equilibrium with high attendance in all the four treatments, although not always the Nash equilibrium is reached (see figures 3, left). In particular, within treatment NP, efficient coordination of all the 16 participants is achieved later than in treatment OL. In the treatments with less information content, there is still a successful decentralized, though incomplete, coordination: even with treatment H seminars converge to equilibria with relatively high participation (the stable equilibrium $P = 14$ in seminar SW1 and the unstable equilibrium $P = 15$ in seminar SW2). Remind that within this treatment, the individual risk is higher because subjects do not know the exact attendance.

In treatment E, attendance in seminar SW1 reached $P = 12$, smaller than that of the Nash equilibrium ($P = 14$). The individual with $H = 0$ never participated, and he reported in the questionnaire that he did not understand the meaning of zero threshold. The individual with threshold $H = 9$ decided not to participate from the start and did not try to learn. Otherwise the system could have reached the stable Nash equilibrium with $P = 14$. In seminar SW2 the individual with $H = 16$ after two attempts with unsuccessful results never participated again. This is why the attendance at convergence was $P = 15$.

Subsequent periods: seminars SM1 and SM2. In treatment NP, these seminars converged to the Pareto optimal equilibrium although slower than seminars SW1 and SW2 (see figures 3, right). Notice the change in the dynamics in seminar SM1: initially the system seems to evolve toward the Nash equilibrium with $P = 7$, but its dynamics changes suddenly after period 5. As already pointed out, at that period seminars SW1 and SW2 reached their full attendances. This success probably induced the individuals not attending so far to take a risk in seminars SM1 and/or SM2. The attendance recovers and rises monotonically to reach the Pareto optimal equilibrium at the 11th period. The change is induced by the subjects with high thresholds, that we interpret as presenting successful teaching behaviors.

In treatment H, both seminars reach intermediate attendances, that fluctuate close to the stable Nash equilibria. Interestingly, these fluctuations are systematically positive: the attendances are equal or larger than the ones expected at the Nash equilibrium. Fluctuations are due to risky attempts of individuals with higher thresholds, that we may interpret as trying to induce full attendance through unsuccessful teaching. These fluctuations are more conspicuous in treatment E. Remember that with this treatment, the only way to explore the actual attendance is to participate, at the risk of loosing the endowment, i.e. getting a vanishing payoff.

Summary. The main result of our experiments is that even without complete information an efficient coordination is possible. Even if the information given to the players is strictly private —i.e. only individual earnings— the attendance reached at least suboptimal equilibria with good participation. This result is in contrast with the earlier experimental observation of Devetag [2] under similar information treatments: she observed no coordination at all if the subjects know only their previous individual earnings. In fact, Devetag considered a critical mass game in which players have to choose a number i in the range $i \in [0, N]$. If the number of subjects that chose number i is larger than i, they all get the same payoff, which is proportional to i. Choices not satisfying this condition have zero payoff. This game has N pure Nash equilibria. Clearly, the payoffs of Devetag's game explicitly incites coordination of large groups, through the choice of large values of i. On the other hand, having so many possible strategic choices induces a larger uncertainty in anticipating an equilibrium. It would be interesting to explore further how the number of possible strategies and Nash equilibria affect coordination.

5 Learning models

In this section we present our attempts to interpret the experimental outcomes in terms of agents' learning. We simulate a system of $N = 16$ agents that decide repeatedly whether to participate or not, under the four information treatments: OL, NP, H and E. We present two learning models: learning attractions with trembling hand decisions and cumulative proportional reinforcement learning (CPR).

5.1 Learning attractions with trembling hand

Following [1], we assume that agents associate attractions to each of the two strategies, and these attractions determine the probabilities with which strategies are chosen when agents make decisions repeatedly. In problems of binary decisions, it is sufficient to consider only the attraction of one strategy, say participating, with respect to the other. Depending on the informational setting, attractions may be modelized by the difference between the agent's expected number of participants $\hat{P}_i$ and his threshold H_i: $A_i = \hat{P}_i - H_i$, or simply by the expected surplus upon participating, which in our experiments can take either of two values: $A_i = U - W$ or $A_i = V - W$.

Initial states. Before making his first decision, each agent i starts with an attraction $A_i(0)$ that reflects his initial belief. Agents may begin with totally random estimations or estimations possibly correlated with their thresholds.

Then at each period t, each agent i chooses $\omega_i(t)$ following a decision rule that depends on his attraction $A_i(t)$. Once decisions made, agents receive the corresponding payoffs and the supplementary information corresponding to the treatment considered. Then, attractions are updated using a learning rule based on the grasped information as described in 5.1.

Decisions. Since the experimental results show that not always the subjects make the optimal decisions, here we assume trembling hand dynamics [6]: there is a small chance that an agent's hand trembles when he chooses, so that both strategies have positive probabilities to be selected. If an agent's best response is $\omega_i^{BR}(t) = 1$ if $A_i(t) \geq 0$ and $\omega_i^{BR}(t) = 0$ otherwise, then at each period t each agent i chooses a strategy $\omega_i(t)$ according to

$$\begin{array}{ll} \omega_i(t) = \omega_i^{BR}(t) & \text{with a probability } 1-\epsilon \\ \omega_i(t) = 1 - \omega_i^{BR}(t) & \text{with a probability } \epsilon \end{array} \tag{8}$$

where ϵ is a small positive number $0 < \epsilon < 1$, for simplicity assumed to be the same for all the subjects.

Updating attractions. According with the subjects' claims in the questionnaire[3], we assume here that agents are myopic: they respond to the observation of only the previous period. In treatments OL and NP, agent i, unconditionally to whether he participated or not in the previous period, knows the actual number of participants $P(t)$, and updates the attraction according to

$$A_i(t+1) = P(t) - H_i. \tag{9}$$

In treatments H and E, agents update attractions using the earned payoffs π (see eq. (2)). More precisely, in treatment H:

[3] Subjects responded that they used only the observation of the preceding period, disregarding older informations.

$$A_i(t+1) = \begin{cases} \pi_i(t) - W & \text{if } A_i(t) \geq 0, \\ 0 & \text{if } A_i(t) < 0, \text{ and } P(t) \geq H_i, \\ A_i(t) & \text{if otherwise} \end{cases} \tag{10}$$

and in treatment E according to:

$$A_i(t+1) = \begin{cases} \pi_i(t) - W & \text{if } A_i(t) \geq 0, \\ A_i(t) & \text{if otherwise} \end{cases}. \tag{11}$$

Simulation results

On figures 4 we present the simulation results: the number of participants as a function of the period. Treatment OL is simulated as a sequential best response dynamics: agents make their choices in the same order as the experimental rank of the first period decisions. Taking the experimental first period choices as initial conditions, the simulations fit well the empirical results (see figure 4, upper-left). This is neither surprising nor a big performance, since the system converged at the second time step.

We simulate the other three treatments using parallel dynamics: at each period agents make decisions simultaneously. We use as initial conditions the first period experimental decisions of the corresponding treatments. We further take into account the systematic correlations in the convergence times (as it was observed during the experiment). In doing so, we sort seminars according to their convergence times: SW1, SW2, SM2, SM1 in treatment NP, and SW2, SW1, SM2, SM1 in treatments H and E.

Then, the dynamics is the following: in treatment NP, starting with $\epsilon > 0$, we make parallel updates for seminar SW1 first. Then, at a certain period t^0 we put $\epsilon = 0$, so that agents begin to give their deterministic best replies. When seminar SW1 converges at period $t*$, we put exogenously $\epsilon = 0$ for the seminar SW2 starting from this period $t*$. Then, when seminar SW2 converges at period $t**$, we put $\epsilon = 0$ starting from this period for seminar SM2, and so on until seminar SM1 is converged. We use a similar procedure for the other two treatments. Therefore, we manipulate with two parameters: t^0 and ϵ. In treatment E, we adjust the values of ϵ in such a way as to reproduce closely the data (this is why the values of ϵ inside the corresponding legend are not strictly zeros).

The simulated results present similar qualitative behaviors as on figures 2. We have experimented with different combinations of ϵ and t^0, and the ones that correspond to the present plots are indicated inside each legend.

5.2 CPR learning

To model agent's behavior in treatment E we can also apply an alternative learning scenario: the cumulative proportional reinforcement learning (CPR) of Laslier *et al* [3], a particular evaluative feedback [7] algorithm. Following

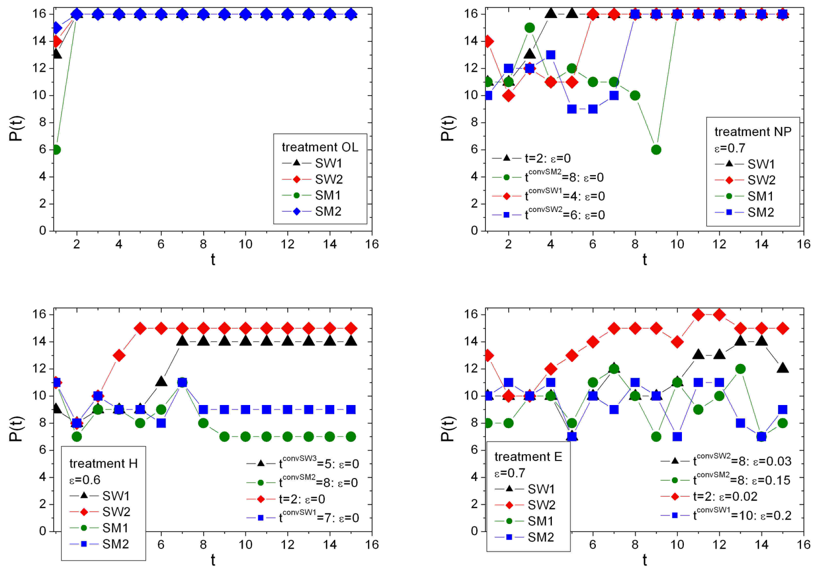

Fig. 4. Simulated results with trembling hand decisions: the number of participants as a function of the period.

this rule, agents choose a strategy with a probability proportional to the cumulative payoff obtained with that strategy. We consider this learning rule in treatment E, because it is the only treatment where information is obtained exclusively through the strategic decisions, like in reinforcement learning scenarios.

According with this rule, each agent i chooses to participate ($\omega_i = 1$) with a probability

$$p_i(t) = \frac{C_i(t)}{\Pi_i(t)}, \tag{12}$$

where $C_i(t)$ is the sum of the payoffs obtained up to t by using the strategy $\omega = 1$:

$$C_i(t+1) = \sum_{t=0}^{T-1} I(\omega_i(t), 1)\pi_i(t) \tag{13}$$

($I(x, y)$ is the indicator function: $I(x, x) = 1$, $I(x, y) = 0$ for $y \neq x$), and $\Pi_i(t)$ the expected cumulative payoff (which is nothing but the sum of cumulative payoffs expected both from participating and not). This rule incorporates the *exploration-exploitation* dilemma since the individuals use more and more often the best performing strategies without eliminating totally any other one.

The simulation results are shown on figures 5. We start with the first period experimental decisions as initial conditions, and we adjust the values

$C(0)$ and $\Pi(0)$ (the same for all the individuals) in order to fit the experimental results (the corresponding numerical values are indicated inside each figure). The simulations results fit the qualitative behavior of the experimental curves for the four seminars. The fits are better than with the probabilistic myopic learning, meaning probably that despite their claim individuals took into account more remote past results than just that of the preceding period.

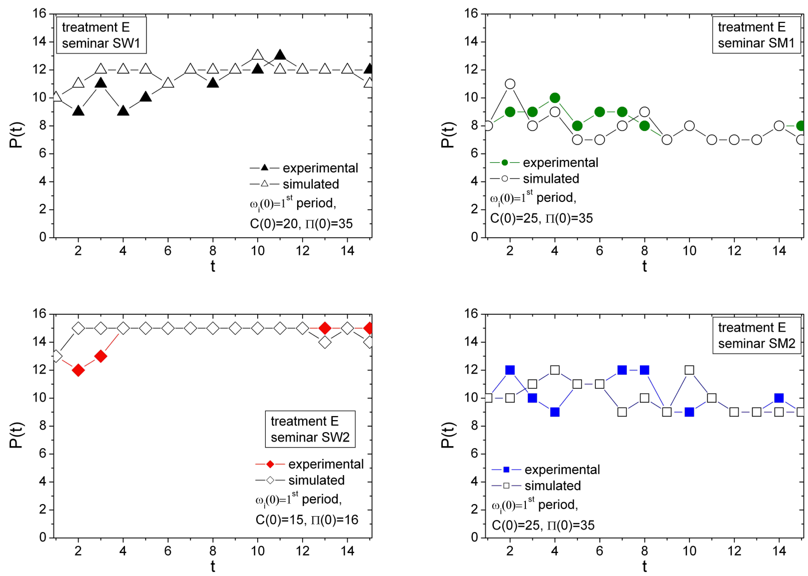

Fig. 5. Treatment E with CPR learning rule: simulated number of participants as a function of the period.

6 Conclusions

In this paper we present experimental results on the dying seminar game, obtained under different information treatments. The purpose of the experiments was to explore the influence of available information on the game outcomes, and to determine the minimum information ingredients which provide an efficient coordination. The main result is the emergence of sufficient decentralized coordination under incomplete information, although not always on a Pareto optimum.

We analyzed the experimental results in terms of initial beliefs and individual myopic learning of the attractions for attending the seminar, obtaining a good qualitative agreement. Simulations show that initial beliefs are crucial in determining the dynamical path of the system. However, the better fits obtained with CPR learning suggest that individuals take into account remote experiences, despite their perception that they only considered the last outcome.

References

1. C. F. Camerer: *Behavioral Game Theory.* (Princeton University Press 2003, Princeton, New Jersey)
2. G. Devetag: Coordination and information in critical mass games: an experimental study. Experimental economics **6**, 53-73 (2003)
3. J.F. Laslier, R. Topol, B. Walliser: A behavioral learning process in games, Games and economic behaviors **37**, 340-366 (2001)
4. J.P. Nadal, D. Phan, M.B. Gordon, J. Vannimenus: Multiple equilibria in a monopoly market with heterogeneous agents and externalities. Quantitative Finance **5(6)**, 557-568 (2006)
5. T.S. Schelling: Micromotives and Macrobehavior. *W.W. Norton and Co, N.LY (1978)*
6. R. Selten: Re-examination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory **4**, 25-55 (1975)
7. R. S. Sutton and A. G. Barto: Reinforcement learning: an introduction. *MIT Press, Cambridge, MA*, (1998)
8. J. Van Huyck, R. Battaglio, R. Beil: Tacit coordination games, strategic uncertainty and coordination failure. American Economic Review **80**, 234-248 (1990)
9. J. Van Huyck, R. Battallio, J. Cook: Adaptive behavior and coordination failure. Journal of Economic Behaviour and Organization **32**, 483-503 (1997)