Dumbing down rational players: Learning and teaching in an experimental game

Antoine Terracol[a,*] Jonathan Vaksmann[b]

[a] *EQUIPPE, Universités de Lille; and CES, Université Paris 1 Panthéon-Sorbonne, CNRS. 106-112 boulevard de l'Hôpital, 75642 Paris cedex 13, France. Tel: (33) 1.44.07.82.51; Fax: (33) 1.44.07.82.47*

[b] *CES, Université Paris 1 Panthéon-Sorbonne, CNRS. 106-112 boulevard de l'Hôpital, 75642 Paris cedex 13, France. Tel: (33) 1.44.07.82.51; Fax: (33) 1.44.07.82.47*

Abstract

This paper uses experimental data to examine the existence of a teaching strategy among boundedly rational players. If players realize that their own actions modify their opponents' beliefs and actions, they might play certain actions to this specific end and forego immediate payoffs if the expected payoff gain from a teaching strategy is high enough. Our results support the existence of a teaching strategy in several ways. After exhibiting some regularities consistent with teaching, we examine more precisely the existence of such a strategy. First we show that players update their beliefs in order to take account of the reaction of their opponents to their own action. Second, we examine whether players actually use a teaching strategy by playing an action that induces a poor immediate payoff but is likely to modify the opponent's behavior so that a preferable outcome might emerge in the future. We find strong evidence of such a strategy in the data and confirm this finding within a logistic model that suggests that the future expected payoff that could arise from a teaching strategy has indeed a significant impact on choice probabilities. Finally, we investigate the effective impact of a teaching strategy on achieved outcomes and find that more tenacious teachers can successfully use such a strategy in order to reach their favorite outcome at the expense of their opponents.

Key words: Game theory, Teaching, Beliefs, Experiment
JEL codes: C72, C91, D83

* Corresponding author
Email addresses: terracol@univ-paris1.fr (Antoine Terracol), jonathan.vaksmann@univ-paris1.fr (Jonathan Vaksmann).

Preprint submitted to Elsevier *6 February 2009*

1 Introduction and motivations

Traditional game theory focuses on strategic interactions among fully rational players, but as is now widely recognized, human reasoning might be bounded. Thus, many approaches have suggested ways to encompass these limits into game theory. While some studies address the theoretical side of the question by examining, for example, convergence supported by evolutionary forces or adaptive rules, other approaches, based on empirical standards, typically use experimental data to provide a more accurate description of players' behavior. In the present study our aim is to contribute to the latter framework.

In the part of the literature designed to describe the way people play their games, several studies analyze the interactions among adaptive players who choose their response according to what they have learned in the past. Some studies (e.g. Roth and Erev, 1995, and Arthur, 1991) focus on reinforcement learning where players determine their actions according to their success in bringing high payoffs in the previous periods of the game. In belief-based learning models (Cheung and Friedman, 1997; Boylan and El-Gamal, 1993; Mookherjee and Sopher, 1994, 1997; Rankin et al., 2000; Fudenberg and Levine, 1998) players use the past history of the game to update their beliefs about what their opponents may play. Finally, other studies (e.g. Camerer and Ho, 1999) take both reinforcement and belief learning as two components of the same learning model, known as the EWA learning model, which has been proven to be very successful in predicting how people behave. All these models provide foundations for equilibrium theory and offer opportunities to model empirically observed behavior.

However, one might think that these approaches are at odds with the foundations of traditional game theory by ruling out perfect rationality and regarding people as merely adaptive. In this paper we attempt to exhibit empirically a behavior that stands between these two extremes.

The starting point of our study is that in the learning models, players' decision process has only a backward component: they respond according to what they have experienced in the past. This rules out any awareness of the fact that opponents might also learn from other players' actions and might thus be influenced by them. However, if a player suspects that his opponent can also learn from the past history of the game, he might attempt to play strategically to drive him into a particular set of actions. That is how a *teaching strategy* might arise (i.e. playing an action that does not necessarily maximize the expected payoff at the current round, but increases the probability of convergence to equilibrium deemed preferable). Thus we aim to exhibit, among adaptive players, a piece of sophistication or, from another point of view, we are actually dumbing down rational players.

What do we know about people strategic reasoning in experimental repeated games? The fact that players might be willing to make good use of strategic interactions in order to manipulate their opponents seems natural and has been referred to in some stylized situations in economics. In repeated public good games for instance, it might be argued that players choose positive contributions in order to set a good example and signal their willingness to cooperate in latter rounds so that they would induce others to do the same. In Cournot games, players could play aggressively by producing a high quantity, hoping that their competitors will reduce their productions in the future. However, it might be noted that, surprisingly, these reasonings have not always proven to be as relevant as expected. In an experimental public good game Offerman et al. (2001) find that players' belief formation process is consistent with naive Bayesian expectations where people do not take strategic interactions into account. However, as noted by the authors, the step-level public good game they use is strategically complicated. This game is characterized by the fact that a funding threshold has to be reached before the good can be provided. It has multiple equilibria, which might appear computationally demanding so that strategic reasoning is made very difficult. On the other hand, Huck et al. (1999) run an experimental Cournot game and find that players tend to imitate the others so that a high production does not induce low productions from the competitors in the future and behavior converges to the competitive (or Walrasian) outcome. In a related research, Dürsch et al. (in press) run a Cournot game where human players face a pre-programmed computer that follows one of some usual learning algorithms; the authors find that human players frequently try to take advantage of the learning process of their adaptive "artificial" opponent and teach it in order to increase their payoff. However, consistent with Huck et al. (1999), the authors also find that imitation is the unique learning rule that prevents human players from teaching their opponents. Moreover, looking at the "human vs human" data from Huck et al. (2004), Dürsch et al. find that when human subjects interact with other human subjects, they tend to behave as predicted by adaptive learning models and do not teach.

These facts seem to indicate that teaching is not relevant in every situation. Indeed, teaching might become an intricate job in too complicated games. Likewise, particular forms of learning (e.g. imitation) might not be compatible with a teaching strategy. Still, rejecting strategic reasoning in game theory might appear strange, and one could suspect that in other environments, there might be some room for strategies more sophisticated than the myopic strategies induced by learning models. Indeed, the literature provides some studies that unambiguously emphasize the limits of adaptive learning models and suggest that players act more strategically than implied by these models. For example, Ehrblatt et al. (2009) analyze how teaching speeds up convergence to a unique (pure strategy) equilibrium. They suggest that when a player's Nash action converges before his beliefs, he must have chosen this action even if it does not maximize his short run payoff (because it is outside his best response set) in order to teach his opponent and get a higher payoff in the long run. They then use a criterion to separate teachers and learners in pairs of players: the player whose actions converge first in a pair is said to be the teacher and the player whose actions converge later is a learner. On the other hand, Camerer et al. (2002) add sophistication to adaptive models by assuming a population composed of two types of players: a fraction of them are supposed to be fully rational and consequently have the ability to exhibit equilibrium behavior while the remaining fraction is regarded as adaptive and only looks backward to choose the current action. Rational players have in mind their own estimation of the repartition (not necessarily the real one) between rational and adaptive players, using their knowledge to outguess their adaptive opponents. Hence the authors examine how people learn in a heterogeneous population of players.

In our approach we let every player have a piece of sophistication, and we regard them as neither extremely sophisticated nor extremely unsophisticated. Thus we investigate teaching when every player has the ability to teach and is willing to lead his opponents to his favorite outcome. In order to make our purpose non trivial, we wanted players' favorite outcomes to be different and thus used a game with diverging interests. We are also interested in studying strategic interaction in pairs of players rather than in a population of players. Indeed, one can reasonably expect that manipulation would be more prominent when players interact with the same opponent throughout the game.[1]

In Section 2 we describe our experimental design. In Section 3, we discuss our results by first exhibiting some regularities that support the existence of teaching. We then turn to a thorough investigation of a teaching strategy. First, we test whether players see their opponents as learners who observe the history of the game and modify their behavior accordingly, which indeed represents a necessary condition for players to adopt a teaching strategy. To this end, we estimate the difference between players' actual (or "true") beliefs, elicited using an appropriate scoring rule, and γ-weighted beliefs (Cheung and Friedman, 1997) where players only consider the past history of their opponents' actions. We then examine whether this difference can be explained by players' own past actions by testing whether they perceive that their own actions can modify their opponents' behavior. In the next step, we investigate whether players actually use a teaching strategy. Teaching implies that players could choose actions with poor immediate expected payoff that are likely to influence the opponent's behavior so that they might lead to a more profitable outcome in the future. Thus, we examine this implication in two parts. First, we test the existence of such strategy by examining players' tendency to depart

[1] For related theoretical approaches on that point, see Ellison (1997) and Schipper (2006).

from best responding by playing actions that might lead to the emergence of a more favorable equilibrium in future. Second, we test if players choose their actions according to the prospective payoff gain that might follow from a modification of their opponents' beliefs. More precisely, we estimate a logistic model where the probability of a given action being played depends not only on the immediate expected payoff, as it is the case when players are regarded are merely myopic, but also on the cumulative future payoff gain induced by the expected modification of the opponent's behavior. Finally, we examine the actual consequences of teaching and analyze whether using a teaching strategy is an effective tool for players to drive the outcome of a game. Section 4 finally concludes.

2 Experimental design and procedures

The experiment was run using the computerized experimental laboratory of the University of Paris 1 Panthéon-Sorbonne from the Summer through the Fall of 2006.[2] No subject had any training in game theory. Each experimental session lasted almost one hour and a half. All sessions consist of 30 repetitions of the game represented below under two strategic treatments we will describe in this section. The written instructions given to the subjects clearly stated that the game would last for 30 repetitions, so this was publicly known by every subject. During the experiment, players were evenly and randomly divided into type 1 and type 2 players. Payoffs were denominated in units of experimental currency and converted into Euros at the end of each session. The subjects, on average, earned approximately €14.4 for their participation. They were paid €3 just for showing up.[3] A translation from the original French instructions given to the participants can be found in an appendix at the end of the paper.

The game we used can be represented by the matrix below where row player is of type 1, while column player is of type 2. It is in fact a reduced form of the duopoly game initiated by Hamilton and Slutsky (1990). In their set-up, X,Y and Z refer respectively to the Cournot, the Stackelberg leader and the follower quantities.

[2] For conducting the experiment we used the experimental software 'Regate' (Zeiliger, 2000).
[3] Throughout the paper, all payoffs will also be denominated in units of experimental currency. Subjects have been paid according to the sum of the payoffs they received during the 30 repetitions of the stage game. Every 200 units of experimental currency could be redeemed for €1.6 after each session.

Payoff matrix

type 1 \ type 2	X	Y	Z
X	(40,52)	(22,46)	(40,52)
Y	(35,40)	(10,20)	(44,46)
Z	(40,52)	(30,60)	(40,52)

This game has many features we desired in our design. First of all, it is easy to understand. Then, it has three non Pareto-rankable Nash equilibria: (X,X), (Y,Z) and (Z,Y) which are not too difficult to calculate or learn deductively.

The multiplicity and non Pareto-rankability of these equilibria are useful features for our purpose. Indeed, as already noted, we are interested in investigating teaching incentives in a set-up where both players are potential teachers and are willing to lead their opponent to their favorite outcome. On the other hand, diverging interests make teaching more useful and consequently more likely to emerge because players have conflicting preferences over equilibria. For instance, at (Y,Z), type 1 players get their best payoff. Thus, it is natural that type 1 players would like to make (Y,Z) emerge. However, one can easily notice that at this equilibrium, the type 2 opponents get their worst equilibrium payoff, hence this equilibrium exhibits a strong conflict of interests. Likewise, (Z,Y) is the best outcome for type 2 players while it is the worst equilibrium in terms of payoff for the type 1 opponents. Then, consistent with the terminology used by Hamilton and Slutsky [4], if we observe convergence to (Y,Z) (resp. (Z,Y)) in a pair, one could say that type 1 (resp. type 2) players take the leadership while the opponents are the followers. Consequently, throughout the paper, we will refer to (Y,Z) and (Z,Y) respectively as type 1 and type 2 leadership equilibria.

Finally, one last interesting feature is the use of asymmetric payoffs that generates interesting testable implications concerning relative teaching incentives across players.

For each session we conducted, we adopted one of two treatments that differ according to the matching protocol that was implemented. More precisely, we performed a fixed-opponent, or ‘Partner’, treatment and a random opponent, or ‘Stranger’, treatment. In both treatments, the type assigned to each player remains the same throughout the experiment. If players attempt to ‘teach’ their opponent, the way in which they are matched when a game is played repeatedly might affect behavior. Thus, running these two treatments allows

[4] More precisely, in Hamilton and Slutsky, (X,X) corresponds to the Cournot equilibrium, while (Y,Z) and (Z,Y) correspond to the Stackelberg equilibria.

us to test the impact of random matching. We recruited 76 subjects, 40 for the Partner treatment and 36 (two sessions of 18 subjects each) for the Stranger treatment.

In each round, before choosing their action, subjects' beliefs were elicited using a proper scoring rule defined below. They were asked to report their beliefs, or predictions, about the likelihood that their opponent would use each of his actions available in the current round (i.e. X, Y or Z).

The belief elicitation procedure takes the classical quadratic form used in the literature. Subjects were asked to report the probability that their opponent will play X, Y or Z. Such a report takes the form of a vector $b = (b_X, b_Y, b_Z)$ where b_a represents the belief held by the subject associated with the action a of his opponent (a =X,Y,Z). A player's payoff when he reports b and his opponent actually uses action a is given by $\left[8 - 4\left((1-b_a)^2 + \sum_{z \neq a} b_z^2\right)\right]$.

This formula means that each subject receives an endowment of 8 units of experimental currency at the beginning of each round and reports his beliefs. The amount $(1-b_a)^2$ subtracted from the initial endowment of the subject corresponds to a penalty for having reported an inappropriate belief for the action a his opponent has played in the current round. Note that this penalty equals 0 when the subject reports a probability $b_a = 1$ and his opponent plays a at the current round. Subjects are also penalized for having stated inappropriate beliefs for the other actions by a subtraction of an amount $\sum_{z \neq a} b_z^2$ from their initial endowment of 8 units of experimental currency. The worst possible guess (i.e. putting all the probability weight on an action that the opponent does not actually choose) leads a payoff of 0 (and explains the normalization constant (4) which appears in the formula). It can be easily demonstrated than according to this computation, risk neutral subjects should tell the truth. [5]

As usual in this kind of design, the reward for reporting beliefs remains small in comparison with the payoffs associated with the game. Indeed, this is an important point because a too large reward could affect subjects' behavior due to the possibility of playing any particular action repeatedly so as to maximize their prediction payoffs at the expense of their game payoffs.

At the end of each round, subjects were informed about the action of their opponent, their game payoff, their prediction payoff, and the game payoff of their opponent for the current round. When deciding in a given round, subjects could always see on their screen information about their own past

[5] Several studies (e.g. Sonnemans and Offerman, 2001, Nyarko and Schotter, 2002) indicate that, with this quadratic scoring rule, players indeed report the truth. Rutström and Wilcox (2006), however, find that an intrusive scoring rule for belief elicitation affects people's behavior.

actions, game payoffs and predictions in earlier rounds. Actions and game payoffs of their opponents for the past rounds also remained present on the screen.[6] Note that in the Stranger treatment, players were only informed about their opponents' actions and game payoffs at the current round, but they were *not* given the whole history of actions and payoffs of each of their successive opponents for the past rounds.

3 Results

3.1 A first overview

Before we turn to a more detailed analysis highlighting the existence of teaching, we start by pointing out some prominent regularities in the data that could indicate the existence of sophistication in players' behavior.

Among the 76 subjects who participated in the experiment, 4 individuals (all in the Partner treatment) reported clearly inconsistent beliefs (for example always putting a probability 1 on a given action, or a probability 1 on each of the three actions in turn, regardless of their opponent's actual behavior) or had no variation in their actions and/or beliefs, thus preventing the statistical analysis of their behavior at the individual level carried out in Section 3.2; we had to discard these subjects in the analysis. Two of these 4 puzzling subjects were paired together, and we discarded this pair in the analysis; obviously, the partners of the 2 others had to be discarded as well. Thus 3 pairs among 20 were discarded in the Partner treatment. All individuals in the Stranger treatment were kept for the analysis. We are left with 34 players (17 pairs) in the Partner treatment, and 36 players in the Stranger treatment.

Teaching involves playing an action that is not necessarily an immediate best response but that is likely to modify the opponents' behavior profitably in the future. Thus, if players use a teaching strategy, they might be led to take 'suboptimal' decisions. Tables 1 to 4 present in the rows players' decisions and in columns their myopic optimal decision given stated beliefs (BestX, BestY, BestZ), separating by type and treatment. Off-diagonal elements give the number of suboptimal decisions. These suboptimal decisions would be regarded as errors in usual learning models, but as we will see in this section and analyze

[6] For the reason noted above, it is natural in this kind of design to make sure that the prediction payoffs are not too prominent in comparison with the game payoffs. That is why subjects were informed about their prediction payoffs on the screen that appeared at the end of each round and recapitulated all the relevant information for the round, but contrary to game payoffs among others, prediction payoffs did not remain present on the screen in later rounds.

in greater details in the remainder of the paper, these decisions exhibit particular patterns that are inconsistent with an error process but consistent with a teaching strategy whereby players try to influence their opponent's future beliefs and behavior by playing actions that are not necessarily in their immediate best response set. Note that X is a best response when players put a zero probability on their opponent playing Y. X then only weakly dominates Z, and players are indifferent between X and Z. Since in this case players do not think their opponent will choose Y, action Z does not support any equilibrium whereas X supports the equilibrium (X,X). Consequently, in what follows, we will retain X as the optimal choice when players are indifferent between X and Z. As can be noticed below, this case remains relatively scarce in the data.

As can be seen in these tables, suboptimal decisions are not evenly distributed, as would be the case if they were random errors by players, but rather concentrated on specific actions; notably on playing Y when Z was the best response (accounting for 44 to 50% of suboptimal decisions). This pattern is particularly interesting because, for both players, deviating from action Z when it is a best response leads to the harshest cost of deviation. It is then hard to interpret these suboptimal decisions as errors since it would mean that players make many more errors when those errors are the most costly. On the other hand, such actions are to be expected from players using a teaching strategy, since their preferred equilibrium is supported by action Y while their worst equilibrium payoff is supported by action Z.

Table 1
Partner treatment, type 1 players

	BestX	BestY	BestZ
X	9	5	51
Y	5	154	75
Z	6	25	180

Table 2
Partner treatment, type 2 players

	BestX	BestY	BestZ
X	5	14	59
Y	2	74	111
Z	2	33	210

One could obviously expect players to use a teaching strategy to make their leadership equilibrium emerge by playing Y even if it is not a best response, so, even if we will turn to a more detailed analysis in section 3.3.1, it can be interesting, as a first approach, to examine the distribution of suboptimal decisions across types and treatments. Table 5 presents, in its first row, the frequency of all possible suboptimal decisions, separated by treatment and type.

Table 3
Stranger treatment, type 1 players

	BestX	BestY	BestZ
X	0	13	41
Y	1	206	89
Z	1	58	131

Table 4
Stranger treatment, type 2 players

	BestX	BestY	BestZ
X	1	6	34
Y	1	50	71
Z	0	41	336

It shows that while type 1 players take slightly fewer suboptimal decisions in the Partner treatment, the reverse is true for type 2 players: their proportion of suboptimal decisions is 53% higher in the Partner than in the Stranger treatment. Two-sample t-tests with unequal variances at the individual level (N=17 in the Partner treatment, and 18 in the Stranger treatment) reveal that type 2 players take significantly more suboptimal decisions in the Partner treatment than in the Stranger treatment (p-value=0.013) while differences in the frequency of suboptimal decisions across treatments is not significant for type 1 players (p-value=0.503). Moreover, in the Partner treatment, a paired t-test at the pair level (N=17) shows that type 2 players take significantly more suboptimal decisions than type 1 players (p-value=0.011). Conversely, type 1 and type 2 players in the Stranger treatment do not take such decisions in significantly different proportions, as a two-sample t-test [7] at the individual level shows (p-value=0.161).

In the three bottom rows of Table 5, we separate suboptimal decisions according to whether they involved playing X, Y or Z: subX is the fraction of suboptimal decisions X among all decisions, and similarly for subY and subZ. We apply the same tests as above. We find that type 2 players take significantly more suboptimal decisions X and Y across treatments (p-value=0.055 for suboptimal decisions X and p-value=0.037 for suboptimal decisions Y). Moreover, we find that type 2 players take significantly more suboptimal decisions Y than type 1 players within the Partner treatment. All the other 9 (5

[7] Because no fixed pairs of players exist in the Stranger treatment, paired t-tests are neither feasible nor adequate. We therefore use two-sample t-tests when testing differences between types in the Stranger treatment. We thus regard our Stranger treatment as a good proxy of a series of one-shot games where players are independent of each other. All tests presented in this study are two-sided.

inter types and 4 inter treatments) tests led to insignificant differences.

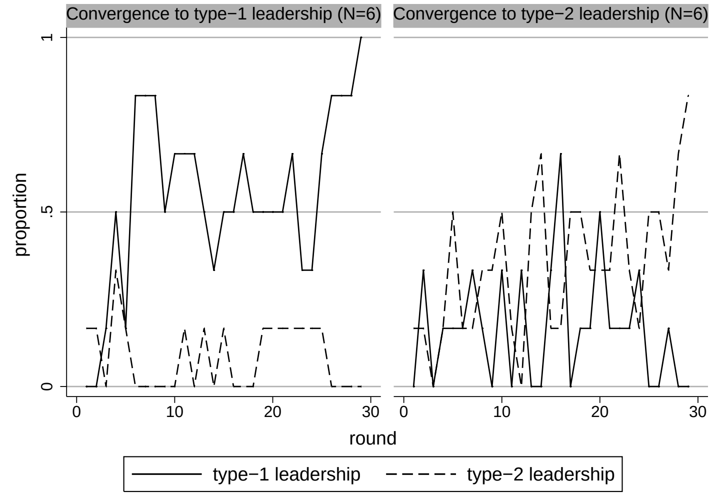
Table 5
Distribution of suboptimal decisions

	Partner		Stranger	
	type 1	type 2	type 1	type 2
All suboptimal decisions	0.327	0.433	0.376	0.283
By type of suboptimal decisions				
subX	0.110	0.143	0.100	0.074
subY	0.157	0.222	0.167	0.133
subZ	0.061	0.069	0.109	0.076

Overall, Tables 1 to 5 show that type 1 players do not behave differently across treatments with respect to suboptimal decisions, neither globally nor when examining specific suboptimal decisions. On the other hand, type 2 players have a higher frequency of suboptimal decisions in the Partner treatment, and this difference is essentially caused by an increase in suboptimal decisions X (+ 6.9 percentage points) and Y (+ 8.9 percentage points). Type 2 players thus seem to be more prone than type 1 players to deviate from their best response when the matching protocol gives them the opportunity to drive their opponents to a specific outcome. Moreover, these deviations tend to favor the action that supports the equilibrium giving them the highest possible payoff at the risk of high deviation costs. Incidentally, giving type 1 players more salient incentives to teach thanks to the fixed-matching protocol does not impact their behavior. This is consistent with the fact that type 1 players benefit less than type 2 players from the emergence of their leadership equilibrium.

Another interesting pattern concerns the description of convergence of pairs of players in the Partner treatment. We say that a pair has converged to a given equilibrium when this equilibrium is the majority outcome in the last five rounds. Among the 17 pairs of players in the Partner treatment, 6 have converged to a type 1 leadership equilibrium, 6 to a type 2 leadership equilibrium, 2 to the (X,X) equilibrium, and 3 have not converged to any equilibrium outcome. We focus on the first two types of convergence as they represent the focal equilibria for both types of players. Figure 1 presents the evolution of the proportion of type 1 and type 2 leadership equilibria, separating by types of convergence. Pairs converging to the type 1 leadership equilibrium do so quite early in the game: from round 7 onwards, type 1 leadership is the majority outcome in almost every round. On the other hand, for pairs converging to the type 2 leadership equilibrium, a fiercer competition seems to have prevailed among players, although type 2 leadership is somewhat more prevalent than type 1 leadership during the early phases of the game.

Fig. 1. Convergence



Thus, type 2 players seem to have taken a more proactive role in groups converging to their preferred equilibrium. Although type 1 and type 2 leadership equilibria had roughly the same frequency in the first rounds in both convergence groups, type 2 players in the first group seem to have abandoned early the prospect of driving their opponent to the (Z,Y) equilibrium, while the reverse appears true for the second convergence group. Looking at players' decisions, we indeed find that type 2 players in the second convergence group tend to play their leadership action Y more often than those in the first convergence group even if Y is suboptimal. This confirms the fact that type 2 players in the second convergence group are more insistent on the emergence of their leadership equilibrium than type 2 players in the first group of convergence. The proportion of suboptimal decisions Y (among all decisions) by type 2 players in the second group is 26.11% while it is 12.22% for type 2 players in the first group. This difference is significant (Student test N=6 in each group, p-value=0.051). On the other hand, the difference of proportions of suboptimal decisions Y for type 1 players across groups is weaker: these proportions are 10% and 18.89% for type 1 players in the first and second group respectively, and the difference is not significant (Student test N=6 in each group, p-value=0.188). Thus, while type 1 players do not seem to behave differently in the two convergence groups, we can notice that type 2 players' behavior is different across groups and the tendency of this particular type of player to insist on its leadership action seems to play a role in convergence. In other words, these results support the fact that type 2 players' behavior is particularly likely to be a determinant of coordination. The remainder of this paper will analyze in greater details the precise determinants, prevalence and consequences of the presence of a teaching behavior in our game.

In sum, this section has emphasized some regularities which tend to indicate that players think more strategically than what usual learning models postulate and are likely to teach their opponents. The rest of the paper will be devoted to testing the existence of a teaching strategy in greater details according to a step-by-step approach: in the following section, we will confirm a precondition for teaching, namely the fact that players think strategically and see their opponents as manipulable learners. Then, in the following section, we will check that players play strategically and actually use a teaching strategy. In the final section, we will investigate the consequences of the use of such a strategy on the outcome of a game.

3.2 Players' beliefs

In order to use a teaching strategy, players must first be aware of the learning process of their opponents (i.e. they must be aware of the fact that their opponents use the past history of the game to form their beliefs). It is thus natural, before we examine whether players actually teach their opponents, to first confirm this precondition. In this section, we test whether players anticipate their opponents' reaction to their *own* action or, in other words, if they see their opponents as learners.

Our strategy to examine whether players realize that they can influence their opponents' behavior will be to analyze whether subjects take their own actions into account when forming their beliefs. Explicitly modeling the way a player's actions influence his own beliefs *via* his opponent's anticipated reaction would undoubtedly lead to an intricate model and would require strong behavioral assumptions. Our aim in this section is not to describe accurately those complex interactions, but rather to test whether players think of their opponents as learners who observe others' actions and modify their behavior accordingly. One way to do this would be to examine directly whether players' beliefs vary according to their own action in the previous round. Beliefs, however, may also depend on the history of the opponents' past actions, as postulated by traditional proxies used to describe players' belief formation process.[8] If it is the case, one must filter out the impact of these past actions to avoid spurious correlations between $a_i(t-1)$, player i's own action in the previous round, and his current beliefs.[9] Hence, we test whether player i's

[8] A typical example would be the Bayesian updating rule.

[9] If players partly base their beliefs on the past history of their opponents' play, then this component will be correlated to their beliefs both in the previous and current rounds and to their action in the previous round (because $a_i\,(t-1)$ obviously depends on player i's beliefs in $t-1$), and one would find a positive correlation between current beliefs and previous action *even if* there is no causal effect of the previous action on the current belief. Note that we do not assume that players necessarily base their beliefs on the history of their opponents' play, but rather allow for the possibility of such a belief formation process.

elicited beliefs about the action a, $a = X, Y, Z$, of player j in round t, denoted $B_i^a(t)$, significantly differ from beliefs that would depend only on the history of his opponent's past actions up to round $t-1$, and if this difference can be explained by $a_i(t-1)$.

Adopt the terminology of Nyarko and Schotter (2002) and refer to beliefs based only on the history of the opponents' actions as "empirical" beliefs. Denote them by $\tilde{B}_i^a(t)$ to distinguish from "true" beliefs $B_i^a(t)$ reported by players. Next define $R_i^a(t)$ as the difference between true and empirical beliefs: $R_i^a(t) = B_i^a(t) - \tilde{B}_i^a(t)$. Since $\tilde{B}_i^a(t)$ is conditional on the history of the opponents' past actions, but not on $a_i(t-1)$, then, if $R_i^a(t)$ depends on $a_i(t-1)$, true beliefs $B_i^a(t)$ must also depend on $a_i(t-1)$. In this case, we may conclude that players think that their opponents modify their behavior according to the history of the game and take this into account into their own beliefs, or in other words they realize that their opponents can learn, which would be sufficient to make teaching possible.

We chose to model empirical beliefs, $\tilde{B}_i^a(t)$, as γ-weighted beliefs (Cheung and Friedman, 1997) where the belief held by player i about the probability that player j will play action a in round $t+1$ is given by

$$\tilde{B}_i^a(t+1) = \frac{\mathbb{1}\{a_j(t) = a\} + \sum_{u=1}^{t-1} \gamma^u \mathbb{1}\{a_j(t-u) = a\}}{1 + \sum_{u=1}^{t-1} \gamma^u} \tag{1}$$

where $\mathbb{1}\{a_j(t) = a\}$ equals one if player j has played action a in round t, and zero otherwise. Actions played in a given round are discounted with time at rate $\gamma \in [0, 1]$. When $\gamma = 0$, this model reduces to the Cournot model where the belief held in period $t$about action a is one if the action has been played in round $t-1$ and zero otherwise; when $\gamma = 1$, the model reduces to the fictitious play model where the belief about a given action corresponds to the frequency with which this action has been played since round 1. The Cheung and Friedman model has been found to perform well empirically to explain people's behavior in games.

As can be seen from Equation (1), γ-weighted beliefs in round t are only conditional on the past history of the actions played by *other* players up to round $t-1$ and are thus good candidates for constructing $\tilde{B}_i^a(t)$.

To do so, we estimate the model of equation (1) at the individual level using the method of minimum mean-squared error [10] along the lines of Nyarko and Schotter (2002). We are thus able to compute estimated empirical beliefs $\hat{\tilde{B}}_i^a(t)$

[10] That is, our estimator is the values of the parameter vector that minimize $\sum_{i,t,a} \left(B_i^a(t) - \tilde{B}_i^a(t) \right)^2$.

that can be interpreted as the largest part of the individual's true beliefs $B_i^a(t)$ that can be explained by the past history of the opponents' actions up to round $t-1$ under the Cheung-Friedman hypothesis. For each a, $a = X, Y, Z$, we then compute $\hat{R}_i^a(t)$, the difference between true (or elicited) beliefs and estimated empirical beliefs in round t, and proceed to examine whether these differences vary according to the action taken by the individual in the previous round, $a_i(t-1)$.

If players think their opponents are adaptive learners, then, as argued earlier, the difference between true and empirical beliefs must depend on $a_i(t-1)$. However, the difference can also vary according to the opponents' past behavior, and, more particularly, according to the opponents' propensity to best respond to the previous action in the past.

Moreover, the size of the difference $R_i^a(t)$ will also depend on the value of player i's empirical beliefs since a high empirical belief leaves less room for a large positive $R_i^a(t)$ than low empirical beliefs. Hence, for each a, $a = X, Y, Z$, we run random-effect panel regressionswith heteroskedasticity-robust standard errors of $\hat{R}_i^a(t)$ on a dummy variable $\mathbb{1}\{a_i(t-1) = \tilde{a}\}$ where a is the best response to $\tilde{a}$ (i.e. X for $\hat{R}_i^X(t)$, Z for $\hat{R}_i^Y(t)$, and Y for $\hat{R}_i^Z(t)$), on thecurrent propensity of the opponent(s) to best respond when player i's previous action was $\tilde{a}$ (i.e. $:\frac{\sum_{\tau=3}^{t} \mathbb{1}\{a_j(\tau-1)=a, a_i(\tau-2)=\tilde{a}\}}{\sum_{\tau=3}^{t} \mathbb{1}\{a_j(\tau-1)=a\}})$ [11] and on the estimated empirical belief of player i, $\hat{\hat{B}}_i^a(t)$.

As mentioned above, the precondition we aim to exhibit in this section is that players are aware of the impact of their actions on their opponents' behavior; thus results in Table 6 focus on the estimated coefficients of $\mathbb{1}\{a_i(t-1) = \tilde{a}\}$ for each belief difference $\hat{R}_i^a(t)$ in each treatment, first for all players, then separating by type. [12] [13] In this table, a positive coefficient means that players' true beliefs regarding a given action have a stronger upward bias (or a weaker downward bias) relative to γ-weighted beliefs when the action for which it is a best response has been played in the previous round. In other words, it indicates that player i will put more weight on his opponent best responding to his own previous action $\tilde{a}$ than implied by γ-weighted beliefs. The coefficients are always positive and significant in the Partner treatment. This indicates that players in this treatment, whatever their type, hold a stronger belief about

[11] The proportion is set at zero before round 3 and when the denominator is zero.
[12] Results for the control variables indicate that, as expected and mainly in the Partner treatment, $R_i^a(t)$ tends to be smaller when $\hat{\hat{B}}_i^a(t)$ is high. No clear pattern emerges for this variable in the Stranger treatment. No clear pattern emerges for the proportion of best responses either, the corresponding coefficients often being insignificant.
[13] There are 17 individuals of each type in the Partner treatment and 18 in the Stranger one.

their opponent playing a when they just played $\tilde{a}$ than when they chose another action. Our interpretation of this finding is that players think that their opponent tries to predict future actions from the past history of the game, will consequently put more weight on the probability that $\tilde{a}$ will be played again in future rounds, and are thus likely to best respond a more frequently. When separating by type, all coefficients remain positive and significant, indicating that both types of player take strategic interactions into account when forming their beliefs. Moreover, the size of the coefficients is also rather homogeneous across types, except for R_i^Z, where type 1 players seem to believe that the impact of having played Y in the previous round will have a larger impact than type 2 players do. This could be explained by the fact that (Y,Z) is the equilibrium where the difference in payoffs between types is minimal, while (Z,Y) is the equilibrium where this difference is maximal. A social norm for equality would then make it harder to drive players to (Z,Y) than to (Y,Z). This is consistent with the pattern depicted in Figure 1, which shows that type 2 players have to be rather insistent to drive the game to their preferred equilibrium.

On the contrary, results for the Stranger treatment show that the coefficients are never significant, indicating that players do not anticipate the impact of their actions on their opponents' behavior in this treatment. The difference in results for the two treatments reflects the fact that the matching protocols differ in the ease with which players can influence their opponents.

To summarize, these results indicate that when players know they will face the same opponent in the next round, they anticipate a greater tendency for the opponent to play a best response to what has just been played than implied by the history of their opponent's past actions. In other words, they seem to believe that their opponent partly bases his actions on the past history of the game. Therefore, they realize that their own actions are likely to influence their opponent's behavior, thus confirming a precondition for teaching. Because in the Stranger treatment players know that pairs are rematched in each round, this conclusion does not hold. They have no reason to put a greater probability on their opponents best responding to their past actions, which implies that the differences between true and empirical beliefs do not differ according to their previous action. The essential conclusion that can be inferred from these results is that there is room for teaching in the Partner treatment, but the actual use of teaching depends does not only on the knowledge of the learning process of the opponent, but also directly on the incentives brought by the payoff matrix. Because the latter is asymmetric, we can expect the two types of players to differ in their use of a teaching strategy, despite having the same knowledge of the potential impact of their actions on their opponent's behavior. The next section addresses the actual use of teaching by players.

Table 6
Differences in belief formation

	Partner			Stranger		
	R_i^Y	R_i^Z	R_i^X	R_i^Y	R_i^Z	
All	0.064^{**} (0.021)	0.109^{**} (0.018)	0.093^{**} (0.019)	0.006 (0.019)	0.006 (0.015)	0.008 (0.017)
type 1	0.070^{*} (0.033)	0.109^{**} (0.025)	0.143^{**} (0.029)	0.032 (0.027)	0.016 (0.020)	0.026 (0.022)
type 2	0.081^{**} (0.026)	0.103^{**} (0.026)	$0.048^{\dagger}$ (0.025)	-0.023 (0.024)	-0.016 (0.023)	-0.022 (0.026)

Robust standard errors in parentheses.
Significance levels : $\dagger$: 10%, $*$: 5%, $**$: 1%

3.3 Teaching a learner

If, as we have shown in the previous section, players *think* strategically and are aware of their opponents' learning process, they have an incentive to *play* strategically and teach adaptive players. We are ready, therefore, to proceed to the next step of our analysis which consists of investigating whether players actually use a teaching strategy by choosing actions with poorer short run payoffs but which will modify an adaptive opponent's behavior in a way that might lead to higher payoffs in the longer run.

3.3.1 Over and under responses

A way of testing whether players use a teaching strategy is to examine, along the lines of Ehrblatt et al. (2009), whether their behavior is consistent with their beliefs about their opponents' actions or if they depart from such an immediate payoff maximizing behavior.

Indeed, as noted in Section 3.1, the existence of a teaching strategy implies that players will not necessarily respond best to their current beliefs about their opponent's behavior and that they are likely to take suboptimal decisions. It will be convenient in what follows to give a more precise description of the way players depart from their best response behavior. If players anticipate a higher future payoff by playing an action which is not in their immediate best response set, they might exhibit a stronger tendency to play this action than implied by their current beliefs. We will refer to such a behavior as an 'over response' behavior. Similarly, one could also say that players 'under respond' by exhibiting a weak tendency to play some actions even if they are in their best response set. Teaching would notably consist in over-responding Y and under-responding Z as these two actions support respectively players' best and worse equilibria in terms of payoff. If players use a teaching strategy, one could expect to observe more over-responses Y and under-responses Z in the Partner treatment than in the Stranger treatment because, as mentioned earlier, we can reasonably assume that rematching players in each round makes teaching riskier.[14] Moreover, because type 1 players have weaker incentives to use a teaching strategy, their over-response rate should be lower than for type 2 players.

We now attempt to test these implications of the existence of a teaching strategy. For each round, we first derive the belief-wise best response defined by $argmax_{a\in\{X,Y,Z\}}\, E_i^a\,(t)$ where $E_i^a\,(t) = \sum_{q=X,Y,Z} B_i^q(t).\pi_i(a,q)$ is player i's expected payoff induced by playing action a in round t, with $\pi_i(a,q)$ player i's payoff when he takes action a and his opponent takes action q, and $B_i^q(t)$ player i's (true) beliefs concerning action q in round t. For each action X, Y and Z, we calculate the number of times where the action has been played despite not being a best response (which gives us a measure of over-response) and the number of times it has not been played despite being a best response (which gives us a measure of under-response).

Note, however, that we do not expect players in the Stranger treatment to play exactly as if they were maximizing their immediate expected payoff in each round. Indeed, other considerations such as fairness or some kind of social norm could induce players to depart from a pure payoff-maximizing behavior. Nonetheless, and irrespective of the introduction of social preferences, the incentives brought by the Partner treatment should lead to statistically significant differences in the over/under response behavior between the two treatments.

Table 7 presents in its third (for the Partner treatment) and fourth (for the Stranger one) columns, for each possible action and type of players, the average number of over-responses. The last column presents test statistics and p-values for a t-test of equality between the means presented in the two previous columns.[15] Table 8 presents the same statistics for under-responses. Thus, these Tables allow us to examine the behavior of each type of players *between* treatments.

[14] Note that in the Stranger treatment, we had 18 subjects in each session, so the probability of being rematched with the same opponent in the next period is almost 11 per cent (i.e. relatively low). Because of space constraints, we could not run a session with more than 18 subjects, and thus we could not decrease this percentage. However, it can be noticed that even with 'perfect' rematching, when players are matched only once with a given opponent, one could not completely rule out a priori the use of a teaching strategy by contagion as soon as players' former opponents can be matched with their future opponents. Note also that players never received information that could allow them to identify their opponents.

[15] Since we separate by types of players, the unit of observation is here the individual player. The number of observations is 17 in the Partner treatment and 18 in the Stranger treatment.

Table 7
Over responses

Action	type	Partner	Stranger	t-stat (p-value)
X	type 1	3.294	3.000	0.287 (0.776)
X	type 2	4.294	2.222	1.990 (0.055)
Y	type 1	4.706	5.000	0.212 (0.834)
Y	type 2	6.647	4.000	2.277 (0.032)
Z	type 1	1.824	3.278	1.360 (0.183)
Z	type 2	2.059	2.278	0.218 (0.829)

Table 8
Under responses

Action	type	Partner	Stranger	t-stat (p-value)
X	type 1	0.647	0.111	1.576 (0.133)
X	type 2	0.235	0.056	1.221 (0.236)
Y	type 1	1.765	3.944	1.817 (0.078)
Y	type 2	2.765	2.611	0.142 (0.888)
Z	type 1	7.411	7.222	0.102 (0.920)
Z	type 2	10.000	5.833	2.432 (0.022)

The over-response behavior of type 2 players is significantly stronger in the Partner than in the Stranger treatment for action Y, as well as (although quantitatively less so) for action X. Type 1 players, on the other hand, do not over-respond significantly differently in the two treatments. Turning now to the under-response behavior, Table 8 shows that type 2 players avoid to playing Z when it is a best response significantly more often in the Partner than in the Stranger treatment. Apart from type 1 players for action Y (where we find that those players under-respond Y significantly more in the Stranger treatment), no other significant differences can be found between the under-response behaviors in the two treatments. These results indicate that type 2 players' behavior is particularly consistent with the existence of a teaching strategy in the Partner treatment (i.e. in the treatment which facilitates the emergence of such a strategy). Indeed, in this treatment, type 2 players exhibit a tendency to forego immediate payoff by not best responding to their beliefs, both in order to make their leadership equilibrium emerge (by over-responding

Y more often) and to deter the emergence of their opponent's leadership equilibrium (by under-responding Z more often). On the other hand, the matching protocol has only a little impact on type 1 players' behavior, and this impact is not consistent with a teaching strategy as it leads to a higher tendency for those players to under-respond action Y in the Stranger treatment. In sum, type 2 players seem more prone to teach their opponents when they are given the opportunity to do so more easily thanks to the fixed-matching protocol. In contrast, the impact of the matching protocol on type 1 players' propensity to teach seems to be smaller. Thus, these findings support the fact that when players interact with the same opponent throughout the game, their behavior is altered a way which is particularly consistent with a teaching strategy for type 2 players.

We now turn to the study of the differences *between* types of players *within* each treatment. Table 9 shows the results of paired t-tests (for the Partner treatment) and two-sample t-tests with unequal variance (for the Stranger treatment). Consistent with previous results, we find that no significant differences exist between types in the Stranger treatment. In the Partner treatment, type 2 players are significantly more likely than type 1 players to over-respond action Y, which supports their preferred equilibrium, and to under-respond action Z, which supports their least preferred equilibrium. Unsurprisingly, no significant differences can be found across types regarding over-response towards Z and under-response towards Y because such strategies would contradict the aim of teaching. These results support the fact that players who have more teaching incentives actually make use of this opportunity. Indeed, type 2 players exhibit a stronger tendency to use a teaching strategy than type 1 players. In the Partner treatment, they over-respond more the action supporting their leadership equilibrium and under-respond more the one supporting the leadership equilibrium of their opponent. In the Stranger treatment, when teaching is made riskier, over- and under-responses of each type of players does not differ significantly.

To summarize, the over- (under-)response behavior of players is consistent with the incentives given both by the matching protocol and by the asymmetry of the payoff matrix. On the one hand, the fixed-matching protocol exacerbates players' tendency to over- and under-respond a way which is particularly consistent with a teaching strategy for type 2 players. On the other hand, differences between types of players show that players who have more to gain from the emergence of their leadership equilibrium (i.e. type 2 players) are more prone to forego a higher immediate payoff by playing out of their best response set and tend to favor future coordination into the equilibrium giving them the highest payoff.

Teaching is a long run strategy which is consequently likely to exhibit a dynamic pattern. More precisely, the finite number of periods (known before-

Table 9
Tests across types

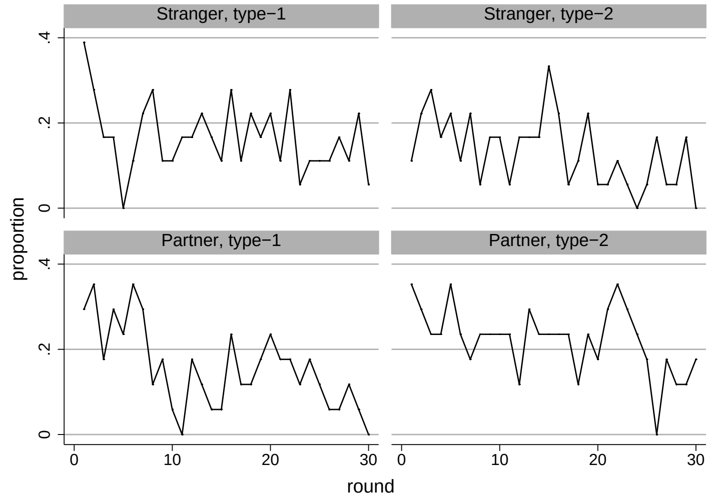
		over response			under response		
Action	treatment	type 1	type 2	t-stat (p-value)	type 1	type 2	t-stat (p-value)
X	Partner	3.294	4.294	1.028 (0.295)	0.647	0.235	1.072 (0.299)
X	Stranger	3.000	2.222	0.823 (0.416)	0.111	0.056	0.589 (0.560)
Y	Partner	4.706	6.647	2.281 (0.036)	1.765	2.765	1.280 (0.219)
Y	Stranger	5.000	4.000	0.773 (0.447)	3.944	2.611	1.221 (0.231)
Z	Partner	1.824	2.059	0.394 (0.699)	7.411	10.000	1.936 (0.071)
Z	Stranger	3.278	2.278	1.058 (0.298)	7.222	5.833	0.856 (0.399)

Note: paired t-test in the Partner treatment and 2-sample t-test in the Stranger treatment

hand by players) should lead to less teaching in later rounds as the gains from teaching become less. Indeed, the gains from a teaching strategy depend on the number of remaining periods during which higher payoffs can offset the losses incurred by not best responding during the initial teaching phase. One way to test this implication would be to examine the dynamics of over-responses towards Y, but other factors unrelated to teaching can also lead to such a decrease. For example, players might also learn during the game how to best respond to their beliefs, thus reducing the number of errors as the game unfolds. Therefore, a decrease in the occurrence of over responses might not necessarily mean less teaching, but it could correspond to less errors. However, one can disentangle the error-correction process from a decrease in genuine teaching by contrasting the evolution of over-responses Y in the two treatments. Indeed, players' propensity to make errors should be irrespective of the matching protocol, but teaching should be sensitive to this. More precisely, if over-responding Y is consistent with a teaching strategy, then the rate at which these over-responses occur should have a stronger decrease in the Partner treatment than in the Stranger treatment since such a decrease should be a typical pattern of teaching, and teaching should be more salient in a fixed-matching protocol. On the other hand, if over-responding Y is an error, then these over-responses should decrease at the same rate across treatments. Moreover, as we do not expect players to use a teaching strategy to make (X,X) or the leadership equilibrium of their opponents emerge, decreases of over-responses X and Z should be mostly due to fewer errors in later rounds, and we should not observe a treatment effect concerning these decreases.

Figure 2 presents the evolution of over-responses Y across rounds, separating by treatment and type. While the rate of over-responses Y decreases in all

Fig. 2. Prevalence of over responses Y



cases, it does so more sharply in the Partner treatment. Student tests with unequal variance (N=34 in the Partner treatment, and 36 in the Stranger treatment) show that the decrease in over responses Y between the first and the last five rounds is significantly higher in the Partner than in the Stranger treatment (p-value=0.082).[16] When examining the evolution of other over responses (X and Z), we find that their decrease is not significantly different between treatments (p-value=0.172).[17] In sum, the dynamics of over responses is consistent with a teaching strategy centered on the emergence of players' leadership equilibria.

3.3.2 Probabilistic choices

3.3.2.1 Model and estimation In this section, we modify learning models by adding a forward-looking component and see if it helps to explain players' behavior. We first present a rationale on how players might evaluate their gains from using a teaching strategy, and then we turn to the empirical specification and estimation results.

Playing against an adaptive opponent implies that every action played will

[16] When considering type 1 and type 2 players separately (N=17 and 18 in the Partner and Stranger treatment, respectively), we find that type 1 players seem to give up teaching more quickly than type 2 players. Their behavior indeed exhibits a significantly stronger decrease in over responses Y in the Partner treatment (p-value=0.098), while the stronger decrease for type 2 players in the Partner treatment is not significant (p-value=0.448).

[17] When type 1 and type 2 players are considered separately, the difference in the evolution of other over-responses remains insignificant (p-values= 0.673 and 0.110 for type 1 and type 2 players respectively).

modify the opponent's future behavior. In order to evaluate the prospects brought by this potential influence, and to decide whether to make use of a teaching strategy, players must first assess the potential gain that would stem from a deviation from their immediate best response.

Playing action k instead of a given reference action r in period t, and hence modifying the opponent's behavior in $t+1$ relative to what it would have been if r had been played, will lead to a difference in the expected payoffs in round $t+1$. Let δ_i^k denote this perceived difference in player i's expected payoffs when he plays k instead of r. Since an adaptive player's beliefs are updated in every round, the influence of past actions on the opponent's beliefs and behavior will decrease with time. Along the lines of Cheung and Friedman (1997)'s model of varying influence of past actions, we assume that players think this influence will decrease geometrically at rate $\beta \in [0, 1)$. If $\beta \to 0$, then the influence of an action played in round t will tend to be only effective for the next round and will have vanished in subsequent rounds. If $\beta \to 1$, then the potential gains from an action played at t will tend to be carried forward unaltered until the end of the game.

Since players must evaluate the overall gain of playing action k in round t, they have to assess the cumulative expected payoff gain, which can be written as

$$\theta_i^k = \sum_{\tau=0}^{T-t-1} \delta_i^k \beta^\tau = \delta_i^k \frac{1-\beta^{T-t}}{1-\beta}. \tag{2}$$

The way we introduce this forward-looking component is similar to the approach of Rutström and Wilcox (2006). It represents to us an elegant and parsimonious way to incorporate teaching. However, the authors use the same rate to describe a decrease in players' perceived influence on the opponent's beliefs (in their set-up, players assume their opponents are γ-weighted learners who compute their beliefs according to Equation (1)) and a corresponding decrease in players' perceived expected payoff gain, represented respectively by γ and β in the present study. We do not follow this intuition here since there is no reason for these two parameters to be necessarily equal.[18] Hence, we introduce this new parameter β to describe the decreasing rate of the influence of past actions, the interpretation of which is given in the above discussion.

If, as our results of section 3.2 suggest, players think that their opponents

[18] Indeed, the way a change in player j's distribution of beliefs will finally impact player i's expected payoff gains is not necessarily that straightforward. Briefly, a change in player j's distribution of beliefs does not necessarily induce the same change in his choice probabilities, which in turn does not necessarily induce the same change in player i's expected payoff gains.

modify their beliefs (and their behavior) according to the history of the game, then they have an incentive to use a teaching strategy. The analysis of Section 3.3.1 has shown that players' behavior was consistent with the existence of such a strategy where players choose actions that do not maximize their expected payoff at the current round, but increase the probability of convergence to an equilibrium they deem preferable. We now turn to a more formal test of this hypothesis by fitting a choice model that includes the cumulative expected payoff bonus δ_i^k.

Our empirical model assumes that player i chooses his actions according to (1) the intertemporal expected payoff from playing a given strategy, consisting of (i) the immediate payoff, and (ii) the cumulative expected future payoff difference induced by playing the action, as defined above by θ_i^k; and according to (2) an intrinsic attraction for the given action, to which the player is attracted for non-pecuniary reasons (such as fairness or another social norms that could influence players' behavior).

Formally, player's i attraction for action a after period t has taken place is denoted $A_i^a(t)$ and can be written

$$A_i^a(t) = \alpha_i^a + \lambda_i \left[E_i^a(t) + \theta_i^a \right]$$

where α_i^a is the intrinsic attraction for action a and λ_i represents the player's 'responsiveness' to his expected (immediate and prospective) payoffs.

Some remarks about the way this model formally extends previous belief-based learning models are worth noting. These usual models frequently postulate that players' attraction for an action at a given time depends linearly on the immediate expected payoff induced by this action and on a "bias parameter". This latter parameter is sometimes viewed as reflecting players' non-pecuniary motives or intrinsic attraction for the action. On the other hand, Battalio et al. (2001) argue that this bias might also reflect players' attempts to drive coordination to their favorite equilibrium, which is the kind of behavior we aim to exhibit in this paper. The way we build players' attractions allows us to separate the two effects supported by these two interpretations. More precisely, usual attractions take the form $\rho_i^a + \lambda_i E_i^a(t)$; here we use a new specification for the bias parameter ρ_i^a, which allow us to identify teaching: $\rho_i^a = \alpha_i^a + \lambda_i \theta_i^a$. By adding θ_i^a, we might indeed identify the effect suggested by Battalio et al. and then α_i^a only reflect players' non-pecuniary motives.

Following Fudenberg and Levine (1998) and many others in the literature, we assume that the probability of a given action being chosen in round t takes a logistic form. To identify such models, one has to define a reference action (action X in our application) and normalize its attraction to 1. Parameters of the components of other actions' attraction must thus be interpreted as the difference with the reference action's parameters: for example α^a is the difference between a's and X's intrinsic attractions; $E_i^a(t)$ is the difference between expected payoff induced by playing a and expected payoff induced by playing X given the (true) beliefs. Probabilistic choices are given by

$$p_i^X(t) = \frac{1}{1+\sum_{q=Y,Z}\exp\left(A_i^q(t)\right)}$$

$$p_i^Y(t) = \frac{\exp\left(A_i^Y(t)\right)}{1+\sum_{q=Y,Z}\exp\left(A_i^q(t)\right)}$$

$$p_i^Z(t) = \frac{\exp\left(A_i^Z(t)\right)}{1+\sum_{q=Y,Z}\exp\left(A_i^q(t)\right)}$$

where $p_i^X(t)$ is the probability that X will be chosen in round t; $p_i^Y(t)$ and $p_i^Z(t)$ are analogously defined.

As Wilcox (2006) has shown, pooled estimation of such learning models can lead to severe bias in the parameters in the presence of heterogeneity in λ. His Monte-Carlo study suggests that a random-coefficient approach, even if the heterogeneity distribution is misspecified, greatly reduces such bias.

We follow this approach and use a random-coefficient model, where λ is assumed to follow either a gamma distribution with shape parameter k and scale parameter θ or a lognormal distribution with parameters μ and σ. Both of these frequently used distributions ensure that λ_i lies on the positive real line.

The final log-likelihood is thus

$$LL = \sum_{i=1}^{N}\sum_{t=1}^{T}\ln\left(\int\left[p_i^X(t)\right]^{\mathbb{1}\{a_i(t)=X\}}\left[p_i^Y(t)\right]^{\mathbb{1}\{a_i(t)=Y\}}\left[p_i^Z(t)\right]^{\mathbb{1}\{a_i(t)=Z\}} f(\lambda)\,\mathrm{d}\lambda\right) \tag{3}$$

where $f(\lambda)$ is the pdf of the gamma or of the lognormal distribution.

The integral in equation (3) is computed using a 32-node Gauss quadrature, and the likelihood is then maximized using standard techniques.

3.3.2.2 Results We estimate the model of equation (3) separately for the Partner and the Stranger treatments. Since the set-up is not symmetric, we have distinguished between type 1 and type 2 players for the parameters that are likely to differ between types of players (intrinsic attractions and prospective payoffs). Other 'psychological' parameters such as β and λ are left equal across types of players since they should not be influenced by payoffs.

Estimated parameters[19] in the Partner treatment (shown in the first two columns of Table 10 for the gamma and lognormal specifications, respectively) are consistent with the existence of a teaching strategy. Moreover, estimated coefficients are rather insensitive to the specification of the distribution of the responsiveness parameter λ. Although type 1 players seem to fear that playing the leadership action Y instead of X might lead to a slightly lower future payoff[20], type 2 players expect a significantly positive future payoff gain from playing Y instead of X. Both types of players anticipate a significant future payoff loss from playing the follower action Z. The ranking of expected future payoff differences resulting from playing Y is consistent with the ranking of payoffs in the underlying equilibria: while type 1 players have little to gain from moving from the (X,X) equilibrium to the (Y,Z) equilibrium, the potential gain is much larger for type 2 players if equilibrium (Z,Y) is reached. However, while the payoff matrix should lead to a lower disincentive for type 2 players to play Z instead of X, estimated δ^Z indicates that the expected payoff gains, although negative, are not significantly different between type 1 and type 2 players in the Partner treatment.

Because players are randomly rematched in each round in the Stranger treatment, incentives for teaching should be weaker in this treatment than in the Partner treatment. Hence, we expect the prospective payoff parameters to be much smaller in absolute value in the Stranger treatment than in the Partner treatment. The last two columns of Table 10 show that type 1 players in the Stranger treatment do not seem to foresee any sizeable payoff gain from playing Y instead of X and a small (and only significant at the 10% level in the gamma specification) loss from playing Z instead of X. Type 2 players expect slightly larger differences in future payoffs, but these expected payoff bonuses are much smaller than in the Partner treatment. Note that some form of "social learning" where players learn about the behavior of the population of opponents (as opposed to their individual opponents in the Partner case) might give rise to a corresponding "social teaching" behavior that could explain the small but significant prospective payoff parameters in the Stranger treatment.

While the impact on the opponents' behavior is unsurprisingly found to be larger in the Partner treatment than in the Stranger treatment, the parameter

[19] Standard deviations in all models have been adjusted for clustering at the individual level.
[20] Possibly because of the 'leaders' warfare' (Y,Y) which is most unfavorable for type 1 players.

Table 10
Estimation results

	Partner		Stranger	
	Gamma	Lognormal	Gamma	Lognormal
Variable	**Coefficient** (Std. Err.)	**Coefficient** (Std. Err.)	**Coefficient** (Std. Err.)	**Coefficient** (Std. Err.)
Intrinsic attractions				
α^Y (type 1)	2.266^{**} (0.649)	1.941^{**} (0.471)	1.410^{**} (0.351)	1.543^{**} (0.281)
α^Y (type 2)	0.015 (0.351)	0.043 (0.577)	1.420^{**} (0.513)	1.644^{**} (0.439)
α^Z (type 1)	1.870^{*} (0.766)	1.478^{**} (0.522)	1.028^{**} (0.256)	1.062^{**} (0.245)
α^Z (type 2)	1.044^{**} (0.336)	1.063^{**} (0.331)	3.011^{**} (0.437)	3.004^{**} (0.440)
Inertia parameter				
β	0.556^{**} (0.106)	0.565^{**} (0.093)	0.858^{**} (0.049)	0.869^{**} (0.029)
Prospective payoffs				
δ^Y (type 1)	$-0.852^{\dagger}$ (0.500)	-0.565 (0.606)	0.097 (0.185)	0.047^{*} (0.023)
δ^Y (type 2)	3.486^{**} (0.674)	3.370^{**} (0.675)	0.643^{*} (0.174)	0.640^{**} (0.138)
δ^Z (type 1)	-1.611^{*} (0.803)	-1.114^{*} (0.480)	$-0.268^{\dagger}$ (0.144)	-0.238^{**} (0.051)
δ^Z (type 2)	-1.546^{**} (0.390)	-1.545^{**} (0.405)	-0.833^{**} (0.264)	-0.604^{**} (0.083)
Heterogeneity parameters				
k	0.378^{**} (0.085)		0.378^{*} (0.150)	
θ	1.585^{*} (0.761)		1.000 (0.717)	
μ		-0.724^{**} (0.064)		-0.598^{**} (0.099)
σ		0.748^{**} (0.036)		1.181^{**} (0.082)
N	1020		1080	
Log-likelihood	-852.961	-850.759	-820.714	-813.433
$\chi^2_{(4)}$	35.67	29.63	52.581	52.28

Significance levels : †: 10% ∗: 5% ∗∗: 1%

β indicates a stronger inertia in the Stranger treatment in comparison with the Partner treatment. Behaviors might indeed evolve faster in a fixed pair of players than in a large population. A reason for this is that players might be quicker in adjusting their beliefs and behavior in a situation where they interact repeatedly with the same opponent compared with a situation where they have to gather information about their current opponent from a population of players.

Finally, the ‘intrinsic attraction’ parameters show that, in both treatments, type 1 players equally favor Y and Z over X and type 2 players favor Z over the two other actions. This might indicate a preference for equality that we will see reflected in the results of the next section.

3.4 Teaching as a coordination device

The literature on learning models generally aims to track players’ behavior without investigating the effective impact of such a behavior on the outcome of the game. Allowing for the use of a teaching strategy might permit us to gain some insights on this question. In other words, do players succeed in teaching their opponent to play a given action? More precisely, applying this reasoning to the present set-up, we can ask the following question: Is equilibrium selection affected when teaching is made riskier?

Our results of the previous sections suggest that type 2 players use this opportunity more than type 1 players do. As a consequence, we expect players to converge to type 2 leadership equilibrium (i.e. to the equilibrium that brings type 2 players the highest payoff) more often in the Partner than in the Stranger treatment.

Figures 3 and 4 graph the evolution of the number of equilibria attained in the Stranger and Partner treatment respectively. Because coordination is easier to achieve in a Partner set-up where players repeatedly interact with the same opponent, differences in the number of equilibria attained might only reflect differences in relative easiness to converge to an equilibrium rather than differences in equilibrium selection due to the existence of a teaching strategy. Thus, in Figures 5 and 6, we rather represent the evolution of the proportion of the various equilibria among the equilibria attained in each round in the Stranger and Partner treatments.

Fig. 3. Number of equilibria, Stranger treatment

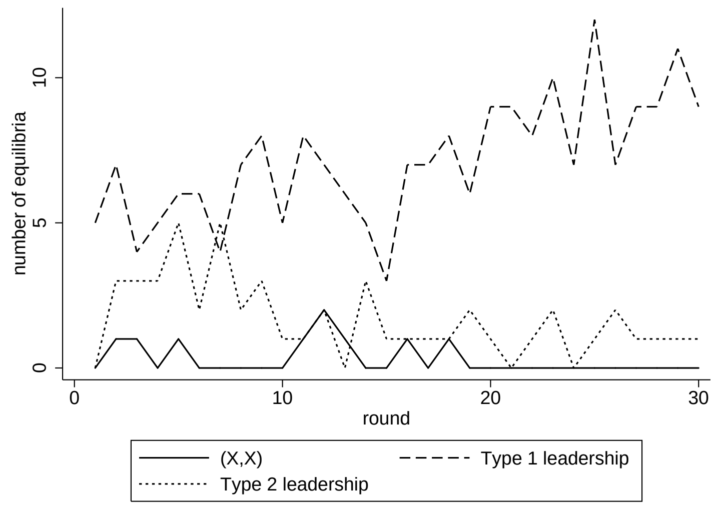


Fig. 4. Number of equilibria, Partner treatment

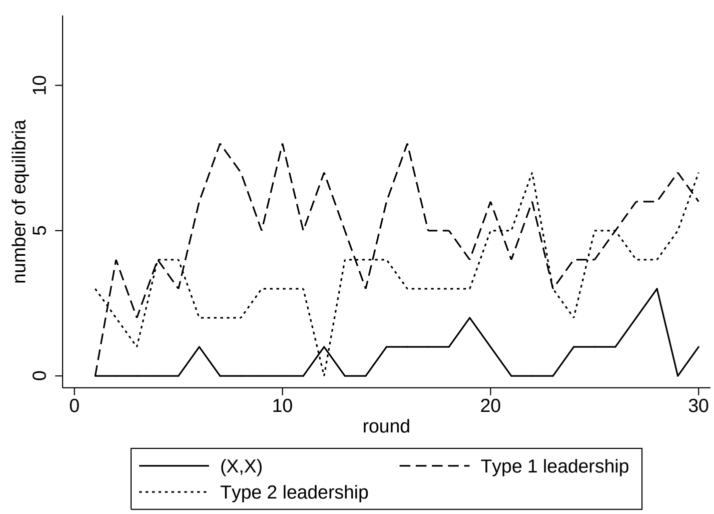


Fig. 5. Proportion of equilibria, Stranger treatment

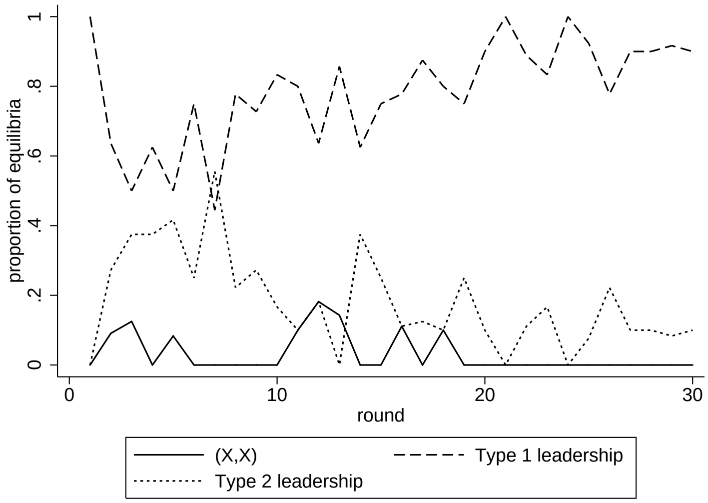
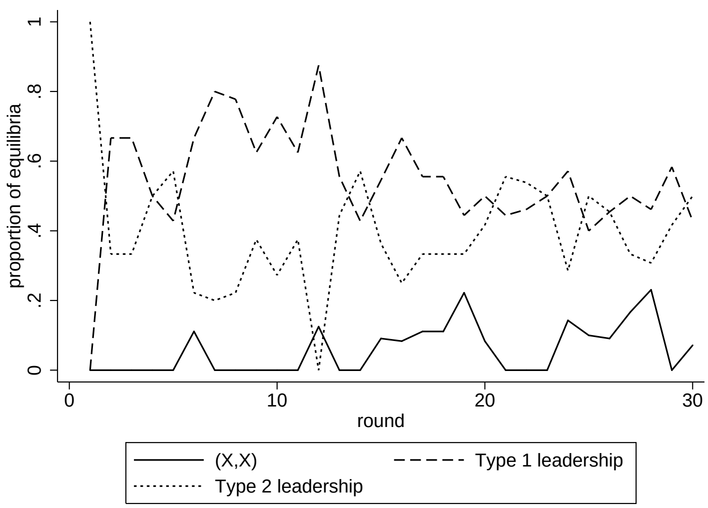


Fig. 6. Proportion of equilibria, Partner treatment



Both sets of Figures show that the number and proportion of type 2 leadership equilibria is larger in the Partner than in the Stranger treatment (50% in the last round in the Partner treatment versus 10% in the Stranger treatment), the reverse being true for type 1 leadership (42.86% in the last round in the Partner case versus 90% in the Stranger case). Moreover, these graphs show a decrease in the emergence of type 2 leadership equilibrium in the Stranger treatment (from an average of 2.7 equilibria, or 29.07% during the first 10 rounds to an average of 1 equilibrium, or 9.6% during the last 10 rounds) and an increase in the Partner treatment (from an average of 2.6 equilibria, or 40.3% during the first 10 rounds, to an average of 4.7 equilibria, or 43.92% during the last 10 rounds), an opposite variation being observed for the emergence of type 1 leadership equilibrium: we find an increase in the emergence of type 1 leadership equilibria in the Stranger treatment (from an average of 5.7 equilibrium, or 67.94% during the first 10 rounds, to an average of 9.1 equilibrium, or 90.4% during the last 10 rounds). Although the number of type 1 equilibria is rather stable in the Partner treatment (from an average of 4.7 in the first 10 rounds to an average of 5.1 in the last 10 rounds), the proportion of type 1 equilibria decreases in this treatment (from an average of 58.6% in the first 10 rounds to an average of 48.05% in the last 10 rounds).[21] The (X,X) equilibrium is also more likely to be attained in the Partner treatment, but the difference between treatments is smaller than for leadership equilibria (7.14% in the Partner treatment versus 0% in the Stranger treatment in the last round).

Table 11 shows the mean count and proportion of each of the 3 equilibria, in both treatments, as well as results from Student tests for the equality of means across treatments.[22]

Table 11
Attained equilibria

Equilibrium	Statistic	Partner	Stranger	t-stat (p-value)
(X,X)	Count	1.059	0.5	1.280 (0.212)
(X,X)	Proportion	0.079	0.037	1.312 (0.200)
Type 1 leadership	Count	8.941	11.888	1.200 (0.239)
Type 1 leadership	Proportion	0.517	0.727	2.075 (0.046)
Type 2 leadership	Count	6.176	2.778	2.197 (0.039)
Type 2 leadership	Proportion	0.404	0.235	1.818 (0.079)

These results show that the matching protocol does significantly change the distribution of attained equilibria, both in proportion and number. Moreover, these changes are consistent with our previous findings that type 2 players are more prone to use a teaching strategy than type 1 players. As a result, the

[21] Note that one could have expected even more important variations in proportions of leadership equilibria in the Partner treatment without the fact that the only equilibria attained in the first round were 8 type 2 leadership equilibria, thus leading to a proportion of 100% of type 2 leadership equilibria and 0% of type 1 leadership equilibria in the first round.

[22] Tests compare attained equilibria at the pair level in the Partner treatment to attained equilibria for type 1 players in the Stranger treatment. Using type 2 players instead of type 1 players does not change the results. There are 17 observations in the Partner treatment and 18 in the Stranger treatment.

number of type 2 leadership equilibria is much higher in the Partner treatment, mainly to the detriment of type 1 leadership equilibria.

Although type 1 leadership predominates in both treatments (possibly because players have in mind a norm of equality, (Y,Z) being the equilibrium where the differences between players' payoffs is minimal [23]), the opportunity of a teaching strategy in the Partner treatment leads to a doubling (both in absolute value and in proportion) of the number of type 2 leadership. Equilibrium selection thus appears to be significantly affected by the matching protocol in a way that is consistent with the use of a teaching strategy by type 2 players, who have the greatest incentives to deviate from a standard payoff-maximizing behavior.

4 Conclusion

Adaptive-learning models have proven to be successful in describing how people behave in games, yet in these models, players look only at the past history of the game to choose their current actions, not paying attention to the fact that these current actions could possibly influence their opponents' behavior in the future. In other words, in these models players do not try to outguess their opponents and consequently do not use a teaching strategy by choosing actions that do not necessarily maximize their immediate expected payoff but might lead to a preferable outcome in the future. Taking this consideration into account might help to track players behavior more accurately and might also allow us to gain some insights in predicting the attained outcome of a game.

This paper has used experimental data to examine whether players use a teaching strategy aimed at modifying their opponents' beliefs and actions in order to reach a preferable outcome. We ran a 'Partner' treatment where players were matched in fixed pairs during the whole game and a 'Stranger' treatment where players were randomly rematched at each period and had thus no particular incentive to 'teach' their opponent. Our results indicate that players indeed use a teaching strategy and suggest that these considerations are relevant to determine the outcome achieved in a game.

First, players' behavior exhibits some patterns which are not consistent with

[23] As noted earlier, this social norm, while not formally tested in this paper, is reflected in the 'intrinsic attraction' parameters α^a of Section 3.3.2. These parameters indicate that, regardless of their beliefs, players have a tendency to favor the (Y,Z) equilibrium, this preference being larger in the Partner treatment than in the Stranger one.

usual approaches designed to track people's behavior in a repeated game but are to be expected from players using a teaching strategy. We indeed found that most suboptimal decisions correspond to very costly deviations, which contradicts the postulates of an error process. On the other hand, these deviations were consistent with a teaching strategy where players aim to make their leadership equilibrium emerge. This regularity appeared more prominent in the Partner treatment where teaching is made easier and for players who have the most to gain from a teaching strategy. Moreover, the latter type of players precisely seems to play a proactive role in coordination.

We then turned to a thorough investigation of the existence of a teaching strategy. We first tested whether players thought of their opponents as belief-learners who base their beliefs on the past history of the game, which is a necessary condition for players to use a teaching strategy. Our results show that players anticipate their opponents' reaction to their own actions and take this reaction into account when forming their own beliefs, thus confirming the precondition that players are aware of the fact that their opponents can learn.

In a second step, we examined whether players actually try to take advantage of this knowledge that they can influence their opponents' beliefs and actions. To do so, we checked whether players' behavior is consistent with the existence of a teaching strategy and whether the matching protocol had an impact on players' tendency to use such a strategy. We found that, especially when the treatment favors the emergence of teaching, players are likely to depart from a best response behavior by choosing the action which supports a preferable outcome and avoiding the action supporting their worst equilibrium payoff. Moreover, this behavior is prominent for players who have the strongest teaching incentives. The dynamics of this over-response behavior is also consistent with teaching, as players tend to teach less in later rounds. We then estimated a logistic model which confirmed this tendency. More precisely, the model suggests that when given the opportunity to teach their opponents, the cumulative expected payoff that players could gain by modifying their opponent's behavior had a statistically significant influence on their propensity to play a given action. Again, the model highlights a greater propensity for type 2 players to base their actions on a teaching strategy.

Finally, we investigated the effective relevance of teaching on equilibrium selection and found that teaching indeed drives coordination significantly so that more tenacious teachers are more likely to make their favorite equilibrium emerge when they can directly teach their opponents.

Acknowledgements

We wish to thank Maxim Frolov, Claude Meidinger, Bruno Rock and participants in various seminars for their help and valuable discussions, and the Centre National de la Recherche Scientifique for financial support. All remaining errors are ours.

References

Arthur, W. B., 1991. Designing economic agents that act like human agents: A behavioral approach to bounded rationality. American Economic Review 81, 353–359.

Battalio, R., Samuelson, L., Van Huyck, J., 2001. Optimization incentives and coordination failure in laboratory stag hunt games. Econometrica 69, 749–764.

Boylan, R. T., El-Gamal, M. A., 1993. Fictitious play: A statistical study of multiple economic experiments. Games and Economic Behavior 5, 205–222.

Camerer, C., Ho, T.-H., 1999. Experience weighted attraction learning in normal-form games. Econometrica 67, 827–874.

Camerer, C., Ho, T.-H., Chong, J.-K., 2002. Sophisticated EWA learning and strategic teaching in repeated games. Journal of Economic Theory 104, 137–188.

Cheung, Y.-W., Friedman, D., 1997. Individual learning in normal form games: Some laboratory results. Games and Economic Behavior 19, 46–76.

Dürsch, P., Kolb, A., Oechssler, J., Schipper, B. C., in press. Rage against the machines: How subjects learn to play against computers, Economic Theory.

Ehrblatt, W. Z., Hyndman, K., Ozbay, E. Y., Schotter, A., 2009. Convergence: An experimental study of teaching and learning in repeated games, mimeo.

Ellison, G., 1997. Learning from personal experience: One rational guy and the justification of myopia. Games and Economic Behavior 19, 180–210.

Fudenberg, D., Levine, D., 1998. Theories of Learning in Games. Cambridge: MIT Press.

Hamilton, J. H., Slutsky, S. M., 1990. Endogenous timing in duopoly games: Stackelberg or Cournot equilibria. Games and Economic Behavior 2, 29–46.

Huck, S., Normann, H-T., Oechlsser, J., 1999. Learning in Cournot oligopoly-An experiment. The Economic Journal 109, C80–C95.

Huck, S., Normann, H-T., Oechlsser, J., 2004. Two are few and four are many: Number effects in experimental oligopoly. Journal of Economic Behavior and Organization 53, 435–446.

Mookherjee, D., Sopher, B., 1994. Learning behavior in an experimental matching pennies game. Games and Economic Behavior 7, 62–91.

Mookherjee, D., Sopher, B., 1997. Learning and decision costs in experimental constant sum games. Games and Economic Behavior 19, 97–132.

Nyarko, Y., Schotter, A., 2002. An experimental study of belief learning using elicited beliefs. Econometrica 70, 971–1006.

Offerman, T., Sonnemans, J., Schram, A., 2001. Expectation Formation in Step-Level Public Good Games. Economic Inquiry 39, 250–269.

Rankin, F. W., Van Huyck, J. B., Battalio, R. C., 2000. Strategic similarity and emergent conventions: Evidence from similar stag hunt games. Games and Economic Behavior 32, 315–337.

Roth, A., Erev, I., 1995. Learning in extensive games: Experimental data and simple dynamic model in the intermediate term. Games and Economic Behavior 8, 164–212.

Rutström, E. E., Wilcox, N. T., 2006. Stated beliefs versus empirical beliefs: A methodological inquiry and experimental test, mimeo.

Schipper, B. C., 2006. Strategic control of myopic best reply in repeated games, mimeo.

Sonnemans, J., Offerman, T., 2001. Is the quadratic scoring rule behaviorally incentive compatible?, mimeo.

Wilcox, N. T., 2006. Theories of learning in games and heterogeneity bias. Econometrica 74, 1271–1292.

Zeiliger, R., 2000. A presentation of regate, internet based software for experimental economics. `http://www.gate.cnrs.fr/~zeiliger/regate/regateintro.ppt`, Lyon: GATE.

Appendix

Instructions (translated from French) [24]

Thank you for participating in this experimental session. During this session, upon the choices you make, you may be able to earn a significant amount of money which will be paid you in private at the end of the experiment. Your identity and those of the other participants will never be disclosed.

This session contains **30 repetitions** (which will be labelled ”rounds” on your screen). Your final payment corresponds to the sum of the payoffs you earn at each repetition. More precisely, during the 30 rounds of this session, you will make points labelled in Unités Monétaires Expérimentales (UME). At the end of this session, your total payment in UME will be converted into Euros at the rate:

200 UME = €1.6

In addition, you will automatically receive a fix amount of **€3** as a ”thank you” payment.

During this session, you will **not** be allowed to communicate with other participants. If you have any questions, please raise the hand and the experimenter will publicly answer.

Subjects were given one of the two following paragraphs according to the treatment they participated.

Type and matching *(Partner treatment)*

At the beginning of the session, you will be attached a ”type”, you can be either of type 1 or of type 2. **Your type will remain the same for the whole session**. Moreover, you will be matched with a pair partner, picked up at random at the beginning of the session **among the participants whose type is different from yours**. For example, if you are of type 1 (resp. type 2), your pair partner will be of type 2 (resp. type 1). **Your pair partner will be the same for the whole session**.

Type and matching *(Stranger treatment)*

At the beginning of the session, you will be attached a ”type”, you can be either of type 1 or of type 2. **Your type will remain the same for the whole session**. **At the beginning of each round**, you will be matched with a pair partner picked at random **among the participants whose type**

[24] The italicized sentences were not displayed to the subjects.

is different from yours. For example, if you are of type 1 (resp. type 2), all your successive pair partners will be of type 2 (resp. type 1). Your pair partners' identities will never be disclosed.

Your decisions

In each round, every participant can choose among **3 decisions: X, Y or Z**. The payoff associated to your decision in a given round depends on your own decision and the decision of your pair partner. These payoffs are presented in Tables 1 or 2 below if you are respectively of type 1 or of type 2.

Prediction of other people's decisions

Prior to choosing a decision in each round, you will be given the opportunity to earn additional money by predicting the decision your pair partner will take in the current round. Thus, at the beginning of each round, you will be asked the following three questions:

- On a scale from 0 to 100, how likely do you think your pair partner will take decision X?

- On a scale from 0 to 100, how likely do you think your pair partner will take decision Y?

- On a scale from 0 to 100, how likely do you think your pair partner will take decision Z?

For each question you have to key in a number superior or equal to 0. The sum of the three numbers you enter has to equal 100.

For example, suppose that you think that there is a 40% chance that your pair partner will take decision X, a 35% chance that your pair partner will take decision Y and a 25% chance that your pair partner will take decision Z. In this case, you will key in 40 in the upper box on the screen and respectively 35 and 25 in the two other boxes. At the end of each round, we will look at the decision actually made by your pair partner and compare his decision to your prediction. We will then pay you for your predictions as follows. Consider the above example: you entered 40% for decision X, 35% for decision Y and 25% for decision Z. Suppose now that your pair partner actually chooses Y. In this case, your payoff for your predictions will be:

$$4[2-(1-0.35)^2-(0.40)^2-(0.25)^2].$$

In other words, you will be given a fixed amount of $4\times 2=8$ points (in UME) from which we will subtract an amount which depends on how inaccurate your predictions were. To do this, when we find out what decision your pair partner has made, we will take the number you assigned to that decision, in this example 35% (or 0.35) on Y, subtract it from 100% (or 1), square it and multiply by 4. Next, we will take the numbers assigned to the decisions not made by your pair partner, in this case the 40% (or 0.40) you assigned to X and the 25% (or 0.25) you assigned to Z, square them and multiply by 4. These three squared numbers will then be subtracted from the 8 points we initially gave you to determine the final payoff associated to your predictions for the current round.

Note that since your predictions are made before you know what your pair partner has actually chosen, the best thing you can do to maximize the expected size of your prediction payoff is to simply state your true beliefs about what you think your pair partner will do. Any other predictions will decrease the amount you can expect to earn as a prediction payoff. Note also that you can not lose points from making predictions but can only earn more points. The worst you can do is predicting that your pair partner will take one particular decision with 100% certainty but it turns out that he actually takes a different decision. In this case, you will earn 0 point. Similarly, the best you can do is to guess correctly and assign 100% to that decision which turns out to be the actual decision chosen. Here, you will keep the whole 8 points amount that was given to you at the beginning of the current round.

In each round, you will have two minutes to enter a correct report. If you make a mistake in a report, i.e. if you enter three numbers which sum is different from 100, or if your report is incomplete, you will be able to retry as many times as you want subject to the fact that you have enough time left to do so. If the available time runs out while you have not entered a correct report, the game continues and you will take your decision but you will not get any payoff for your prediction at the current round.

The computer screens

In each round, you will enter your predictions and take your decisions on different screens represented below.

In the first screen, you will have to report your predictions. You have to enter one number for each decision in the box next to the corresponding question. You will see on your screen the time remaining to report your predictions, both in figures at the upper right of the screen and represented by a saltbox in the middle of the screen. Below the three questions, you have a calculator that automatically provides you, in the box "**Sum**", the sum of the numbers you enter and show you, in the box "**Rest**", 100 minus the sum of the numbers you have already entered so that it would made computations easier while reporting your predictions. You can change your report at any time provided that you have any time left to do so and when your report sounds to you satisfactory, click **OK** to proceed to the next screen to take your decision.

In the next screen, you will have to pick up a decision among the three decisions available. To do so, you have to click on the box corresponding to your choice. You have as time as you want to take your decision.

Once you have reported your predictions and taken your decision, you will get information about the current round. More precisely, you will see recapitulated on a final screen your decision, the decision of your pair partner, your payoff and the payoff of your pair partner associated to your decisions along with the predictions you reported and your prediction payoff. Your predictions, your decisions, the decisions of your partner, and your respective decision payoffs will remain present during the whole session on the bottom of your screen in the table which recapitulates the history of the game by round, so that you will always be able to track what happened in previous round and you will always see which round you are in. Moreover, the last line in the table reminds you of your type so that you could always look at the payoff tables an appropriate way.

Your Final Payment

The payoff associated to your predictions will be in addition of what you will make with your decisions. Your final payment in UME will simply be **the sum of all payoffs** you will make throughout the 30 rounds of this session; it is this total payoff that will be converted into Euros at the above rate.

Decision payoff by type

Table 1. Payoff associated to the decision of type 1 participants. Your payoff and the payoff of your pair partner given your decision and the decision of your pair partner.

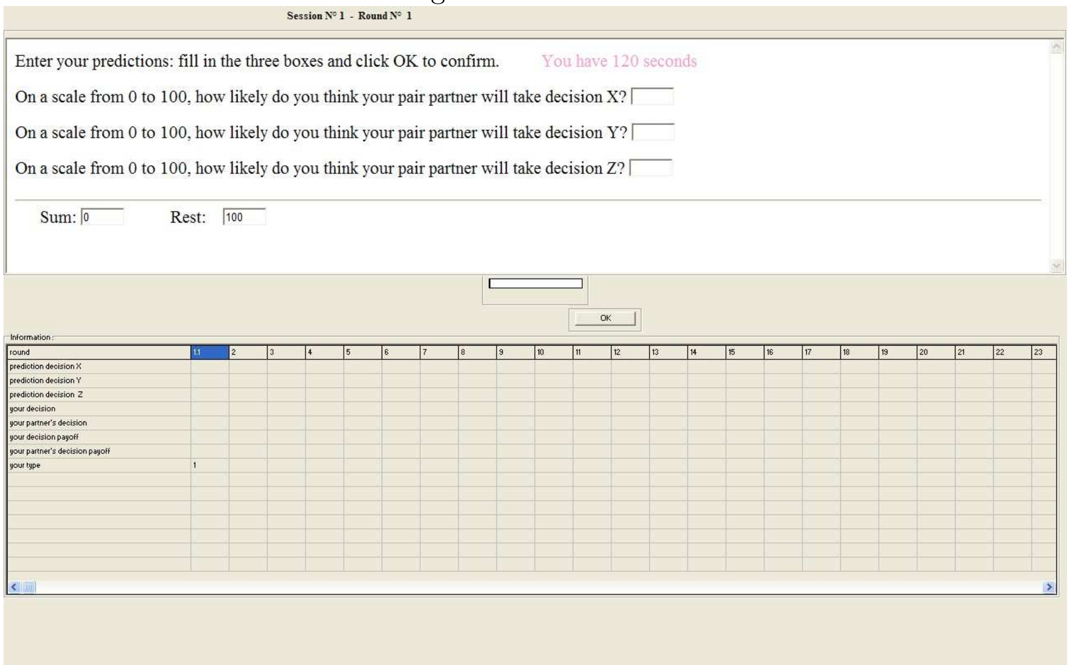
Your decision	Decision of your pair partner	Your payoff	The payoff of your pair partner
X	X	40	52
X	Y	22	46
X	Z	40	52
Y	X	35	40
Y	Y	10	20
Y	Z	44	46
Z	X	40	52
Z	Y	30	60
Z	Z	40	52

Table 2. Payoff associated to the decision of type 2 participants. Your payoff and the payoff of your pair partner given your decision and the decision of your pair partner.

Your decision	Decision of your pair partner	Your payoff	The payoff of your pair partner
X	X	52	40
X	Y	40	35
X	Z	52	40
Y	X	46	22
Y	Y	20	10
Y	Z	60	30
Z	X	52	40
Z	Y	46	44
Z	Z	52	40

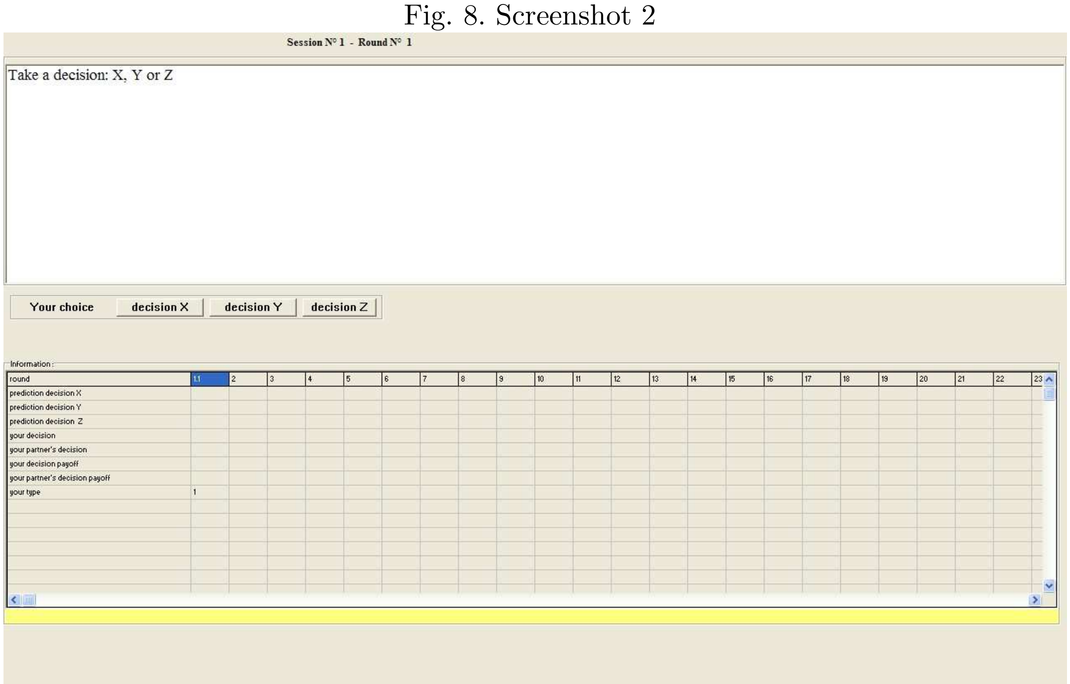
The screen where you will be asked to report your predictions (the last line of the table shows the appropriate type, type 1 in this example)

Fig. 7. Screenshot 1



The screen where you will be asked to take your decision (the last line of the table shows the appropriate type, type 1 in this example)

Fig. 8. Screenshot 2



The following screen was not contained in the instructions given to the subjects. It shows an example of how information were recapitulated at the end of each round.

Fig. 9. Screenshot 3