CONVERGENCE: AN EXPERIMENTAL STUDY OF TEACHING AND LEARNING IN REPEATED GAMES

Kyle Hyndman
Maastricht University and
Southern Methodist University

Erkut Y. Ozbay
University of Maryland

Andrew Schotter
New York University, Center for
Experimental Social Science

Wolf Ze'ev Ehrblatt
Symphony IRI Group

Abstract
Nash equilibrium can be interpreted as a steady state where players hold correct beliefs about the other players' behavior and act rationally. We experimentally examine the process that leads to this steady state. Our results indicate that some players emerge as *teachers*—those subjects who, by their actions, try to influence the beliefs of their opponent and lead the way to a more favorable outcome—and that the presence of teachers appears to facilitate convergence to Nash equilibrium. In addition to our experiments, we examine games, with different properties, from other experiments and show that teaching plays an important role in these games. We also report results from treatments in which teaching is made more difficult. In these treatments, convergence rates go down and any convergence that does occur is delayed. (JEL: C70, C91, D83, D84)

1. Introduction

It goes without saying that Nash equilibrium is an important concept in modern economic analysis. Theoretically speaking, one can interpret a Nash equilibrium as a steady state of a game where players hold correct beliefs about the other players' behavior and best respond to these beliefs. An important question then is how do players achieve this steady state? Is it a belief-led process in which people's

The editor in charge of this paper was Stefano DellaVigna.

Acknowledgments: We would like to thank the editor, Stefano DellaVigna, and four anonymous referees for their detailed and helpful comments. We also acknowledge the helpful comments made by Colin Camerer, Vincent Crawford, Guillaume Fréchette, Ed Hopkins, David Levine, and Chris Parmeter, as well as participants at the C.E.S.S. Experimental Design Workshop and at various seminars and conferences. We also thank Antoine Terracol, Jonathan Vaksmann, Dietmar Fehr, Dorothea Kübler, and David Danz for providing data from their experiments. Financial support from C.E.S.S. and the NSF (SES-0425118) is gratefully acknowledged. A previous version of this paper circulated under the title "Convergence: An Experimental Study".

E-mail: k.hyndman@maastrichtuniversity.nl (Hyndman); ozbay@umd.edu (Ozbay); andrew.schotter@nyu.edu (Schotter); wolly.ehrblatt@gmail.com (Ehrblatt)

Journal of the European Economic Association June 2012 10(3):573–604

DOI: 10.1111/j.1542-4774.2011.01063.x

beliefs converge and then, through best-responding, their actions follow, or do actions converge first and then pull beliefs? In this paper, we experimentally examine this question.

Much of the literature studying this question depicts player behavior as a backward-looking process in which beliefs are formed using historical data on the actions of one's opponent and actions are determined by a deterministic or stochastic best response to these beliefs. In such a world, for convergence to a Nash equilibrium to occur, beliefs must lead actions (since actions are a best response to beliefs). Examples of this kind of model in microeconomics include Fudenberg and Levine (1998), Hopkins (2002), and Camerer and Ho (1999), while representative examples from macroeconomics include Sargent and Marcet (1989), Cho et al. (2002), and Sargent and Cho (2008).

In this paper we question whether convergence to equilibrium can be achieved via such backward looking models, and suggest that, instead, one needs to examine forward-looking models of behavior in order to explain the process of convergence. In such models, some players (who we will call teachers) choose strategies so as to manipulate the future choices of their (possibly myopic) opponent. An early example of such a model is Ellison (1997) who shows that a single rational player interacting in a population of myopic players may be able to move the population to a Nash equilibrium if she is patient enough and if myopic players update quickly enough. More recently, Camerer et al. (2002) incorporate forward-looking behavior by extending their earlier EWA model to include a fraction of sophisticated players who use EWA to forecast the behavior of adaptive players.[1] In their model, a teacher is someone who takes into account the effects of current actions on future behavior. While our results support many of the qualitative features of Camerer et al. (2002),[2] our results suggest a more nuanced view of forward-looking behavior.

In order to study the process of convergence, and the role of teaching in this process, we initially conducted experiments in which subjects played a 3×3 normal form game for 20 periods in fixed pairs and then, after being re-matched, played another 3×3 game also for 20 periods. One of the games was solvable via the iterated elimination of strictly dominated actions, while the other game was not, though both games had a single pure strategy equilibrium. In addition, in both games the pure strategy Nash equilibrium determined payoffs on the Pareto frontier of the set of payoffs available. In our view, such games had the best shot of converging because the question of what to teach is fairly obvious (i.e. the Nash equilibrium). Our results in these treatments demonstrate that many subjects are willing to repeatedly choose their Nash equilibrium action for a number of periods, despite the fact that it is not, at least initially, a best response to their stated beliefs. Such players we will call *teachers* and such behavior we call *teaching*. Therefore, like Camerer et al. (2002), we view teachers as those players who may accept lower short-run payoffs and not best respond

1. Other discussions of sophisticated behavior can be found in the last chapter of Fudenberg and Levine (1998) as well as Crawford (2002) and Conlisk (1993a,b).

2. For example, the presence of forward-looking players facilitates convergence, and in environments where teaching is hard, convergence rates fall.

to current beliefs, in order to influence the opponent's beliefs and *teach* her to play some strategy which will lead to higher long-run payoffs. As we will presently show, about half of the pairs in our baseline treatment converge to the pure strategy Nash equilibrium, and half do not. However, in many of the pairs that do not converge, we demonstrate behavior consistent with a player trying to teach his opponent to play the Nash equilibrium, before ultimately giving up.

This suggests to us, different from Camerer et al. (2002), that teaching is really a higher-order learning process in which the teacher actually learns about how the other player learns. While we do not provide a formal model of teaching here (see Hyndman et al. (2009) who sketch a stylized empirical model of teaching that is also consistent with our results), the basic intuition for the model, and also for our view of teaching, is that a teacher starts off believing that his opponent updates her beliefs very quickly based on past actions. Given such a belief about his opponent, it may be optimal to choose the Nash action, even though it is not currently a best response to stated beliefs. If the teacher's belief about how fast a learner he faces is substantiated, then the pair will converge. However, if the other player proves to be too sluggish in her behavior, after a few periods, teaching may no longer be optimal, in which case the teacher may *give up*.

While focusing on games with a unique pure strategy equilibrium and payoffs on the Pareto frontier makes studying teaching relatively easy, because it is fairly clear that any teaching will be to the Nash equilibrium, it creates a problem in that it is difficult to distinguish between the teaching of Nash equilibrium and the teaching of other focal points, such as Stackelberg equilibrium. To address this issue, we analyzed data from other experiments that follow a similar methodology as our paper. Specifically, we analyze data from Terracol and Vaksmann (2009), Hyndman et al. (2009), and Fehr et al. (2009). In Terracol and Vaksmann (2009) the authors study a game with three pure strategy Nash equilibria which are Pareto incomparable while Hyndman et al. (2009) examined behavior in four games, each with two Pareto rankable pure strategy equilibria and a (common) mixed strategy equilibrium. Fehr et al. (2009) study a game with a unique pure strategy equilibrium which is Pareto dominated by another, nonequilibrium, strategy profile.

These data reinforce our results that teaching is an important factor in the convergence process, but also suggest that what subjects attempt to teach is context dependent, and that their willingness to teach depends on the incentives to do so. For example, despite the presence of a *compromise* equilibrium in which both subjects receive an intermediate payoff, the subjects in Terracol and Vaksmann (2009) vigorously attempt to teach their preferred equilibrium. Because of this conflict, convergence appears to be delayed. In Hyndman et al. (2009), because the two Nash equilibria are Pareto rankable, the question of what to teach is relatively moot—they teach the efficient equilibrium. However, teaching is much more prevalent when the gains to successful teaching are large and the short-run cost of teaching is small, than when teaching incentives are less favorable. Finally, in Fehr et al. (2009) there is again some uncertainty about what should be taught: the Nash equilibrium, or the strategy profile that dominates it (which coincides with one player's Stackelberg equilibrium).

Of those pairs who converge, about half converge to the Nash equilibrium and half converge to the Stackelberg equilibrium. While teachers are present in both groups, the results would seem to indicate that most teaching is done with an aim to reaching the efficient Stackelberg outcome, rather than the inefficient Nash equilibrium.

As a further robustness check on the importance of teaching, we changed aspects of our original games in order to make teaching more difficult. In particular, if teaching facilitates convergence, then by making teaching more difficult, we should see less teaching behavior and also less frequent convergence. We changed our original games in three ways. First, we modified our 3×3 games by adding a strategy for each player. Our conjecture is that by making the game more complex, teaching should be more difficult. Second, we took our original 3×3 games but employed a random matching protocol, rather than the fixed matching of our original experiments. In this case, since subjects are anonymously rematched with a different subject each period, the incentive to engage in long-run behavior is diminished. Finally, we ran a treatment identical to our original 3×3 games with fixed matching, but where players only had access to their own payoffs. In this treatment, because of their limited information about payoffs, subjects could not compute the Nash equilibrium, which makes teaching difficult since one does not know what to teach. As expected, the more difficult we make teaching the less convergence we find. Note, however, that since the convergence rates predicted by the backward looking models should be invariant to all of these changes, our results present further evidence against these models.

In the next section we describe our experimental design and procedures in greater detail. Section 3 provides the results from our baseline treatments where we highlight the important role of teaching in achieving convergence. In Section 4, we re-examine the data generated by other experiments and also provide the results of our two treatments designed to make teaching more difficult. Finally, Section 5 provides some concluding remarks.

2. Experimental Design, Procedures and Definitions

2.1. Experimental Design and Procedures

In order to answer the questions posed in the Introduction, we conducted a number of different experiments, the details of which are given in Table 1. All experiments were run on inexperienced subjects recruited from the undergraduate population at New York University. The experiments were programmed in z-Tree (Fischbacher 2007) and conducted at the Center for Experimental Social Science. Each session typically lasted 1 to 1.5 hours and subjects' mean payoffs were $19.14 across all treatments, not including a $7.00 show-up fee.[3]

In the first treatment, called the All Payoff (AP) Treatment, subjects played one of the games depicted in Figure 1 for 20 periods with a fixed partner and with the payoffs

3. Instructions are available at http://faculty.smu.edu/hyndman/Research/HOSE-Instructions.pdf.

TABLE 1. Summary of experimental treatments.

Treatment	Task	Game(s)	# Subjects	Matching	Payoffs	# Periods
AP	Beliefs/Actions	DSG/nDSG	64	fixed	all	20/20†
$AP_{4\times4}$	Beliefs/Actions	DSG/nDSG	20	fixed	all	20/20†
RM	Beliefs/Actions	DSG	20	random	all	20+40‡
RM	Beliefs/Actions	nDSG	20	random	all	20+40‡
OP	Beliefs/Actions	DSG/nDSG	72	fixed	own	20/20+40♯
NB	Actions	DSG/nDSG	40	fixed	all	20/20†

†Subjects played one game for 20 periods and then (after being rematched) the other game for 20 periods.
‡Subjects played for an initial 20 periods and were then asked to play 40 more periods.
♯Subjects played one game for 20 periods, and then (after being rematched) the other game for 20+40 periods as in ‡.

	$A1$	$A2$	$A3$
$A1$	$51, 30$	$35, 43$	$93, 21$
$A2$	$35, 21$	$25, 16$	$32, 94$
$A3$	$\mathbf{68, 72}$	$45, 69$	$13, 62$

(a) DSG

	$A1$	$A2$	$A3$
$A1$	$12, 83$	$39, 56$	$42, 45$
$A2$	$24, 12$	$12, 42$	$\mathbf{58, 76}$
$A3$	$89, 47$	$33, 94$	$44, 59$

(b) nDSG

FIGURE 1. Games used in the experiments.

of both players visible. They were then randomly rematched and played the other game in Figure 1 for 20 periods. Figure 1(a) depicts a dominance solvable game (DSG), with a unique Nash equilibrium which is in pure strategies. In contrast, Figure 1(b) presents a game which is not dominance solvable (nDSG). This game actually has one pure strategy equilibrium and two mixed strategy equilibria.[4]

The games chosen had Nash payoffs that are on the Pareto frontier, and the Nash equilibrium payoffs were not symmetric. We chose games with these properties for several reasons. First, because our interest was in convergence, we wanted games with a unique pure strategy equilibrium since we assumed that such games would facilitate convergence and avoid the coordination problems inherent in games with multiple equilibria. Second, since the equilibria are on the Pareto frontier, there is no way that subjects could jointly do better for themselves by devising some out-of-equilibrium rotation strategy. In addition, there is no way a subject can teach his or her opponent to play something other than Nash and do better for himself if his opponent is an effective best responder. Of course, these two points also mean that our games are not well-designed to address the question of *where* subjects converge to and *what* players may attempt to teach their opponent (e.g. Stackelberg, Pareto efficiency, etc.). We will use the games run by other experimenters to analyze these questions.

In each period, subjects were asked to make two decisions. The first was to choose the action for that period. The second was to state their beliefs regarding their partner's

4. One is strictly mixed with expected payoffs of (40.5, 58.2). The other puts no weight on the Nash actions and has expected payoffs of (37.05, 69.86). We find no evidence for convergence to either mixed equilibrium.

(a) DSG

	$A1$	$A2$	$A3$	$A4$
$A1$	44, 72	51, 30	35, 43	93, 21
$A2$	36, 18	35, 21	25, 16	32, 94
$A3$	16, 65	**68, 72**	45, 69	13, 62
$A4$	52, 68	58, 67	39, 78	11, 72

(b) nDSG

	$A1$	$A2$	$A3$	$A4$
$A1$	12, 83	39, 56	42, 45	47, 54
$A2$	24, 12	12, 42	**58, 76**	49, 61
$A3$	61, 63	79, 24	21, 13	13, 81
$A4$	89, 47	33, 94	44, 59	31, 11

FIGURE 2. Games used in the 4×4 experiments.

action in that period.[5] The action decision was rewarded according to the relevant game matrix, while the belief reports were rewarded using a quadratic scoring rule (QSR). All payoffs from the action choices and the belief predictions were then summed up to give subjects their final payoff.

In addition to the AP treatment we also ran three others to help substantiate our conclusions. In the $AP_{4\times4}$ treatment, we followed the exact same procedures as in the main treatment but subjects played a 4×4 game as shown in Figure 2. Our belief is that the larger the game, the more complex it is and, therefore, the more difficult should teaching be.

We ran two random matching (RM) sessions (one for each game). In this treatment, subjects were randomly matched each period over the 20-period horizon of the experiment—however, they kept their same role, as row or column player throughout. They were informed only of the outcome of their interaction at the end of each period.[6] In contrast to the AP treatment, subjects only played either the DSG or nDSG game. After the initial 20 periods were completed, we surprised the subjects and told them that the experiment would last for 40 more periods. This was done to check if behavior would change if the horizon was increased. In our third treatment, Own Payoff (OP), we replicated the conditions of the original AP treatment except that subjects are only able to see their own payoffs and *not* the payoffs of their opponent. As with the RM treatment, we surprised subjects after their final 20-period interaction and asked them to play the game for 40 more periods.

As we mentioned in the introduction, these treatments were run in order to better isolate the role of teaching. If teaching is important for convergence and we make it more difficult to teach, then we should see less convergence. In contrast, since backward-looking learning models do not rely on teaching, they should be immune to these changes. For example, if our conjecture is correct, a random matching protocol, by reducing teaching incentives and highlighting myopic play, should decrease the rate of convergence relative to the AP treatment.[7] Similarly for the OP treatment: for most of the popular learning theories (e.g. reinforcement learning, EWA, fictitious play,

5. Eliciting beliefs has become common in experimental economics. See, among others, Terracol and Vaksmann (2009), Costa-Gomes et al. (2001), Haruvy (2002), Costa-Gomes and Weizsäcker (2008), Fehr et al. (2009) Huck and Weizsäcker (2002), and Dufwenberg and Gneezy (2000).

6. This is one of the three ways in which random matching feedback could be given (see Fudenberg and Levine 1998, pp. 4–7, and Hopkins 2002).

7. As Fudenberg and Levine (1998, p. 4) point out, with fixed pairs, subjects may think “that they can ‘teach’ their opponent to play a best response to a particular action by playing that action over and over.”

noisy fictitious play, etc.), the elimination of one's opponent's payoffs should have no impact on behavior or convergence rates.[8]

Finally, there is a strand of the literature which studies whether or not the act of eliciting beliefs changes the behavior of subjects.[9] While the evidence on this front is mixed, Rutström and Wilcox (2009) suggests that eliciting beliefs may encourage more strategic thinking in games with asymmetric payoffs. To address this issue, we conducted one treatment in which subjects only chose actions (i.e. we did not also elicit beliefs). This is our NB treatment.

2.2. Definitions

We first give our definitions of convergence in actions and beliefs. We say that player i has *converged in actions* in period $t_i^a \leq 18$ if, for all $t \geq t_i^a$, player i chooses his Nash equilibrium action. We insist that the player chose the Nash action for at least three consecutive periods before the end of the game so that we don't count players as converging because they randomly chose the Nash action in the final period. If player i converges in actions in period t_i^a, while his opponent converges in period t_j^a, we say that the game converges in period $\tilde{t}^a = \max\{t_i^a, t_j^a\}$.

To describe convergence in beliefs let $b_i(t)$ denote player i's belief about player j's period t action choice. Define the Nash Best-Response Belief Set (NBR_i) for player i as the set of beliefs for which player i's best response is to choose her Nash action. We say that player i has *converged in beliefs* in period t_i^b if, for all $t \geq t_i^b$, $b_i(t) \in NBR_i$.[10]

These definitions are rather strict and make convergence difficult to achieve. We feel they are less ad hoc than other possible definitions since, for any other definition we could use, we would have to call a play path of actions convergent even if at some points along the path subjects would not be playing their Nash actions. In our definition, once a game converges it converges and no deviations are allowed.[11]

3. Results: The AP Treatments

In this section we describe the differences in the behavior of pairs of subjects whose play converged to the Nash equilibrium in the AP treatment and those who did not. Our aim is to demonstrate that teaching facilitates the process of convergence. After

8. For more on how behavior is different when players do not have access to their opponent's payoffs see Partow and Schotter (1993), Mookherjee and Sopher (1994), Costa-Gomes et al. (2001), Fehr et al. (2009) and the references cited therein.

9. For a more detailed overview, see Rutström and Wilcox (2009) and the references cited therein.

10. Recall that nothing in the definition of a pure-strategy Nash equilibrium says that beliefs must be degenerate on one's opponent's Nash action. All that is required is a set of beliefs for which it is a best response to play one's own Nash action.

11. There were three instances of subjects having a high frequency of Nash equilibrium play, with a final period deviation that we labeled as convergent. There was also one pair that we labeled as nonconvergent because one of the pair members only chose his/her Nash action in periods 19 and 20, despite his/her opponent having played the Nash action from period 1.

TABLE 2. Actions, beliefs and best response: "Successful" teaching. (Nondominance solvable game: Nash equilibrium (A2, A3)).

	Row player			Column player	
Period	Action	Best response	Nash played	Action	Best response
1	A3	A3		A3	A2
2	A3	A3		A3	A2
3	A3	A3		A3	A2
4	A3	A2		A3	A2
5	A3	A2		A3	A2
6	A2	A2		A2	A2
7	A3	A3		A3	A2
8	A3	A2		A3	A2
9	A2	A2	√	A3	A3
10	A2	A2	√	A3	A3
11	A2	A2	√	A3	A3
12	A2	A2	√	A3	A3
13	A2	A2	√	A3	A3
14	A2	A2	√	A3	A3
15	A2	A2	√	A3	A3
16	A2	A2	√	A3	A3
17	A2	A2	√	A3	A3
18	A2	A2	√	A3	A3
19	A2	A2	√	A3	A3
20	A2	A2	√	A3	A3

Note: Highlighted cells (in the "Action" column) indicate that the subject played his/her part of the Nash equilibrium in that period. A √ indicates that the Nash equilibrium was the observed outcome in that period.

this descriptive exercise, we present a more formal econometric analysis of a set of hypotheses about teaching and convergence.

3.1. Examples of Successful and Unsuccessful Teaching

3.1.1. A Successful Teaching Episode. To get a flavor for what (successful) teaching looks like, consider Table 2, which shows the history of one convergent pair in the nDSG game. While the time series for other convergent pairs look different, they all tell a similar story: teachers recognize the Nash equilibrium fairly early and choose their part in it repeatedly in an effort to teach. The only question is whether their opponent catches on quickly enough.

For both the row and column players, for each period we note the action chosen as well as the action that was a best response to his/her stated beliefs. Highlighted cells in the "Action" column indicate that the player chose his/her part of the Nash equilibrium. Finally, in the "Nash Played" column, a √ indicates that the Nash equilibrium was the observed outcome for that particular period.

Since this pair of subjects played the nondominance solvable game, the Nash equilibrium is (A2, A3) where the row player chooses A2 and the column player chooses A3. There are many interesting features of this interaction that are illustrative of our point. First, according to our definition, this game converges in period 9. In this pair, the column player is the teacher and starts to play his Nash action in period 1, despite the fact that it is not a best response to his beliefs, and continues to do so until period 6, despite the fact that the row player always chose his non-Nash action, A3. In period 6 he gives up and chooses A2, which is a best response to his beliefs in that period. This might have ruined this pair's chances of convergence except for the fact that, in that period, the row player finally chose his Nash action. Seeing this, the column player resumes teaching by choosing A3, despite the fact that it is still not a best response to do so. Finally, in period 9 the game converges.

3.1.2. A Failed Teaching Episode. The previous example showed that when a teacher is combined with a fast enough learner convergence to the Nash equilibrium may occur.[12] Of course, approximately half of our pairs failed to converge. As we will argue in what follows, a failure to converge is more about beliefs not being updated quickly enough and less about an inability to best respond.

Consider Table 3, which shows the actions and best responses to beliefs for a game that did not converge. This table corresponds to the dominance solvable game, so the Nash equilibrium is (A3, A1). In this pair, we say that the row player is the teacher since she chose A3 in periods 1 through 10 despite the fact that it was a best response to beliefs in only two of those periods. On the other hand, the column player appears to be a particularly dim fellow. Despite the row player choosing A3 for ten consecutive periods his beliefs never actually updated sufficiently so that A1—the Nash action—was a best response. Even more striking is the fact that the column player actually chose the Nash action in periods 4, 7, and 8, yet somehow did not learn (fast enough) that continuing to play the Nash action would be to his benefit. This mistake turns out to be quite costly for the column player. Had he continued to play the Nash action from period 4 onwards, his earnings from the game would have been 37% higher. Finally, after round 10 the row players gives up teaching and basically plays a best response to his beliefs from then on.

3.2. Convergence and Teaching

3.2.1. The 3 × 3 Games. We begin with a thorough analysis of the results from our AP treatment because, of all the environments we considered, it is the most conducive to teaching.[13]

12. Of course, we cannot distinguish between a *fast enough* learner and a *patient enough* teacher, but we will continue to use this terminology.

13. Throughout our analysis, we assume that subjects report their true beliefs. Recent work by Costa-Gomes and Weizsäcker (2008) has called this assumption into question. It is possible that some of our

TABLE 3. Actions, beliefs and best response: A "Failed" teaching episode. (Dominance Solvable Game: Nash Equilibrium (A3, A1)).

	Row player			Column player	
Period	Action	Best response	Nash played	Action	Best response
1	A3	A1		A2	A2
2	A3	A3		A3	A2
3	A3	A1		A3	A2
4	A3	A1	√	A1	A2
5	A3	A1		A3	A2
6	A3	A1		A3	A3
7	A3	A1	√	A1	A2
8	A3	A1	√	A1	A2
9	A3	A3		A3	A2
10	A3	A1		A3	A2
11	A1	A1		A1	A2
12	A1	A1		A1	A2
13	A3	A1		A3	A2
14	A1	A1		A1	A2
15	A1	A1		A3	A2
16	A1	A1		A1	A2
17	A3	A1		A2	A2
18	A1	A1		A1	A2
19	A1	A1		A2	A2
20	A1	A1		A1	A2

Note: Highlighted cells (in the "Action" column) indicate that the subject played his/her part of the Nash equilibrium in that period. A √ indicates that the Nash equilibrium was the observed outcome in that period.

Using the previously given definition of convergence, 17 of 32 pairs converged in the dominance solvable game, while 16 of 32 pairs converged in the nondominance solvable game. As can be seen in Table 4, the frequency of Nash actions over the first ten periods was 56.7% in the DSG game and 45.8% in the nDSG game. These frequencies increased by slightly less than ten percentage points over the last ten periods of the game. As the frequency of Nash actions increased, so too did the best response rate. Indeed, the improvement in the best response rate was more dramatic than was the improvement in the frequency of Nash actions. Also, note that those pairs who converged did so slightly faster in the DSG game than in the nDSG game (5.1 versus 7.2).

Of course, just by looking at convergence rates, we cannot say anything about whether or not teaching was going on. As we have said, teachers are those subjects who are willing to take suboptimal (in the short run) actions in order to influence the beliefs of their opponent with the intention of leading them to a more desirable long run outcome. Therefore, if teaching is going on, and if teachers are teaching the

results are sensitive to this assumption. However, a full investigation of this issue is beyond the scope of the paper.

TABLE 4. Convergence and teaching in the AP treatments.

		3 × 3 Games		4 × 4 Games	
		DSG	nDSG	DSG	nDSG
Percentage of pairs converging	Periods	53.1	50	40	40
Frequency of Nash actions	1–10	56.7	45.8	31.0	35.5
	11–20	64.5	55.3	48.0	44.0
Frequency that action was a best response to stated beliefs	1–10	59.5	59.8	47.0	50.0
	11–20	70.3	76.1	60.0	66.5
Frequency that Nash action chosen, *conditional upon subject not best responding to stated beliefs*	1–10	61.8	46.3	34.9	25.0
	11–20	54.2	41.1	25.0	13.4

Nash equilibrium, then we would expect to see, conditional upon subjects not best responding to their stated beliefs, a high frequency of Nash action choices (i.e. higher than 1/3). Furthermore, since teaching is an investment, we should see it decline in the latter half of the game. These results are on display in Table 4 for both games. As can be seen, over the first ten periods the frequency that the Nash action was chosen conditional upon subjects not best-responding was 61.8% in the DSG game and 46.3% in the nDSG game—both of which are significantly greater than 1/3.[14, 15] Next observe that in both games, such behavior declines in the latter half of the game, and, indeed, for the nDSG game, we cannot reject that the frequency is equal to 1/3.[16] Therefore, it appears that, when subjects do not best respond, very often they choose the Nash action. Such behavior lends supports to our conjecture that some subjects are attempting to teach their opponent the Nash equilibrium.

Next, observe that if some subjects are trying to teach their opponent to play the Nash equilibrium, and if such teaching is ultimately successful, then it should be the case that those subjects within a pair who converged first in actions should have actions which converge *before* their beliefs converge. We turn to this now. In Table 5, we focus on those pairs that converged to the Nash equilibrium and divide them into two subgroups: those subjects who, within their pair, converged to the Nash equilibrium first (Early Convergers) and those who converged second (Late Convergers).[17] For each subgroup, we calculate the average period of convergence in actions as well as convergence in beliefs.

14. Respectively, for DSG_{1-10} and $nDSG_{1-10}$, we have $t_{61} = 6.53$ and $t_{63} = 2.97$. Subscripts denote degrees of freedom. The tests correct for clustering at the subject level.

15. Since row players in the DSG game have a dominated action, one might argue that the threshold should be 1/2. This higher threshold is easily surpassed, with row players choosing the Nash action 69.6% over the first 10 periods when they were not best-responding.

16. Specifically, $t_{40} = 1.19$.

17. There are some pairs who converged in actions simultaneously. These subjects are qualitatively identical to early convergers in the strict sense, and so, have been placed in the "Early" group.

TABLE 5. Average period of convergence in actions and beliefs.

	3 × 3 Games				4 × 4 Games			
	DSG		nDSG		DSG		nDSG	
	Early	Late	Early	Late	Early	Late	Early	Late
Actions convergence period	3.38	7.69	5.67	8.71	4.75	9.5	2.25	5.75
Beliefs convergence period	7.33	8.38	8.83	8.57	8.25	9.25	5.00	7.75
Difference (Beliefs − Actions)	3.95	0.69	3.17	−0.14	3.5	−0.25	2.75	2.00
Number of Subjects	21	13	18	14	4	4	4	4
Mann–Whitney test: Early versus Late Convergers	2.605		3.936		1.78		0.31	

We now argue that, among those pairs of subjects that converged to the equilibrium, the subject who converged first demonstrates behavior consistent with that of a teacher, while the subject who converged later, demonstrates behavior more consistent with that of a follower. Support for this is on display in Table 5 where we show the period of convergence for both actions and beliefs, broken up by whether the subject converged (in actions) first or second. As can be seen, in the dominance solvable game, the difference between belief and action convergence for early convergers is approximately 3.95 periods, while for the nondominance solvable game, it is 3.17 periods. On the other hand, for late convergers, the difference between the period of convergence in actions and beliefs is not distinguishable from zero. Note that for both games, the Mann–Whitney rank sum test allows us to reject the null hypothesis that early convergers and late convergers come from the same population in favor of the alternative that early convergers have beliefs which converge after actions, while late convergers do not.

While obviously one subject is likely to converge in actions before the other, the fact that actions converge *before* beliefs for early convergers, combined with the fact that actions and beliefs converge simultaneously for late convergers, suggests that there is a fundamental difference between these two subgroups of players. In particular, this result is consistent with our claim that early convergers were, in fact, teachers, while late convergers were more passive followers. The reader should note, however, that the classification of subjects as late versus early convergers uses the same variable (the period of convergence in actions) that we also use to construct the variables in Table 5. Therefore, subjects that we classify as early convergers *could* have a higher value of "Difference (Beliefs − Actions)".[18] This potential for bias becomes more likely the earlier in the game that the early converger converged in actions. While we do not have a satisfactory way to deal with this issue, we note that if we exclude from

18. For example, if the period of convergence of actions and beliefs are independent.

our analysis those subjects who converged in the first period, our results are actually strengthened.[19]

Another interesting fact is that teachers appear to be born and not raised. That is, amongst the subjects whom we labeled as early convergers, for the DSG game, 15 out of 21 subjects actually chose their Nash action in the first period as opposed to only 3 of the 13 late convergers. In the nDSG game, the corresponding numbers were 10 out of 18 as opposed to 3 out of 14. For both the DSG and nDSG games, a two-sample proportions test reject the null hypothesis that the frequency of Nash actions in period 1 is the same for early and late convergers: $Z_{DSG} = 2.745$ ($p < 0.01$) and $Z_{nDSG} = 1.95$ ($p = 0.051$), respectively.

3.2.2. The 4 × 4 Games. We ran one session with 20 subjects in which they first played a 4 × 4 dominance solvable game and then, after being re-matched, played a 4 × 4 nondominance solvable game. The experimental procedures were exactly the same as those in the AP treatment. The games played were as in Figure 2. As was the case for the 3 × 3 games, each game had a single pure strategy Nash equilibrium. For the dominance solvable game, this was the *unique* equilibrium, while for the nondominance solvable game there were also two mixed strategy Nash equilibria. Notice also that the games used are identical to the 3 × 3 games but for the addition of another strategy for each player.[20] We now examine whether this additional complexity makes teaching more difficult.

For both the dominance solvable and nondominance solvable game, 4 of the 10 pairs converged to the Nash equilibrium according to our definition, which is slightly lower than the approximately 50% of the pairs that converged in the 3 × 3 games (the difference is not statistically significant). In terms of teaching behavior, the results are similar, though somewhat less pronounced than in the 3 × 3 games. For example, from Table 5 we see that there is a similar difference between early and late convergers in the 4 × 4 games. For both the DSG and nDSG games, those who converge to the Nash equilibrium first have actions which converge strictly before beliefs (on average three periods before). For late convergers, actions and beliefs converge nearly simultaneously for the DSG game, while for the nDSG game, actions actually converge before beliefs. Indeed, in the nDSG game, the difference in the period of convergence of actions and beliefs is very similar for early and late convergers. We do not have an adequate explanation for this, apparently contradictory, finding.

3.2.3. Does Belief Elicitation Influence Behavior? As already mentioned, it has been noted in the literature that the act of eliciting beliefs may encourage players to think more strategically, which may lead to different convergence rates than if we did not elicit beliefs. To examine this issue, we ran our NB (No Beliefs) treatment. Detailed

19. For the DSG game, the Mann–Whitney test statistic becomes 3.709, while for the nDSG game it becomes 3.483.

20. For the dominance solvable game, delete the fourth row and the first column to recover the 3 × 3 game. Similarly, for the nondominance solvable game, delete the third row and fourth column.

TABLE 6. Behavior in the No Beliefs treatment.

	Periods	DSG	nDSG
Percentage of pairs converging		65.0	25.0
Period of convergence		6.46	7.70
Frequency of Nash actions	1–10	56.5	30.8
	11–20	75.0	36.0

results from this treatment are given in Table 6. As can be seen, in the DSG game 65% of the pairs converged to the equilibrium, which is actually slightly *higher* than the convergence rate with belief elicitation, though the difference is not statistically different ($p = 0.4$). On the other hand, for the nDSG game, the convergence rate of 25% is lower than the 50% convergence rate when we elicit beliefs, though this difference is not significant at the 5% level ($p = 0.074$). From this, we conclude that there is no *strong* systematic bias in terms of convergence rates, though our nDSG results suggest there could be an effect of belief elicitation in some games. One can see that there is also no systematic bias in terms of the *period* of convergence to the Nash equilibrium. Comparing Table 5 with Table 6, we see that subjects converged to the Nash equilibrium slightly later in the NB treatment, though for neither game is the difference statistically distinguishable (DSG: $p = 0.2$; nDSG: $p = 0.7$). Given this evidence, we believe that the act of eliciting beliefs did not change behavior in a systematic fashion, though in light of our results for the nDSG game, future studies with a wider array of games and larger sample sizes would seem warranted to determine conclusively if and when belief elicitation does affect behavior.

3.3. Nonconvergence

Until now we have spent our time on convergence and the role that teachers play in fostering it. However, half of our pairs in the AP treatment failed to converge so it would be interesting to discover why. It is our claim that a failure to converge is a failure of belief formation and not of an inability to best respond. In particular, what nonconvergers do wrong is to update too sluggishly. Therefore, when paired with a teacher trying to lead the way to equilibrium, it will take many periods for one's beliefs to enter the Nash Best Response Belief Set and this protracted delay may cause the teacher to give up before convergence occurs.

3.3.1. Failed Teaching. As we have said, teachers are those subjects who are willing to take sub-optimal actions in the short run in order to influence the beliefs of their opponent and drive them to a *better* long-run outcome. Of course, there is no reason to expect that all teachers are ultimately successful. Therefore, if we look at those subjects who did not converge, we would expect to see some instances where one player attempts to teach the other player the Nash equilibrium before ultimately giving

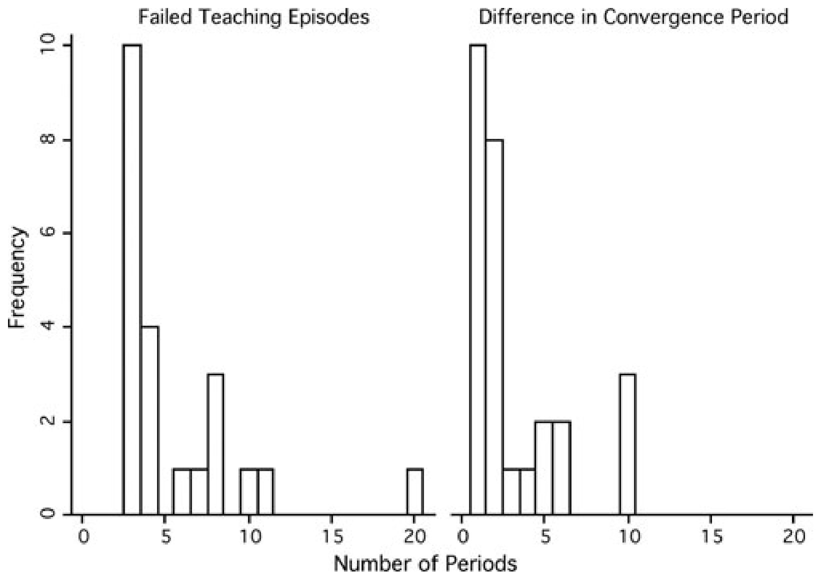


FIGURE 3. Histograms: Length of failed teaching episodes & difference in convergence period between early and late convergers (AP treatment).

up. In Table 3 we gave one such example. Here we seek to understand whether such behavior is prevalent.

Indeed, such behavior is not uncommon amongst those players who did not converge. If we define a *failed teaching episode* as one in which a player chose her Nash action for three or more consecutive periods, then there are a total of 22 such episodes by 20 different players. Of these 22 episodes, 16 of them were true teaching episodes in the sense that the beliefs of the failed teacher were outside the Nash best response set when teaching began. If we are more conservative and insist that the Nash action be chosen for four or more periods, then there are only twelve such instances, and nine of them began when the player was not best responding to her beliefs. The left panel of Figure 3 plots a histogram of the length of failed teaching episodes. As can be seen, most such episodes were only three or four periods in duration, though some were much longer, suggesting that these players were extremely patient. The average length of a failed teaching episode was 5.64 periods.

Our earlier analysis showed that among the pairs that converged, one of the players took the role of a teacher. The current analysis shows that teachers are present even in pairs that do not converge. Therefore, one important factor in determining whether a pair converges would seem to be the *quickness* of the follower. Among the 33 pairs that converged to the Nash equilibrium, 27 pairs had one subject converge strictly before the other. The right panel of Figure 3 shows the difference in the period of convergence between the early and the late converger for each of these 27 pairs. As can be seen, for 19 of 27 pairs, the difference in convergence period is three or fewer periods and the average difference in convergence period is 3.15 periods. Formally a t-test rejects the hypothesis that the length of failed teaching episodes is identical to the difference in convergence period ($t_{47} = 2.48, p = 0.017$).[21] In other words, when a teacher is paired with a follower capable of learning quickly it is unlikely that the teacher will have to teach for more than three periods before convergence occurs. The fact that some teaching episodes lasted substantially more than three periods suggests that excessive sluggishness is an important part of the explanation for nonconvergence.

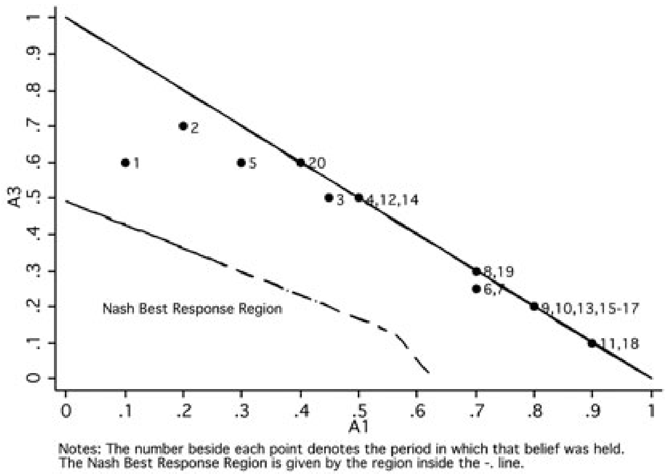
REMARK 1. Before moving on, we note that there were a number of instances in which players took the same (non-Nash) action repeatedly. However, what is striking about such behavior is that it does not appear that there was an underlying teaching motive. In particular, of the 50 instances in which a player chose the same non-Nash action for 4 or more consecutive periods, in 36 of them the action was initially a best response to their stated beliefs. In contrast, in only 3 of the 12 instances in which the Nash action was chosen for four or more consecutive periods was it initially a best response to stated beliefs. This suggests that there is actually something special about the Nash action to a certain subset of players which leads them to choose it repeatedly despite it not being a best response.

3.3.2. Hypotheses. To demonstrate that those subjects who do not converge do so largely because of a failure of beliefs to update quickly enough, we must show two things. First, that there is no difference in the frequency of best response between convergent and nonconvergent players. Second, we must show that those subjects who do not converge do, in fact, update their beliefs more sluggishly. We consider the former first and state it formally as follows.

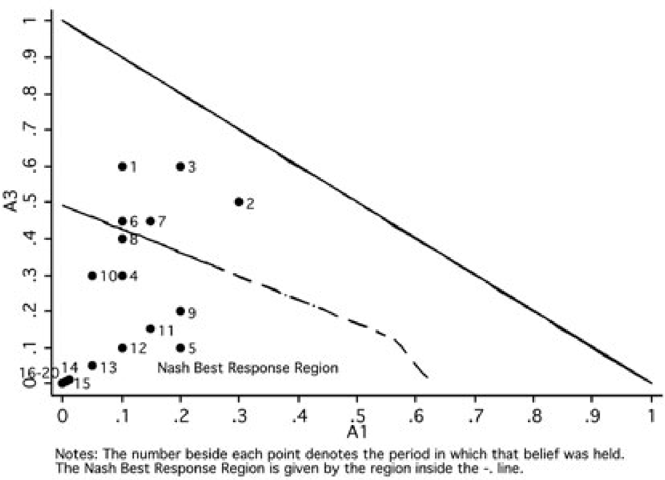
NONCONVERGENCE HYPOTHESIS. *Subjects who do not converge to Nash equilibrium best respond with at least as great a frequency as do those subjects who do converge to Nash equilibrium.*

As evidence in favor of this hypothesis, note that in the DSG game, those subjects who converged best responded 45.1% of the time (up to and including the period of convergence), while those subjects who did not converge actually best responded 52% of the time. Using a two-sample proportions test, we are unable to reject the hypothesis that convergent and nonconvergent players have the same best response rate ($Z_{DSG} = 1.56$, $p = 0.118$). For the nDSG game, those subjects who converged best responded 48% of the time, while those who did not converge did so 60% of the time. In this game, there is a statistically significant difference between convergent and nonconvergent players, but it actually goes is the *opposite* direction. That is, subjects who do not converge actually have a higher best response rate ($Z_{nDSG} = 3.13$, $p < 0.01$). Although we cannot conclusively say that an inability to best respond does not contribute to nonconvergence, the fact that subjects who do not converge best responded with at least as high a frequency as those who did converge, suggests that the primary explanation for nonconvergence lies elsewhere.

21. As can be seen in Figure 3, one subject played his/her Nash action for all 20 rounds. Even if we exclude this outlier, the difference is still statistically significant ($p = 0.03$).



(a) A nonconverging player



(b) A converging player

FIGURE 4. Sequence of beliefs for a nonconverging player and a converging player.

We will go into more detail about this in the next section but before we do let us pause and take a look at the beliefs of subjects who did and did not converge. Consider Figure 4 where we present the simplex of beliefs and the time paths of beliefs of two subjects—one that *did not* converge to the Nash equilibrium, while the other did converge to the Nash equilibrium. The subjects depicted both played the nDSG game and were in the role of the column player. In both subfigures, the point (0,0) represents the case in which a player holds degenerate beliefs that her opponent will play his Nash strategy. The beliefs on the two non-Nash actions are then given by a point in the (x, y) plane and the area enclosed by the dashed line represents the Nash Best Response Set. That is, if beliefs lie inside this set, it is a best response for the player to choose *her*

Nash action. The numbers beside each point denote the period in which the subject held those particular beliefs.

As can be seen, for the player who did not converge, her beliefs actually *never* entered the Nash Best Response Set. In other words, if subjects are capable of best responding and best respond to the beliefs we elicited, then this is clear proof that failure to converge is a result of a failure in beliefs and not in actions. In contrast, the bottom panel of Figure 4 shows the sequence of beliefs for a typical subject that did converge. Here, after some initial periods outside the Nash Best Response Set, the beliefs of the subjects enter the set and very shortly become degenerate.

The subjects depicted in Figure 4 are the rule and not the exception: there is a dramatic difference in the frequency with which beliefs enter the Nash Best Response Set when comparing pairs that converged to those that did not. For convergent pairs their beliefs were in the Nash Best Response Set 79.7% and 68.0% of the time in the DSG and nDSG games, respectively, while for nonconvergent pairs these same percentages were 18.7% and 12.5%, respectively. This is strong evidence that the beliefs of those players who do not converge to Nash equilibrium spent very little time inside the Nash Best Response Set.

3.4. A Formal Analysis of Convergence and Nonconvergence

The previous results indicate that the belief formation process is a key element in whether player pairs converge. What appears to be important for convergence is a teacher along with a follower who updates quickly. In order to test this hypothesis, we must first define a metric for how quickly subjects update their beliefs.

The simplest metric one can use for measuring how quickly subjects update their beliefs upon receiving new information is due to Cheung and Friedman (1997). In that paper, the authors assume that subjects form beliefs based on historical data with geometrically declining weights. That is, more recent information is weighted more heavily than is older information. More formally, we denote by $\Gamma_i^k(t+1)$ play i's belief that his opponent will choose strategy k in period $t+1$, and we write this as

$$\Gamma_i^k(t+1) = \frac{1_t\left(a_k^j\right) + \sum_{u=1}^{t-1} \gamma^u 1_{t-u}\left(a_j^k\right)}{1 + \sum_{u=1}^{t-1} \gamma^u}. \tag{1}$$

Here $1_t(a_k^j)$ is an indicator function which takes on the value of 1 when j plays action k in period t and 0 otherwise.

The lone parameter, γ, captures the rate at which new information is incorporated into a player's belief. Specifically, when $\gamma = 0$, player i believes that his opponent will choose the same action in period $t+1$ as she did in period t. Such extreme beliefs are called *Cournot beliefs*: a subject with such beliefs discards all but the most recent observation. At the other extreme, when $\gamma = 1$, the beliefs about any given strategy are simply given by the empirical frequency with which that strategy has been chosen in the past and each observation is given equal weight. These beliefs are called *fictitious-play beliefs*. Therefore, the closer γ is to 1, the more *slowly* do beliefs respond to new information, while the closer γ is to 0, the more *quickly* do beliefs respond to new information.

To explain the relationship between the γ and pair convergence consider a teacher playing the Nash equilibrium for a couple of periods. If her opponent has a relatively low γ, her opponent's beliefs will rapidly enter the best response set. If her opponent also best responds to his beliefs, then the game will converge. On the other hand, if the teacher's opponent has a relatively high γ, then his beliefs will update only very slowly in response to teaching and may not enter the Nash Best Response Set before the teacher gives up teaching. In such a case the game will not converge. Since many games, including those that do not converge, have teachers, one might conjecture that what separates successful from unsuccessful teaching is the γ of the follower in the pair. If the follower has a low γ and the teacher is persistent, then we expect convergence while if the follower has a high γ, then we expect convergence to be less likely. These facts allow us to state our first convergence hypothesis:

CONVERGENCE HYPOTHESIS 1A—THE γ DISTRIBUTION. *The distribution of the γ for nonconvergers stochastically dominates the distribution of the γ for followers (late convergers).*

To estimate the γ used by each subject we take the sequence of elicited beliefs $\{b_{i,t}\}_{t=1}^{20}$ for each player i over the 20 periods of the experiment and compare it to what that sequence would have been if the subject formed their beliefs using the Cheung–Friedman belief model which would produce, for a given γ, a sequence of beliefs denoted as $\{b_{i,t}(\gamma)\}_{t=1}^{20}$. We estimate γ by searching for that γ that minimizes the sum of squared prediction errors. That is, given a sequence of choices, $\{b_{i,t}\}_{t=1}^{20}$ we find the $\gamma \in [0, 1]$ that minimizes

$$SSE(\gamma) = \sum_{t=1}^{20}\sum_{k=1}^{3}\left(b_{i,t}^{k} - b_{i,t}^{k}(\gamma)\right)^2, \tag{2}$$

where we sum over all periods $t = 1, \ldots, 20$ and all three possible actions $k = 1, 2, 3$.[22]

In Figure 5, we present the cumulative distributions of our estimates of γ for each subject. The results are separated by game (DSG and nDSG) and also by players' status as either a nonconverger or a follower (late converger). First observe that, for both games, when comparing the distributions of γ for followers and nonconvergers, there is a clear pattern of first-order stochastic dominance. Indeed, the distribution of γ is skewed towards much higher values for nonconvergers. Moreover, for both the DSG and the nDSG game, there is a mass of subjects with γ estimated to be 1. What looks like first-order stochastic dominance in the figure is supported statistically, for both games, via one-sided Kolmogorov–Smirnov tests, which are reported in Table 7.

22. To obtain our estimates, we used the Differential Evolution optimization procedure, as implemented in MATLAB, proposed by Storn and Price (1997).

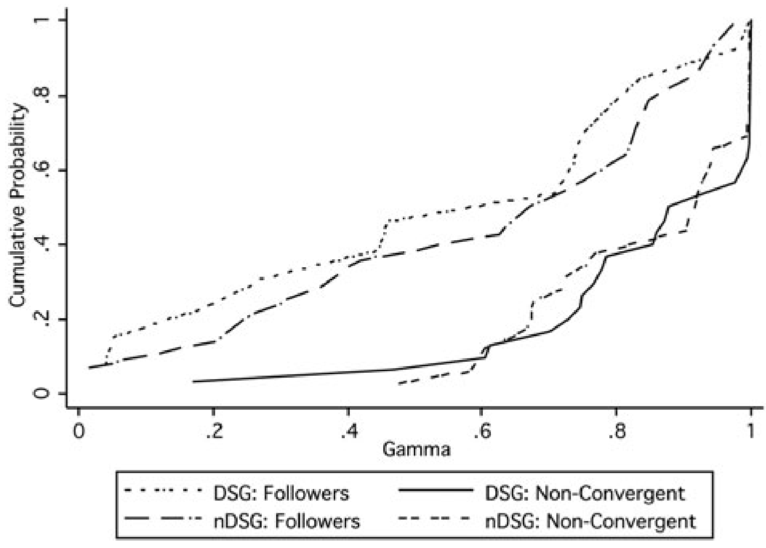


FIGURE 5. Empirical distributions of γ: AP treatment.

TABLE 7. One-sided hypothesis tests for the estimates of γ.

		DSG Game		nDSG Game	
		Nonconvergers	Followers	Nonconvergers	Followers
Test of Means	μ	0.848	0.556	0.848	0.617
	t	3.57^{***}		3.33^{***}	
Mann–Whitney	z	2.91^{***}		2.71^{***}	
Kolmogorov–Smirnov	D	0.480^{***}		0.388^{**}	

*** significant at 1%; ** significant at 5%. Reported significance levels are based on **one-sided** alternative hypotheses.

As can be seen, for the DSG game, we can reject the hypothesis that the distributions of γ are the same for followers and nonconvergers in favor of the alternative that the distribution of γ for nonconvergers first-order stochastically dominates the distribution of γ for followers at the 1% level. For the nDSG game, we reach the same conclusion, though the significance level is only 5%. Table 7 also reports test statistics for a *t*-test of means, as well as the Mann–Whitney test. In all cases, we easily reject the respective null hypotheses that the populations are the same. We take this as substantial evidence in favor of Convergence Hypothesis 1A.

While we prefer to focus on individual level data, we can also conduct a similar exercise pooling across players and estimate γ jointly with the subjects' stochastic best response precision. Since we also argue that a failure to converge is *not* a failure of best responding, we estimate a model of γ-weighted beliefs with stochastic best response. More precisely, using the beliefs defined by (1), if we define the expected utility of choosing action k in period t as $\mathbb{E}_t[\pi(a_k, a_{-i})] + \epsilon_k$, for $k = 1, 2, 3$, where ϵ_k has a

Type I extreme value distribution, with ϵ_i and ϵ_j for $i \neq j$ independently distributed, then we can define the probability that action k will be chosen in period t as

$$\Pr[A_t = k] = \frac{\exp(\lambda \mathbb{E}_t \pi[(a_k, a_{-i})])}{\sum_j \exp(\lambda \mathbb{E}_t[\pi(a_j, a_{-i})])}, \tag{3}$$

where λ measures the precision with which this player best responds.

Our convergence hypothesis can then be restated and expanded to include both γ and λ. Before stating our hypothesis or discussing results, it is important to make clear our purpose with this exercise. Our underlying theme throughout the paper is that two features help convergence. First, the presence of a player who understands the Nash equilibrium and tries to influence the beliefs of his/her opponent, even if it means choosing actions which are not a best response to static beliefs. Such players we have called teachers. Second, the opponent of the teacher must be a sufficiently fast learner for, if not, the teacher will give up and return to static best responses before the game has converged. Our previous analysis has shown that there are teachers present in virtually all games that converged and also in many of the games that did not converge. Therefore, if our conjectures are correct, then what separates those who converge from those who do not is the speed at which beliefs are updated. Thus, it makes sense for us to examine the behavior of two groups: followers (late convergers) and nonconvergers. In order to be consistent with our conjectures, we would expect to find that γ is lower for followers than for nonconvergers (i.e. followers update their beliefs more quickly than nonconvergers) and that λ is the same for both groups (i.e. both followers and nonconvergers best respond equally as well). We state this formally as follows.

CONVERGENCE HYPOTHESIS 1B—THE JOINT γ/λ HYPOTHESIS. *The estimate, γ^{FOL}, for followers should be less than the estimate, γ^{NC}, for nonconvergent players. Moreover, the estimated λ should not be different.*

We test this hypothesis on the pooled data of all subjects in each treatment. However, there is some question about what the appropriate sample to use in the estimation is. After a pattern of convergence has been well established, virtually all subjects are best responding to their stated beliefs. Therefore, when comparing λ between convergent and nonconvergent players, we feel that it does not make sense to include a lot of post-convergence periods. Therefore, for this comparison, we estimate the model using only data up to and including *two periods* after convergence.[23] On the other hand, the dependence on history of beliefs, and hence on γ, still seems important, even after convergence has occurred. Therefore, for this comparison, we will make use of the full sample.

The reader can see that Table 8 lends support to Convergence Hypothesis 1B. Pooling the data across both followers and nonconvergers, we are able to conduct

23. Although two periods post-convergence is somewhat arbitrary, we feel that it gives subjects time to recognize that convergence has been achieved, but is still short enough so that the estimates of λ are not inflated by a lot of post-convergence best responses. Whether we use data up to convergence or two periods post-convergence, the results do not qualitatively differ.

TABLE 8. Estimates of γ and λ: γ-weighted beliefs with stochastic BR.

	DSG (FOL)			nDSG (FOL)		
	conv. + 2	all data	DSG (NC)	conv. + 2	all data	nDSG (NC)
λ	0.043 (0.0091)	0.098 (0.0269)	0.033 (0.0036)	0.051 (0.0111)	0.096 (0.0134)	0.048 (0.0037)
γ	0.423 (0.1210)	0.503 (0.0876)	0.913 (0.0788)	0.497 (0.0799)	0.608 (0.0550)	0.727 (0.0567)
n	126	260	600	151	280	640
LL	–126.41	–216.15	–623.9	–137.91	–187.44	–620.28

Notes: Standard Errors are generated via a jack-knife procedure: For each of 150 replications, we randomly drew (without replacement) a sample of approximately 70% of the players and estimated the model for γ and λ. NC: nonconvergent pairs; FOL: followers.

a series of likelihood ratio tests on the estimated parameters. First, compare the estimates of λ for followers and nonconvergers. For the DSG game, a likelihood ratio test gives a test statistic of $\chi_1^2 = 0.925$, while for the nDSG game, the same test statistic is 0.075. Therefore, in both games, we see that the estimates λ are statistically indistinguishable.[24]

Now consider the estimates for γ. First observe that, whether we use the full sample or only the restricted sample, for both games the estimate of γ for followers is less than the estimate for nonconvergers. In terms of statistical significance, using the full sample, for the DSG game, we have a likelihood ratio test statistic of 8.02, which is highly significant. In contrast, for the nDSG game, the same test statistic is only 2.22, which is not significant at the 5% level. Interestingly, if we use only the restricted sample, then the difference between nonconvergers and followers is also significant ($\chi_1^2 = 4.05$) at the 5% level. This would seem to be due to the fact that γ is actually smaller when we use the restricted sample, though we do not have an adequate explanation for why this would be so.

4. Robustness: Other Treatments and Experiments

4.1. Games Used in Other Experiments

We now investigate the robustness of our findings, and show that our results are transferable to other environments, by using data from experiments run by other investigators using an experimental design similar to ours but employed on games with different structures. We first present the games we will discuss and then analyze them using our teaching hypothesis.

24. Using all the data to compare the estimates of λ, the λ's of followers and nonconvergers are statistically different. Specifically, for the DSG and nDSG games, respectively, $\chi_1^2(DSG) = 40.04$ and $\chi_1^2(nDSG) = 26.51$.

	X	Y	Z
X	**40, 52**	22, 46	40, 52
Y	35, 40	10, 20	**44, 46**
Z	40, 52	**30, 60**	40, 52

FIGURE 6. The Terracol and Vaksmann game.

(HL)

	X	Y
X	**40, 45**	8, 37
Y	39, 0	**12, 32**

(HH)

	X	Y
X	**40, 45**	0, 37
Y	37, 0	**12, 32**

(LL)

	X	Y
X	**20, 45**	8, 37
Y	19, 0	**12, 32**

(LH)

	X	Y
X	**20, 45**	0, 37
Y	17, 0	**12, 32**

FIGURE 7. Games From Hyndman et al. (2009)

Terracol and Vaksmann (2009). In a closely related paper, Terracol and Vaksmann (2009) had subjects play the game in Figure 6 for 30 rounds in fixed pairs. This game has three pure strategy equilibria, which are marked in bold. The equilibrium at (X, X) is in weakly dominated strategies; however, this equilibrium can be seen as a kind of compromise since both players receive their second highest equilibrium payoff. Hence, this game is fundamentally different from the 3×3 games used in our experiment.

Hyndman et al. (2009). In a follow-up paper to ours and Terracol and Vaksmann (2009), Hyndman et al. (2009) study the incentives that subjects have to teach their opponent to play a particular Nash equilibrium. They study four different 2×2 coordination games, each with two Pareto rankable equilibria and a mixed strategy equilibrium. The games are on display in Figure 7. The incentives for the row player to teach are varied on two dimensions, while those of the column player are held constant. The two dimensions studied are the so-called *teaching premium* and the *teaching cost*. The teaching premium represents the gain to a player by moving from the inefficient to the efficient equilibrium, while the teaching cost captures the loss to a player who chooses action X even though it is not a best-response. In Figure 7, the first letter below each game indicates whether the teaching premium was high or low, while the second letter indicates whether the teaching cost was high or low. The authors' hypothesis is that the prevalence of teaching should be increasing in the teaching premium and decreasing in the teaching cost.

Fehr et al. (2009). The authors use a similar design to study the evolution of strategic behavior in a repeated game, which is given in Figure 8. Notice that this game has a unique pure strategy Nash equilibrium at (X, X), which is attainable through the iterated deletion of strictly dominated actions. Unlike our dominance solvable game,

	X	Y	Z
X	$\mathbf{78, 68}$	$72, 23$	$12, 20$
Y	$67, 52$	$59, 63$	$78, 49$
Z	$21, 11$	$62, 89$	$89, 78$

FIGURE 8. The Fehr et al. game

TABLE 9. Results from other experiments: convergence.

	Periods	TV	HTV[‡] HL	HH	LL	LH	FKD (X, X)	$(Z, Z)^{\dagger}$
Percentage of pairs converging	(20 Period)	17.6	52.9	37.5	42.1	26.7	25.9	22.2
Percentage of pairs converging	(30 Period)	52.9	NA	NA	NA	NA	NA	NA
Frequency of Nash Actions	1–10	40.6	52.4	57.2	58.2	53.0	24.1	43.7
	11–20	40.3	61.8	55.3	58.9	42.0	30.7	43.5
	21–30	42.9	NA	NA	NA	NA	NA	NA

[‡]We report only convergence to the Pareto efficient equilibrium. For each treatment, respectively, 17.6%, 25%, 15.8% and 40% of the pairs converged to the Pareto inefficient equilibrium.
[†]In all groups, the column player had a last-period deviation to the static best-response.

however, the Nash equilibrium is Pareto dominated by (Z, Z). Note that Z is also the column player's Stackelberg action, while the row player's Stackelberg action is X.

4.2. Results From Other Experiments

As can be seen, the previously outlined six games used in the other experiments have very different characteristics than the games we chose to study. As such, by studying their data, we can hope to see whether teaching is present in games with other properties and, if so, whether *what* subjects attempt to teach differs across these games.

Tables 9 and 10 replicate our earlier analysis of the AP treatments for these six games. Examining convergence rates, we see that after 20 periods, very few pairs managed to converge to a Nash equilibrium in Terracol and Vaksmann (2009), while after 30 periods, almost half of the subjects converged to an equilibrium. Thus, the presence of multiple Pareto incomparable equilibria seems to make convergence more difficult. In the Hyndman et al. (2009) games, we see that the highest convergence rate (to the efficient equilibrium) was achieved when teaching was easiest (high premium, low cost), and the lowest convergence rate was achieved when teaching was most difficult (low premium, high cost). In terms of convergence to the inefficient equilibrium, the highest convergence rate (40%) occurs when teaching is most difficult. Finally, in Fehr et al. (2009), only about 26% of pairs converged to the Nash equilibrium, while another 22% converged (but for a last period deviation) to the column player's Stackelberg equilibrium, which Pareto dominates the Nash equilibrium.

Just by looking at convergence rates, it is difficult to say that teaching was going on. Therefore, in Table 10, we report the frequency with which subjects chose a

TABLE 10. Results from other experiments: teaching.

			HTV‡					
	Periods	TV†	HL	HH	LL	LH	FKD	
Frequency that action was a best response to stated beliefs	1–10	52.9	63.3	71.6	74.7	71.3	59.4	
	11–20	62.4	75.3	85.7	82.9	85.7	64.8	
	21–30	70.6	NA	NA	NA	NA	NA	
							X^*	Z^{**}
Frequency that Nash action chosen, *conditional upon subject not best responding to stated beliefs*	1–10	52.5	92.0	94.5	84.4	74.4	14.2	69.4
	11–20	45.3	90.5	88.7	78.5	74.4	10.5	77.9
	21–30	51.0	NA	NA	NA	NA	NA	NA

‡We report the frequency with which subjects chose the efficient Nash action, conditional upon subjects not best-responding to beliefs. †This refers to the frequency with which subjects chose their preferred Nash equilibrium action (i.e. action Y), conditional upon subjects not best-responding to beliefs. *This refers to the frequency with which subjects chose action X (which is each player's Nash action, and the row player's Stackelberg action), conditional on subjects not best-responding to beliefs. **This refers to the frequency with which subjects chose action Z (which is the column player's Stackelberg action), conditional on subjects not best-responding to beliefs.

particular action conditional on not best-responding to stated beliefs. For Terracol and Vaksmann (2009), we report the frequency that subjects chose their *preferred* Nash equilibrium action conditional on not best-responding to beliefs; for Hyndman et al. (2009) we report the frequency with which subjects chose the efficient equilibrium action conditional on not best-responding to beliefs; finally, for Fehr et al. (2009), we report both the frequency that subjects chose X (the Nash action) or Z (column's Stackelberg action), conditional on not best-responding to beliefs. The results here indicate that subjects are willing to take statically suboptimal actions in order to influence the ultimate outcome of the game. The results also shed light on the question of what subjects try to teach, which we now turn to.

4.3. What do Teachers Try to Teach?

Since our games had a unique pure-strategy Nash equilibrium on the Pareto frontier, the fact that our subjects engaged in teaching offers us little insight into what they were attempting to teach: what else would one teach but Nash? Hence, there was no way for a teacher to benefit from trying to teach her opponent to play in a non-equilibrium manner (e.g. be a Stackelberg leader) or to alternate between cells in a way to increase her payoff. Such was not the case in the experiments performed in the three outside studies we discussed. Here, because the payoffs were either not on the Pareto frontier or because of the existence of multiple equilibria, there were more opportunities to teach, more varied things to teach, and therefore more opportunities for us to gain insights into the motives of teachers.

One might hypothesize that there are two motivations for teaching. In one, the teacher is a payoff maximizer who uses teaching to convince her opponent that she is committed to choosing an action which either yields her preferred equilibrium outcome (if many equilibria exist) or to establish herself as a Stackelberg leader hoping to influence the beliefs of her opponent and show that she is committed to playing Stackelberg equilibrium. Under this motive the point of teaching is to alter beliefs in a self-serving manner or to build a reputation.

An alternative might be to lead one's opponent to an outcome that is preferred on ethical grounds, such as fairness or efficiency. Here teaching attempts to make one's opponent aware of the existence of this *preferred* outcome and show a willingness choose it.

The data generated by the three external experiments tend to support the former hypothesis that teaching is self-serving and aimed at influencing the expectations of one's opponent. For example, in the Terracol and Vaksmann (2009) experiments, one might imagine that the equilibrium (X, X) is attractive on two grounds: it maximizes the sum of equilibrium payoffs and each player gets his/her middle equilibrium payoff, making it fair in some sense. However, it was never the case that players converged to this equilibrium: Four times convergence was to column's Stackelberg outcome and five times it was to row's Stackelberg outcome. That is, people appear to teach the equilibrium which is most attractive to themselves, rather using teaching to lead to the compromise payoff (X, X).

In the Fehr et al. (2009) experiments the strategic dilemma for the players is the fact that while there is a unique Nash equilibrium (X, X) in pure strategies, the outcome (Z, Z) Pareto dominates it. Therefore, a logical teaching strategy might be to try to teach one's opponent to play Z in exchange for her reciprocation (this would be consistent with the second view of teaching). Beyond the fact that (Z, Z) Pareto dominates the Nash equilibrium, action Z is also the column player's Stackelberg action, meaning that the column player has an additional incentive to choose Z, even if it is not a best response. In Table 10, we reported that Z was chosen 69.4% of the time, conditional on subjects not best-responding, over the first ten periods. If we break this up across player roles, we see that row players chose Z 49% of the time, while column players chose Z 85.4% of the time. One can, perhaps, attribute the difference between row and column players as arising from the column player's additional, self-interested, motivation to teach Z, above and beyond any shared concerns for efficiency.

In terms of game outcomes, we see that seven of 27 pairs converged to the Nash equilibrium, while six of 27 pairs essentially converged to the Stackelberg equilibrium, though in all cases, the column player had a last period deviation to the static best-response, Y. However, in terms of teaching, Table 10 indicates that subjects spend most of their efforts trying to lead the way to the column player's Stackelberg equilibrium.[25] Moreover, given the fact that the column players always deviated in the last period to static best-response, it seems that this teaching was, potentially, more of the opportunistic variety.

25. To be sure, there is some bias in these numbers. Stackelberg teaching in the Fehr et al. (2009) game, if successful, will always be suboptimal in the static sense for the column player. In contrast, in our games, since convergence is to a Nash equilibrium, the teacher's actions will eventually be a best response to stated beliefs post-convergence.

Finally, in the Hyndman et al. (2009) experiment, we have a clear case where what to teach should be obvious—subjects should teach each other to play the equilibrium which Pareto dominates the other. This is exactly what they do, though they respond to incentives, with teaching being most prevalent when the teaching premium is high and the teaching cost is low and being least prevalent under opposite conditions.

The punch line, therefore, seems to be that subjects teach to alter the beliefs of their opponents in an effort to increase their payoff. When, as in our original 3×3 games, the unique Nash equilibrium is also Pareto optimal, they teach Nash. When Stackelberg-like equilibria exist, those are taught to the exclusion of more equitable Nash outcomes while if there is a joint Pareto best equilibrium, it is what is selected.

4.4. The RM and OP Treatments

The results presented in Section 3 showed the importance of teaching in facilitating convergence to Nash equilibrium in the two games we have considered when teaching is relatively easy. We also believe that the analysis so far in Section 4 demonstrates that teaching plays an important role in games with other properties, such as multiple equilibria (either Pareto rankable or not) or a game with a unique, Pareto inefficient equilibrium. In this section we study behavior in other environments in which teaching should be more difficult. If teaching plays a role in facilitating convergence, then if we make teaching difficult, we should see less of it. To do this we turn to the RM and OP treatments.

For both the RM and OP treatments, we expect teaching to be difficult. In the RM treatment, subjects were randomly rematched each period, dramatically reducing the incentives to teach, and certainly making teaching more difficult.[26] In contrast, in the OP treatment subjects did not know their opponent's payoffs so that they were unable to calculate the Nash equilibrium. Clearly, this makes teaching virtually impossible since a subject does not necessarily know what to teach or even how to interpret his opponent's response. Note that standard backward looking models would not predict that convergence rates differ in these treatments from those seen in our AP Treatment.

The evidence presented in Table 11 is consistent with our hypothesis that teaching facilitates convergence to a unique Nash equilibrium. As can be seen, after 20 periods the frequency of Nash actions is lower for both the RM and OP treatments and in both the DSG and nDSG games. In all cases, a proportions test shows the difference is highly statistically significant (in all cases $p \ll 0.01$). Next, if we look at the frequency that subjects chose a best response to their stated beliefs, we see that in periods 1–10, there are no substantial differences; in fact, in all but one case, subjects in our RM and OP treatments actually best responded slightly more often than in the AP treatment. However, over periods 11–20, the best response rate increased much more in the AP treatments, likely owing to the greater convergence to equilibrium. Finally, if we look

26. See, however, Ellison (1997) who shows that a single rational player interacting in a population of myopic players may be able to move the population to a Nash equilibrium if she is patient enough.

TABLE 11. Behavior in other treatments: convergence and teaching.

		AP		OP		RM	
	Periods	DSG	nDSG	DSG	nDSG	DSG	nDSG
Frequency of Nash actions	1–10	56.7	45.8	29.2	18.6	28.5	14.5
	11–20	64.5	55.3	32.9	17.9	36.0	7.0
	21–30	NA	NA	33.7	17.6	52.0	4.5
	31–40	NA	NA	37.6	22.6	72.5	7.0
	41–50	NA	NA	46.6	23.5	89.5	3.5
	51–60	NA	NA	48.7	24.4	93.0	6.0
Frequency that action chosen was a best response to stated beliefs	1–10	59.5	59.8	56.9	63.2	69.0	61.0
	11–20	70.3	76.1	62.1	68.1	57.5	64.5
	21–30	NA	NA	62.1	65.9	68.5	65.5
	31–40	NA	NA	60.3	73.8	74.5	61.0
	41–50	NA	NA	62.4	76.2	81.0	61.5
	51–60	NA	NA	65.3	75.6	84.0	57.5
Frequency that Nash action chosen, *conditional upon subject not best responding to stated beliefs*	1–10	61.8	46.3	50.3	37.4	67.7	33.3
	11–20	54.2	41.1	42.9	31.7	56.5	18.3
	21–30	NA	NA	41.7	29.3	69.8	13.0
	31–40	NA	NA	36.4	28.1	58.8	17.9
	41–50	NA	NA	46.2	24.7	57.9	7.8
	51–60	NA	NA	39.4	19.3	59.4	14.1

at the frequency of times that subjects chose the Nash action conditional on not best-responding to beliefs, we see that, in all cases (and particularly for the nDSG game) it is lower in the RM and OP treatments than in the AP treatment. This also suggests that subjects are not attempting to teach their opponent to play the Nash equilibrium.

Recall that we ran our OP and RM treatments for an additional 40 periods to see if we observe delayed convergence. Table 11 also has these results. In the OP treatment, there was more frequent Nash play after 60 periods but the frequencies were still below those of the corresponding 20-period frequencies from the AP treatment. In the RM treatment, for the nondominance solvable game, increasing the length of play had no effect on convergence, while for the dominance solvable game, in the final 10 periods of play, the Nash action was chosen 93% of the time.

Thus when teaching is more difficult, as is the case with the RM and OP treatments, convergence rates go down. However, it appears to be true that, given enough time, in some environments convergence rates may rise even when teaching is difficult. This is particularly true of the dominance solvable game in the RM treatment.

Although some teaching may be present, as evidenced by the high frequency of Nash action choices that were not best responses to stated beliefs, we believe that the high convergence rate is due to subjects learning how to iteratively delete dominated strategies. To illustrate the iterative dominance principle at work, consider Figure 9, which shows the frequency of actions taken by row and column players each period in the DSG game. For this game the first strategy to be deleted is A_2 for the row player, and as can be seen, row players virtually never play that strategy over the course of the

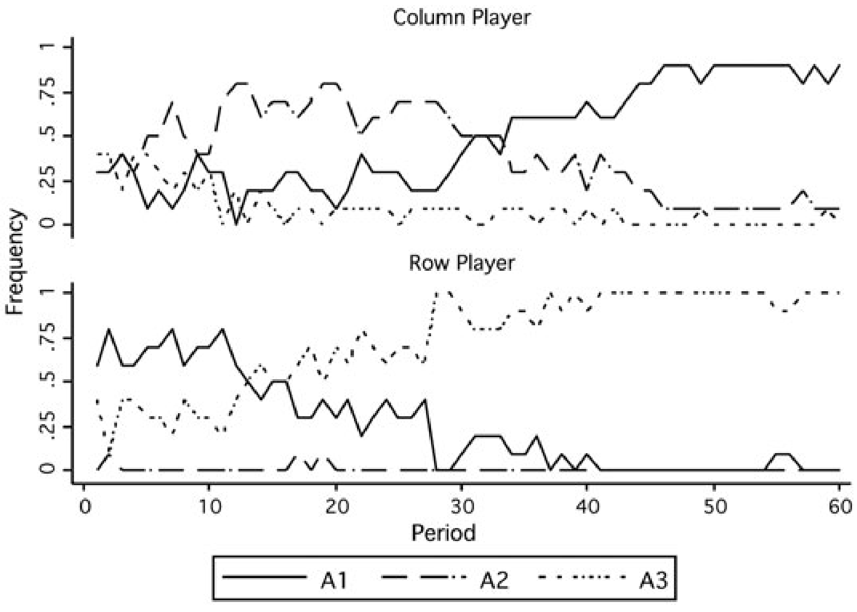


FIGURE 9. Frequency of action choices (Random Matching: DSG).

experiment. The next strategy to be eliminated is A_3 for column. As we see, by period 13 the mean use of strategy A_3 for column drops below all the other strategies and stays there throughout the rest of the game. Clearly it is the second strategy eliminated. This leads to the Row Player's strategy A_1 being eliminated by period 40 etc. A similar pattern can be seen for the beliefs of the subjects.[27]

While convergence was achieved via iterated deletion of dominated actions in the DSG game, no such possibility was present in the nDSG game. As Table 11 shows, subjects almost never chose the Nash action. Moreover, there is no evidence that teaching was taking place: over periods 1–10, the frequency of Nash actions that were not also best responses to stated beliefs was only 33.3%, which would be consistent with random choice. In all other 10 period ranges, the corresponding frequency was never more than 20%. Clearly, there is very little evidence of Nash play or attempted Nash teaching. The question remains as to what explains subjects' behavior. If we estimate the γ/λ model described in Section 3.4, we obtain $\lambda = 0.09$ and $\gamma = 0.84$, and we easily reject that $\lambda = 0$ ($p \ll 0.01$). Moreover, the average frequency of play corresponds very closely to a quantal response equilibrium (with $\lambda \approx 0.8$). Therefore, it seems that a model of stochastic best response to stated beliefs, where beliefs update according to the past history of observed actions is the leading explanation of behavior in this treatment.

27. The corresponding figure showing the iterative process for beliefs is available upon request.

5. Conclusions

This paper has attempted an investigation of the process through which people playing games converge to an equilibrium—a state where their beliefs about the actions of their opponents are confirmed. The results of our experiments support our hypothesis that, in the class of games that we study, teaching plays an important role in the process of convergence to Nash equilibrium. More precisely, in the two person, 3×3 games we used, those pairs that successfully converged did so through a process quite different from the backward looking process described in much of the learning literature. Rather, convergence seems to be an action led process where one player, the teacher, takes it upon herself to lead the way to the Nash equilibrium by repeatedly choosing her Nash action despite the fact that it is not a best response to her beliefs. Successful convergence matches such a teacher with a fast learner, i.e. someone who places sufficient weight on recent history. Nonconvergence appears, predominantly, to be the result of beliefs updating too sluggishly, rather than an inability of subjects to best respond.

While our experiments have focused on games with a single pure strategy Nash equilibrium with payoffs on the Pareto frontier, we also examined data from other experiments, whose games had different properties than our own. These results are also consistent with the hypothesis that teaching has an important role in the process of convergence. However, these games also demonstrate that *what* subjects attempt to teach depend on the properties of the game being played. When the game has multiple, Pareto rankable equilibria, teaching is exclusively towards the efficient Nash equilibrium, though the extent of teaching depends on the strength of the incentives to do so. When games have multiple, Pareto incomparable equilibria, players try to teach their most preferred equilibrium. Finally, when the game has an equilibrium which is Pareto dominated by another strategy profile, a non-negligible subset of players try to teach their opponent to play Pareto dominating strategy profile, while others are content to teach their way to the Nash equilibrium, even though it is inefficient. These last two games suggest that one cannot ignore Stackelberg equilibria when studying the process of convergence and teaching.

We also examined the robustness of our results with a series of other experiments. In particular, if teaching plays an important role in the process of convergence, then if teaching is made more difficult, we should see less, and possibly delayed, convergence. In one treatment we increased the strategy space, in another we only showed players their own payoffs (OP treatment) and, lastly, we had a treatment with random matching (RM treatment). What we found is that as teaching becomes more difficult convergence becomes rarer. This is interesting because, according to a large segment of the learning literature, these treatments are expected to have either no impact (in the OP treatment) on the convergence rates of many backward looking processes or actually enhance it (in the RM treatment). Our results are strongest for the nondominance solvable game, where convergence rates after 60 periods were still well below the 20-period convergence rate of our baseline. Interestingly, in the dominance solvable game for the RM treatment subjects were able to converge after 60 periods, though the procedure they used appeared to follow an iterative deletion of dominated actions process rather than a teaching process, which may explain why it took so long for convergence to occur.

There is still a lot of work to be done. If what we have uncovered here is replicated, investigators may want to span a wider set of games and environments to see if teaching is relevant in all, some or none of them. For example, in multi-person games the ability of a player to teach is diluted by the actions of others. Still environments with different feedback or communication rules may foster teaching and hence convergence.

References

Camerer, Colin F. and Teck-Hua Ho (1999). "Experienced-Weighted Attraction Learning in Normal Form Games." *Econometrica*, 67, 827–874.

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong (2002). "Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games." *Journal of Economic Theory*, 104, 137–188.

Cheung, Yin-Wong and Daniel Friedman (1997). "Individual Learning in Normal Form Games: Some Laboratory Results." *Games and Economic Behavior*, 19, 46–76.

Cho, In-Koo, Noah Williams, and Thomas J. Sargent (2002). "Escaping Nash Inflation." *Review of Economic Studies*, 69, 1–40.

Conlisk, John (1993a). "Adaptation in Games: Two Solutions to the Crawford Puzzle." *Journal of Economic Behavior and Organization*, 22, 25–50.

— (1993b). "Adaptive Tactics in Games: Further Solutions to the Crawford Puzzle." *Journal of Economic Behavior and Organization*, 22, 51–68.

Costa-Gomes, Miguel A. and Georg Weizsäcker (2008). "Stated Beliefs and Play in Normal Form Games." *Review of Economic Studies*, 75, 729–762.

Costa-Gomes, Miguel A., Vincent P. Crawford, and Bruno Broseta (2001). "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica*, 69, 1193–1235.

Crawford, Vincent P. (2002). "Introduction to Experimental Game Theory." *Journal of Economic Theory*, 104, 1–15.

Dufwenberg, Martin and Uri Gneezy (2000). "Measuring Beliefs in an Experimental Lost Wallet Game." *Games and Economic Behavior*, 30, 163–182.

Ellison, Glenn (1997). "Learning from Personal Experience: One Rational Guy and the Justification of Myopia." *Games and Economic Behavior*, 19, 180–210.

Fehr, Dietmar, Dorothea Kübler, and David Danz (2009). "Information and Beliefs in a Repeated Normal-Form Game." SFB 649 Discussion Paper 2008-026, Technische Universität Berlin.

Fischbacher, Urs (2007). "z-tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10, 171–178.

Fudenberg, Drew, and David K. Levine (1998). *The Theory of Learning in Games*. The MIT Press.

Haruvy, Ernan (2002). "Identification and Testing of Modes in Beliefs." *Journal of Mathematical Psychology*, 46, 88–109.

Hopkins, Ed (2002). "Two Competing Models of How People Learn in Games." *Econometrica*, 70, 2141–2166.

Huck, Steffen and Georg Weizsäcker (2002). "Do Players Correctly Estimate What Others Do? Evidence of Conservatism in Beliefs." *Journal of Economic Behavior and Organization*, 47, 71–85.

Hyndman, Kyle, Antoine Terracol, and Jonathan Vaksmann (2009). "Learning and Sophistication in Coordination Games." *Experimental Economics*, 12, 450–472.

Mookherjee, Dilip and Barry Sopher (1994). "Learning Behavior in an Experimental Matching Pennies Game." *Games and Economic Behavior*, 7, 62–91.
Partow, Zeinab and Andrew Schotter (1993). "Does Game Theory Predict Well for the Wrong Reasons: An Experimental Investigation." Working Paper 93-46, C.V. Starr Center for Applied Economics, New York University.
Rutström, E. Elisabet and Nathaniel T. Wilcox (2009). "Stated Beliefs versus Inferred Beliefs: A Methodological Inquiry and Experimental Test." *Games and Economic Behavior*, 67, 616–632.
Sargent, Thomas J. and In-Koo Cho (2008). "Self-confirming Equilibrium." In *The New Palgrave Dictionary of Economics*, edited by Steven Durlauf and Larry Blume.
Sargent, Tom and Albert Marcet (1989). "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models." *Journal of Economic Theory*, 48, 337–368.
Storn, Rainer and Kenneth Price (1997). "Differential Evolution—A Simple and Efficient Heuristic for Global Optimization Over Continuous Spaces." *Journal of Global Optimization*, 115, 341–359.
Terracol, Antoine and Jonathan Vaksmann (2009). "Dumbing Down Rational Players: Learning and Teaching in an Experimental Game." *Journal of Economic Behavior and Organization*, 70, 54–71.