Experimental Research about Learning in Inductive Game Theory

Yu Wang[1,a], Ming-Sheng Chen[2,b,*], Xiao Han[3,c], Fei-Na Yu[4,d]

196 West Huayang Rd., Yangzhou University, Yangzhou, Jiangsu Prov., China

[a]52964280@qq.com, [b]1183023766@qq.com, [c]949401659@qq.com, [d]1165354853@qq.com

*Corresponding author

Keywords: Experimental economics, Learning, Inductive game theory.

Abstract. In this article we expect to research how do players who are incompletely informed learn about the game from the perspective of experiment. We find that players can form and update their personal views about the rules of the game through interactions and then often choose the best replying strategy according to the Nash equilibrium of the games. The experimental data indicates that there are some factors can influence the effect of learning. Moreover, players are more likely to choose their strategies according to QRE (quantal response equilibrium).

Introduction

The former researches about learning in games always assume that the rules or the formation of the game is the common knowledge of all players, irrespective of classical game theory, evolutionary game theory and behavioral game theory. Kaneko (2013) noted that former researches neglected to explore how players perceive the game itself and how the cognitions about the game of players updating as time goes on. However, different from game theory and evolutionary game theory, behavioral game theory attempts to use the theory of learning in games to explain the changes of strategy selection. It provides a particular perspective for us to concern how players learning and acting in a game. Devetag (2003) asserted, in her experimental study, that the players' ability of learning the rules of the game was closely related to the ability of players' short-term memory which was a measurable proxy of their cognition. Hankie, et al. (2005) explored the strategy learning of players, which occurs in two phases, the long-run phase and the short-run phase. Branas-Garza (2012) identified players' cognitive ability that was the basement of behavior of players in the beauty contest game by both Raven and Cognitive Reflection tests. But behavioral game theory didn't research the rules learning in games. However the developing inductive game theory does that. Kailai (1993) suggested that, when players had incomplete objective knowledge of the game, the rules of the game in the players' mind may not be the objective description of the game, and it could be more likely to be the players' subjective cognition of the game. Moreover, the equilibrium of the game was also the subjective equilibrium. Oechssler (2002) pointed out that most of the players set up their own subjective game on the basis of the understanding of the objective game, then acted to the Nash equilibrium of the subjective game and the action were usually very close to the Nash equilibrium of the objective game. Kaneko (2006) argued that, the cognition of boundedly rational players about the rules of game was not given exogenously, but inducted from the previous learning of the game. Hanaki (2009) applied computer simulation method to study the influence of long-term memory and short-term memory in the process of learning in the repeated game.

In this paper, we will research that how the players who did not fully understand the game learning the game itself and how to choice the strategy in the experiments, considering that players have limited reasoning ability and can learning in playing the game.

Inductive game theory and game learning

Kaneko& Kline (2008) proposed the inductive theory. They thought that the bounded rationality of players in the case of incomplete understanding the game can cognate the formation of the game by inductive learning. Inductive learning was closely related to the memories of past actions of

players.

Hanaki & Ishikawa (2009) considered that players' cognitive ability is related to their short-term memory and long-term memory ability and by computer simulation technology to study the following two questions:

Personal views how to form

Hanaki, et al. (2009) put out a learning model where it assumes the players are bounded rational, with two forms of memory, short-term memory and long-term memory. Short-term memory records the outcomes of the game, and it will disappear thereafter. a short-term memory can be transformed into a long-term memory which will exists in players' mind forever, if it repeated for a value that we call it cognition threshold before it disappears.

The short-term memory length of player i is represented by mi and the threshold, ki. So only if a outcome (si , sj) be repeated ki times in mi, it can be a long-term memory, and then player i will realizes his or her payoff. The long-term memory may changes over time, and then the personal views also change.

Let $\prod$ be the objective payoff matrix of the game under consideration, and let $\prod$i represent the part of the payoff matrix that player i receives.

$$\prod{}^i = \begin{bmatrix} \pi^i\left(s_1^i, s_1^j\right) & \dots & \pi^i\left(s_1^i, s_{n^j}^j\right) \\ \dots & \dots & \dots \\ \pi^i\left(s_{n^i}^i, s_1^j\right) & \dots & \pi^i\left(s_{n^i}^i, s_{n^j}^j\right) \end{bmatrix} \tag{1}$$

In which ni and nj are the numbers of strategies in Si and Sj, and $\pi i:Si \times Sj \to R$ is the payoff function of player I, $\pi i(si,sj) \neq 0$, si $\in$ Si, sj $\in$ Sj; i, j $\in$ {1,2},i$\neq$j.

Let $\tilde{\prod}^i$ be the personal view of the game for player i at period t.

$$\tilde{\prod}{}^i(t) = L^i(t) \cdot \prod{}^i \tag{2}$$

Li (t) be the state of the long-term memory in player i's mind at period t. Li (t) takes value one, if (si,sj) is a long-term memory of player i at period t, and it is zero otherwise. So $\tilde{\prod}^i$ is one when player i has learned (si,sj) at period t, and it is zero otherwise.

How do players learn to choose strategies

In their research, Hanaki, et al. assumed that in each period, players chose their actions based on their attractions for each action. Players learned through the evolution of attractions.

Let $A_s^i(t)$ denote player i's attraction for s $\in$ Si at period t. Let $P_s^i(t)$ be the probability that player i chooses action s at period t.

$$p_s^i(t) = \frac{e^{\lambda^i A_s^i(t)}}{\sum_{k \in S^i} e^{\lambda^i A_k^i(t)}} \tag{3}$$

Parameter λ^i represents the favored extent of actions with higher attractions. When $\lambda^i = 0$, all actions are equally likely to be chosen. Actions with higher attractions become more likely to be chosen. But the attraction of each action will be different gradually.

Mclkelvey and Palfrey (1995, 1998) introduced the solution concept in game theory, quantal response equilibrium. In QRE, at each period players with bounded rationality do not choose the best response with probability 1 but they always choose better strategies that brings higher expected

payoff i with higher probabilities. In their research, QRE is often specified by logit equilibrium (LQRE).In a logit equilibrium player's strategies are chosen according to the probability distribution:

$$P_{ij} = \frac{\exp(\lambda \sum_{s_{-i}} P(s_{-i})u_i(s_i,s_{-i}))}{\sum_{s_k} \exp(\lambda \sum_{s_{-i}} P(s_{-i})u_i(s_k,s_{-i}))}$$
(4)

λ represents the sensitivity of different expected payoff of players. The initial value of λ is 0 but as it increases; players' responses become more accurate and gradually converge to Nash equilibrium. $P(s_{-i})$ represents the believe about other players' action, and players computerize their expected payoffs according to it

Experimental design

In our computerized experiment involved two 2x2 normal form games; we studied players' learning behavior. The games in experiment was divided into two blocks that were used respectively, Block 1and Block 2. The first game was a prisoners' dilemma where strategy profile (Y, Y) is the unique and dominant Nash equilibrium for both players. The second game was a stag–hunt game with two strict Nash equilibria in pure strategies, (X, X) and (Y, Y). The equilibrium (X, X) is payoff dominant, while (Y, Y) is risk dominant.

We invited 26 freshmen to the experiment where they were anonymously matched. We assumed that the strategy profile of each players were common knowledge and all the players only known their own payoff but don't know the others payoff after each round.

In block 1, players with one opponent played 25 rounds, and after 15 rounds, all players received a questionnaire composed of a payoff matrix, and four questions:

		Player 1	
		X	Y
Player 2	X	40，40	50,12
	Y	12,50	25,25

Fig.1 Block 1

		Player 1	
		X	Y
Player 2	X	24,24	24,0
	Y	0,24	40,40

Fig.2 Block 2

		Player 1	
		X	Y
Player 2	X	a	b
	Y	c	d

Fig.3. the payoff of each player

(1)What is the value of a?
(2)What is the value of b?
(3)What is the value of c?
(4)What is the value of d?

Players should answer these questions, and after finishing the rest 10 rounds, they need answer the questions again. In block 2, players only need finish the questionnaire after 15 rounds. After the experiment the total payoff of all players will be converted into RMB according the proportion of 1%.

Results

Can players learn the game

Table 1 shows the percentage of correctly answered questions in two periods of block 1, separate for each question. We found that about 1/4 players got the correct value of a, b, and c, while about 1/2 players got the correct value of d, after first 15 rounds. While in the second period there are more than 2/3 players correctly got a; more than 3/4 players got d; and the ones correctly got b and c about increased by 1/2; in addition, the incretion of correctly got all values was about 1/3. Table 2 shows that subjects answered significantly more questions correctly for the coordination game in Block 2 than Block 1, after first period. The reason may be that block 2 was simpler than block 1 in cognition and players had got some experiences from learning block 1.

Specially, we found that about 3/4 players' strategy choice shifted after finishing the questionnaire by comparing round 11-15 and 21-25 of block 1. And about 4/5 players who shifted their strategy to chose strategy Y more than 4 times in round 21-25. So we summarize that, most players could choose their best replying strategies according to inductive game ∏i, i.e., players can construct their inductive games by leaning through repeated game, and act according to it. Through, some players' inductive games ∏i were partly like the objective game∏, some were totally different from it.

Tab.1 Percentage of correctly answered questions in Block 1

Question	pre 15 rounds	latter 10 roungds
Q1	0.23	0.69
Q2	0.23	0.42
Q3	0.27	0.54
Q4	0.46	0.77
Four correct	0.08	0.38

Tab.2 Percentage of correctly answered questions in Block 2

Question	
Q1	0.81
Q2	0.54
Q3	0.58
Q4	0.73
four correct	0.38

How players change the choice of strategies

Players would continually update their personal view about the game in the repeated game, by observing the shift of strategies.
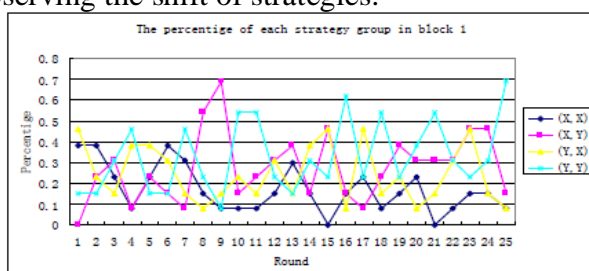

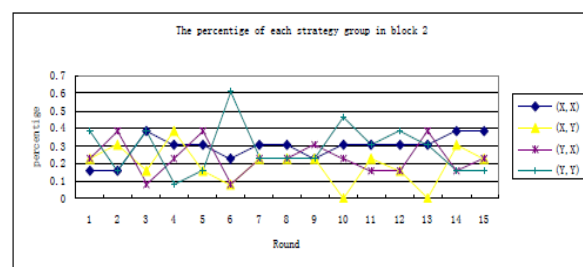Fig.4 percentage of strategy profiles in block1   Fig.5 percentage of strategy profiles in block2

We have noted that players had one Nash equilibrium (Y,Y) in block 1.While they had two in block 2,(X,X) and (Y,Y), but (X,X) was Pareto superior to (Y,Y). Fig.4 shows that only the percentage of Nash equilibrium (Y, Y) was gradually increasing, while the percentage of (X, X) was

gradually decreasing. The percentage of (X, Y) and (Y,X) is in a state of shock, but appeared a decreasing trend. Fig.5 shows that only the percentage of Pareto Nash equilibrium (X, X) was gradually increasing, the rest of the strategy profiles were present a downward trend.

Especially, we want to know whether players choose Nash equilibrium strategy after forming the personal view about the game. In fig.4 we can see that most players gradually turned to choosing Y in last 10 rounds of block 1. i.e., with the time goes, (Y, Y) will be thought to be the Nash equilibrium by players. In fig. 5 we can see that in the last five rounds of block 2 strategy X was adopt by most players while the others besides Y were abandoned over time. Because in most players' inductive game, X is the best replying strategy. Therefore, after constructing subjective game, players will consciously choose the inductive Nash equilibrium strategies and the chosen Nash equilibrium strategies were always approach to the objective ones. In model 1, we find that players didn't shift their choosing to the Nash equilibrium in the first 15 rounds because of bounded rationality. Initially, λ in QRE is 0, they chose their strategies with equal probability stochastically, but in the repeated process, players would altered their believes and the value of λ increased. So their choosing was gradually approaching and converging to Nash equilibrium.

The influence factors of players' strategy choice

In our experiment, we found three main factors that influence the players' strategy choice. First of all, it is external information. Camerer (2003) found that after some period of time of strategies learning players can find the best replying strategy in the research. Kaneko& Kline (2008, 2013) noted that by inductive learning the past actions, players could cognize the form or rules of the game eventually. But they all thought that it would take a long time. In our experimental research, we found that in the process of repeated game players provided with the external information would take shorter time in cognizing the form or rule of the game. Secondly, We think that players' cognitive ability has a important influence on their strategy learning. Hirshleifer(2000) and Camerer(2003) demonstrated the cognitive differences of players impacted on their decision-making from the perspective of theory and experiment respectively. Through the analysis of the experiment results, we found that in the repeated game the stronger the strategy learning ability of players, the faster for them to find the best strategy. In the second period of model 1, about 3/4 players chose strategy 2 over the probability of 1/2, who chose strategy 2 at an average probability of 1/2. That means they tended to try different strategies. By contrast, the other players chose Nash equilibrium strategy, strategy 2, at the average probability of 1/4. Thirdly, we consider the gender as a important factors. Eckel & Grossman (2001) found that the strategic behavior of men and woman had significantly difference in the ultimatum game experiments. We also received a similar observation in our experiment. We found that in the two phases of model 1, male players tented to chose strategy 1, while female players tended to chose strategy 1 and strategy 2 at the probability of 1/2.

Conclusion

In this paper, we used experimental research to study the players with bounded rationality how to learn and act in inductive game theory. Players initially had little knowledge about the game they played in our experiment. They got experience through playing the game repeatedly. Based on their experience, they not only learned which action would bring about a higher payoff but also formed their view about the game they are playing. It is noteworthy that personal views and behaviors of players coevolve in playing the game. i.e., The players' personal view and behaviors influence each other in the process. Our conclusions are: (1) In the process of repeated game, players can form their personal view by learning; (2) After forming personal view, players choose best replying strategy consciously according to it. (3) Players can choose Nash equilibrium strategy in accordance with the inductive game and form behavior convergence.

In future, we may extend our research in several dimensions. Because we found some interesting questions needed to explore. We may investigate how players to act in a game in which they won't

be supplied with any hints about the formation of game. We will research whether players behave in the same way when they are initially only informed about the set of actions to themselves but they can observe the opponent's action choosing and payoff after each interaction. Possible work is also to run current experiment with larger games and more rounds to research players how to learning and acting in more complex situation.

References

[1]Devetag, G., Warglien, M., 2008, playing the wrong game: an experimental analysis of relational complexity and strategic misrepresentation, Games and Economic Behavior, 62,364-382.

[2]Kalai, E., Lehere, E., 1995, Subjective games and equilibria, Games and Economic Behavior, 8,123-163.

[3]Camerer, C., 2003, Behavioral game theory: experiments in strategic interaction, Princeton University Press.

[4]Hanaki, N., Ishikawa, R., Akiyama, E., 2009, Learning games, Journal of Economic Dynamics & Control, 33, 1739-1756.

[5]Kaneko, M., Kline, J., 2008, Inductive game theory: A basic scenario, Journal of Mathematical Economics, 44, 1332–1363.

[6]Oechssler, J., Schipper, B., 2003, Can you guess the game you are playing, Games and Economic Behavior, 43,137–152.

[7]Keenan Huang, 2010, Inductive game theory and the endogenous evolution of institution, Economic Research, 4, 134-146.

[9]Kaneko, M., Kline, J.J., 2013, Partial memories, inductively derived views, and their interactions with behavior,Econ Theory ,53,27–59.

[10] Pedro Dal Bó and Fréchette,G., 2011 ,The evolution of cooperation in infinitely repeated games: Experimental Evidence,The American Economic Review,1,411-429.

[11]Branas-Garza,P.,García-Munoz,T.,González,R.,Cognitive effort in the Beauty Contest Game, Journal of Economic Behavior & Organization, 83,254–260.

[12] Akiyama, E., Ishikawa, R.,Kaneko, M., Kline, J.J., 2008 , A simulation study of learning a structure: mike's bike commuting, SSM.DP.1190.

[13] Fudenberg, D., Levine, D.K., 1998. The Theory of Learning in Games. MIT Press, Cambridge, MA.

[14]Hanaki, N., Sethi, R., Erev, I., Peterhansl, A., 2005. Learning strategy. Journal of Economic Behavior and Organization 56, 523–542.

[15]Hirshleifer, D., Welch, I., 2000, An Economic Approach to the Psychology of Change: Amnesia, Inertia and Impulsiveness. Working Paper, Ohio State University.

[16]Camerer, C., Hogarth, R., 1999, The effect of financial incentives in economics experiments: A review and capital-lab-production framework. Journal of Risk and Uncertainty, 18, 7-42.

[17]Richard, M., Thomas, P., 1995, Quantal Response Equilibria for Normal Form Games, Games and Economic Behavior 10: 6–38.

[18]Richard, M., Thomas, P., 1998, Quantal Response Equilibria for Extensive Form Games, Experimental Economics 1: 9–41.