



A Modeling of Game Learning Theory Based on Fairness

Qingquan He

School of Business, Central South University

Changsha 410083, China

E-mail: qingquan_he@126.com

School of Science, Hunan Agriculture University

Changsha 410128, China

Yulei Rao and Jie Xu

School of Business, Central South University

Changsha 410083, China

The project is supported by the National Natural Science Foundation of China(70872111), the Ministry of Education Post-project of Key Research Institute of Humanities and Social Science (06JHQZ0010) and Education Department Foundation of Hunan Province(08c398)

Abstract: By incorporating fairness factor in the EWA (experience-weighted attraction) learning model, we develop an extended game learning model called FGL model. We use psychological effect instead of material effect to modify strategy's payoff and attraction, and to study the equilibrium movement further in dynamic Games. That participants have fair thinking will, in turn, lead to their psychological function changes. Compared with EWA learning model by simulating the decision-making in Ultimatum Game, we find FGL model converges to equilibrium strategy faster.

Keywords: Fairness, Game, Learning model

1. Introduction

Traditional theories of economics are based on rational hypothesis which assumes people only pursue their own material self-interest. However, many famous economists such as Simon (1955), Arrow (1981), Samuelson (1993) and Sen (1995) believed that people are bounded rational rather than perfect rational in reality. The hypothesis which economic-man is self-interest has been challenged since 1980s. Many experimental economists have proved that participants are altruism, and have strategy learning and fair thinking in games, which contravene the behaviors forecasted by standard game theory in many different game experiments (Guth, Schmittberger, and Schwarze. (1982); Forsythe, Joel, Savin and Martin (1994); Camerer, Thaler, Richard (1995); Roth (1995); Fehr, Alexander and Schmidt (2007); Ernst Fehr, Jean-robert Tyran (2008); Qingquan He and Yulei Rao (2009)). Explaining these phenomena is the main focus of experimental economics and other related economic theories. Theoretical methods to explain these anomalies mainly include learning model and fairness model.

Learning model assumes that people are bounded rational, but it makes the material payoffs as utilities. There are three main learning models: belief learning model (Brown (1951); Milgrom and Roberts (1991); Hon-snrir, Monderer, and Sela (1998); Sela (2000); Berger (2005)), reinforcement learning model (Gale, K. Binmore and Samuelson (1995); Roth and Erev (1995)) and EWA learning model (Camerer, and Ho (1999)). Crawford (1995) thought players have belief learning in games, namely, players can take account of previous behavior by other players or themselves to update their beliefs about what others will do in the future, then choose a best-response strategy accordingly to maximize their expected payoffs. However, due to the payoffs of historical strategies, some players also may repeat successful strategies and abandon failed strategies. That is reinforcement learning. (Note 1) Experience-weighted attraction learning model (we called EWA learning later), designed by Camerer and Ho (1999), combines the most appealing elements of reinforcement and belief learning models. Learning models can explain game players' behavior better, however, the learning process forecasted by it is slower so that the strategy equilibrium can't be converged to fast.

Rabin, who developed the Fairness model, pointed out that player's utility is not just equal to material payoff, but depends on others' payoffs (Rabin (1993)). That is to say, people may often respond to others' intentions on their certain

behavior: people are willing to return kindness to those who they think are kind, and retaliate to those unkind, regardless of cost. Rabin defined a “kindness function” to measure others’ kindness or behavioral intention in his model. However, there may be multi-equilibriums forecasted by his model. For example, there may be multi-equilibriums when material payoff is smaller but psychological payoff is more important relatively. The model can’t forecast which equilibrium will emerge at last, when both fairness equilibrium and unfairness equilibrium exist and meet self-fulfilling prophecy.

Players don’t represent only learning behavior but also fairness thinking in games. As the learning models don’t take player’s fairness thinking into account, it forecasts player’s learning process slowly so that they can’t make strategies converge to equilibrium fastly. This paper attempts to incorporate fairness factor into learning model to form the game learning model based on fairness (ab. FGL model). Our purpose is to enhance the veracity of the forecasted equilibrium results, to improve player’s slowly learning process and solve the problems such as the imperfect learning effect and so on.

2. The Game Learning Model Based on Fairness (FGL Model)

We incorporate the fairness factor into the EWA learning model to form the FGL model. In the EWA learning model, every strategy has an “attraction” (Note 2), which implicates the choice probability of a certain strategy. The basic ideas of this model are as follows: for player i , there are m_i strategies (indexed by k) which have an initial attraction denoted $A_i^k(0)$. Denote i ’s k th strategy by s_i^k , strategies chosen by i and other players (denote $-i$) in period t by $s_i(t)$ and $s_{-i}(t)$, and player i ’s payoffs of choosing s_i^k by $\pi_i(s_i^k, s_{-i}^k(t))$, and the attraction and experience weight of i ’s k th strategy in period t by $A_i^k(t)$ and $N(t)$. Obviously, both $A_i^k(t)$ and $N(t)$ vary with the time t , that is, $A_i^k(t)$ and $N(t)$ are updated after every period. The EWA attraction $A_i^k(t)$ and the experience weight $N(t)$ (Note 3) updating equations from Camerer and Ho^{[21]-[23]} are:

$$A_i^k(t) = \frac{\phi N(t-1)A_i^k(t-1) + [\delta + (1-\delta)I(s_i^k, s_{-i}(t))]\pi_i(s_i^k, s_{-i}(t))}{N(t-1)\phi(1-\rho) + 1} \quad (1)$$

$$N(t) = N(t-1)\phi(1-\rho) + 1 \quad (2)$$

Parameter ϕ reflects the decay of previous attractions owing to forgetting or deliberating shedding of old experience when the learning environment is changing. ϕ is between zero and one. $N(t-1)$ is the experience weight of period $t-1$.

Function $I(s_i^k, s_{-i}^k(t))$ is an indicator function of the model which equals to 0 or 1 under different conditions, as

follows: $I(s_i^k, s_{-i}^k(t)) = \begin{cases} 1 & s_i^k = s_{-i}(t) \\ 0 & s_i^k \neq s_{-i}(t) \end{cases}$. In the equation 1, the term $[\delta + (1-\delta)I(s_i^k, s_{-i}(t))]$ implies that a weight

of one is put on the payoff term when the strategy being reinforced is the one the player chose ($s_i^k = s_{-i}(t)$), but the

weight on forgone payoffs from the other strategies ($s_i^k \neq s_{-i}(t)$) is δ . The parameter δ is the weight placed on forgone payoffs, which is presumably affected by imagination and the reliability of information about forgone payoffs, when $0 < \delta < 1$. The parameter ρ controls the rate at which attractions grow. It also captures influences of different models on attractions growing in the process of game learning.

Attractions can be mapped into choice probabilities using a logit response function by Camerer and Ho. The choice probability of player i ’s strategy k in period $t+1$ is:

$$p_i^k(t+1) = \frac{e^{\lambda A_i^k(t)}}{\sum_{k=1}^{m_i} e^{\lambda A_i^k(t)}} \quad (3)$$

Where λ is the response sensitivity.

We can learn that this learning model has taken history experience and the beliefs of players which have effects on game behavior. We can also learn that the material payoffs $\pi_i(s_i^k, s_{-i}(t))$ remain the same during the whole game process. However, people’s beliefs change with the game moving, then the payoffs of strategies change. Moreover, we believe that game players have fairness thinking. They adjust their strategies according to whether their opponents’ behaviors are fair or not. They may choose retaliatory strategies, even at a cost to themselves, a bad belief given on their opponents. However, if their opponents are kind, as “gift exchange”, they may choose some strategies which benefit both or even more to their opponents.

Well then, we replace material effect $\pi_i(s_i^k, s_{-i}(t))$ (Note 4) with fairness effect function $U_i(a_i, b_j, c_i)$ (Note 5) to measure psychological effect changes which are brought by the changes of game opponents' beliefs. Denote player i 's strategy by a_i , the strategy player i believes player j is choosing by b_j , while the strategy that player i 's belief about what strategy player j believes i is choosing by c_i . Then, denote the fairness effect function as follows:

$$U_i(a_i, b_j, c_i) = \Pi_i(a_i, b_j) + \tilde{f}(b_j, c_i)[1 + f(a_i, b_j)] \quad (4)$$

$\Pi_i(a_i, b_j)$ is the material effect when both game players choose strategies a_i, b_j respectively. $f_i(a_i, b_j)$ (ab. f_i) measures the degree of kindness and generosity player i to player j . $\tilde{f}_j(b_j, c_i)$

(ab. $\tilde{f}_j$) measures the degree of kindness which player i believes j to himself, while $\tilde{f}(b_j, c_i)[1 + f(a_i, b_j)]$ represents the psychological effect function brought by fair motivation.

Then the strategy attraction (Note 6) of FGL model is:

$$A_i^k(t) = \frac{\phi N(t-1)A_i^k(t-1) + [\delta + (1-\delta)I(s_i^k, s_i(t))][U_i(s_i(t), s_{-i}(t), s_i'(t))]}{N(t-1)\phi(1-\rho) + 1} \quad (5)$$

Here, $s_i'(t)$ represents the third belief about what strategy player i believes player j believes what strategy he is choosing. Other symbols have the same meaning as above.

According to our model, we can calculate the attraction of each strategy in every period. Then these attractions can be transformed into the choice probabilities according to equation (3).

3. The Compare of the Simulation to the Ultimatum Game

To verify the astringency and forecast ability of FGL model, we carry out a computer simulation to the ultimatum game, and compare our results with the results simulated by EWA learning model.

The ultimatum game is a game about dividing some amount of money or goods. There are two participants in the game: a proposer and a responder. At first, the proposer makes a take-it-or-leave-it offer $(1-x, x)$. x is the money for the responder, while $(1-x)$ is left for himself. The responder responds to the offer then. If he accepts the division, then both people earn the specified amounts and the game is over. If he rejects, they both get nothing. Also, the game is over.

Suppose 100 yuan is divided between the two participants in the ultimatum game. The proposer's strategy set composed of ten strategies what are the proportion offered to the other $\{0-10\%, 11-20\%$,

$21-30\%, \dots, 91-100\%\}$, symbolized by $\{S_1, S_2, \dots, S_{10}\}$. While the strategy set of the responder is $\{\text{accept, refuse}\}$. Suppose the offers from proposer obey the uniform distribution, then the average payoff of each strategy is corresponding to $\{95, 85, 75, 65, 55, 45, 35, 25, 15, 5\}$. When the Proportion offered to the responder from the proposer is lower comparatively, the responder rejects at most time. Otherwise, offers are rarely rejected. Using the experiment results of Camerer, the expected payoffs of the proposer of every strategy are listed as Tab.1.

As Tab.1, the strategies with lower offer such as S_1 and S_2 have higher mean payoffs, but the expected payoffs of the proposer are quite low due to the high rejection rate. The rejection rate decreases with the increasing division proportion, which leads to increase the expected payoffs. The expected payoff reaches maximum at S_5 but then decreases. So, S_5 should be the optimal strategy according to the expected payoffs. But people may not choose the optimal strategy due to the temptation of the average payoffs at first. They may adjust their strategies gradually through learning and fair thinking to reduce the choice probability of strategy S_1 or S_2 , and increase the choice probability of fair strategy such as S_5 relatively. So S_5 becomes the equilibrium strategy of the game.

We use EWA learning model and FGL model to simulate the ultimatum game separately. Suppose in the first period ($t=0$), the choice probabilities of the strategies ($S_1, S_2, \dots, S_{10}$) of proposers are $(65\%, 25\%, 5\%, 3\%, 2\%, 0, 0, 0, 0, 0)$. When the division proportion is lower than 40%, the responder refuses the propose, (supposed by the experiment results of Roth). Thus the initial attractions of the strategies are: $A^1(0) = 136.371$, $A^2(0) = 115.027$, $A^3(0) = 79.0786$, $A^4(0) = 67.6683$, $A^5(0) = 58.629$, $A^6(0), \dots, A^{10}(0) = 0$, and the response sensitivity $\lambda = 0.0407$.

All the values of parameter δ, ρ, ϕ are between 0 to 1 but they all depend on the actual circumstances.

When the players are changed, the parameter will change. According to the documents of Camerer, δ is the weight placed on forgone payoffs. In the ultimatum game, the opportunity payoffs of the forgone strategy is low, as the strategy's expected payoff is lower. Suppose $\delta = 0.36$. ρ controls the rate at which attractions grow, which implicates the learning ability of participants. While in general, it is between 0.65 and 0.85, here we suppose $\rho = 0.82$. ϕ reflects the decay of previous attractions, which decays slowly in general. Also we suppose $\phi = 0.97$. $N(0)$ is the strategy's initial weight, supposing $N(0) = 1$.

We calculate updated attraction of each strategy by the EWA learning model and FGL model in period 2 ($t=2$). Under the EWA model, according to the initial values and the parameters' values above, taking the expected payoffs in Tab.1 as

the material payoffs, we calculate attractions and choice probabilities of strategies by equation (1), see Tab.2. We can get the choice probabilities of strategies in ten periods through iterative updating calculation, see Fig.1.

The processes simulated by our model are as follows. We can calculate each strategy's fair payoffs as equation (4) due to the expected payoff as Tab.1 at first.

Using the fair payoffs instead of the expected payoffs, we update each strategy's attraction by equation (5), and then calculate its choice probability by equation (3). see Tab.4. We can get the choice probabilities of the strategies in next nine periods through iterative updating calculation, see Fig.2.

From Fig.1 and Fig.2, we find both models reflect the decrease of S_1 's choice probability and the increase of S_5 's with the game. Figure 1 shows that, under the EWA learning model, the choice probability of S_5 exceeds the choice probability of S_1 for the first time in period 6. However, Fig. 2 shows that, under the FGL model, the choice probability of S_5 has exceeded the choice probability of S_1 for the first time in period 4, and the proposer has no more tendency to change the strategy from this on. The choice probability of this strategy becomes higher. The results show that both models can simulate the changing process of choosing strategy. But FGL model converges to equilibrium strategy S_5 earlier than EWA learning model, which implicates our FGL model has stronger forecast ability.

<Insert Figure 1 here>

<Insert Figure 2 here>

4. Conclusion

As participants have both learning and fair thinking in repeated games, put the fair factors into learning model can predicate equilibrium more accurately, and solve the problem that the players learn slowly and disefficiently in learning models. The main conclusions of this paper are as follows:

- (1) Putting the fair factor into EWA model, and replacing material payoffs with psychological effect function, FGL model revise strategies' payoffs and attractions.
- (2) Compared the EWA model with FGL model through the simulation results of ultimatum game, the latter model can predicate the choice probability of each strategy and equilibrium more accurately.
- (3) Comparing FGL with the EWA according to the experiment results, both of the models can capture players' decision process accurately. The equilibrium converges to strategy S_4 simulated according to experiment 1, and which is S_5 according to experiment 2. However, the decision process simulated by our model is closer to the actual process of players, so FGL model has stronger explanatory and forecasting power than the EWA learning model.

References

- Arrow, Kenneth J. *Optimal and voluntary income redistribution*. In: Rosenfield, Steven (ed). *Economic welfare and the economics of soviet socialism: essays in Honor of Abram Bergson*. (1981). Cambridge University Press.
- Berg, Joyce, John Dickhaut., and Kevin McCabe. (1995). Trust, reciprocity and social history. *Games and Economic Behavior*. 10: 122-142.
- Berger, U. (2005). Fictitious play in $2 \times n$ games. *Journal of Economic Theory*. 129:139-154.
- Brown, G.W. (1951). *Iterative solution of games by Fictitious play*. In: *Activity Analysis of production and allocation* (T.C. Koopmans, Ed.), pp:374-376. Wiley: New York.
- Camerer, CF., TH Ho., and JK Chong. (2001). Models of Thinking, Learning and Teaching Games. *American Economic Review*, Vol. 93, No. 2 (May), pp.192-195.
- Camerer, CF., TH Ho., and JK Chong. (2004). Behavioral Game Theory: Thinking, Learning and Teaching"[J], paper presented at the Nobel Symposium 2001, Essays in Honor of Werner Guth (Steffen Huck, Ed.). pp. 119-179.
- Camerer, Colin F., and Teck H-Ho. (1999). EWA learning in Normal-form Games. *Econometrica*. 67:827 -873.
- Camerer, Conlin F. (2003). *Behavioral Game Theory*. The University of Princeton Press. New Jersey.
- Camerer, Colin F., and Thaler, Richard H. (1995). Ultimatums, dictators and manners. *Journal of Economic Perspectives*. (9):209-219.
- Crawford, Vincent P. (1995). Adaptive Dynamics in Games coordination. *Econometrica*. January. 63:103-43.
- Fehr Ernst., and Jean-robert Tyran (2008). Limited Rationality and Strategic Interaction: the Impact of the Strategic Environment on Nominal Inertia. *Econometrica*, 76(2): 353-394.
- Fehr. Ernst, Alexander Klein., and Klaus M. Schmidt. (2007). Fairness and Contract Design. *Econometrica*. 75/1:121-154.
- Forsythe, Robert L., Joel Horowitz, N E Savin., and Martin Sefton. (1994). Fairness in simple bargaining games. *Games*

and *Economic Behavior*, (6):347-369.

Gale, J., K. Binmore., and L.Samuelson. (1995). Learning to be imperfect:the ultimatum Games. *Games and Economic Behavior*, 8:56-90

Guth,W., R. Schmittberger., and B.Schwarze. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*. 3: 367-88.

Hon-snr,S., Monderer,D., and Sela,A. (1998). A learning approach to auctions. *Journal of Economic Behavior*. 5:368-386.

Milgrom,P., and Roberts,J. (1991). Adaptive and sophisticated learning in normal formgames. *Games and Economics Behavior*. 3:82-100.

Qingquan He & Yulei Rao. (2009). Fairness thinking and behavioral learning:an experiment test about social preference.*Journal of Hunan agricultural university(social sciences)*.10(3):47-52.

Rabin, Matthew. (1993). Incorporating Fairness into game theory and economies. *American Economic Review*, 83.

Roth A., and I. Erev. (1995). Learning in extensive form games:experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*. 8:163-212.

Roth, A. (1995). Bargaining experiments. *Handbook of experimental economics*. Princeton University Press.

Samuelson. (1993). Paul Altruism as a problem involving group versus individual selection in economies and biology. *American Economic Review*, (83): 143-148.

Sela, A. (2000). Fictious play in 2×3 games. *Games and Economic behavior*. 31:152-162.

Sen, Amartya. (1995) *Moral codes and economic success*. Edward Eldar Press.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics* 69(1): 343.

Notes

Note 1. See Gale,Binmore and samuelson (references[16]), Rothand Erev (references[16]).

Note 2. Attractions represent the degree of strategies attracting players to choose, and it is positive correlative with the choice probabilities of strategies.

Note 3. Camerer denoted $N(t) = \rho N(t-1) + 1$ in References [18][21], but $N(t) = N(t-1)\rho(1-\rho) + 1$ in References [22][23]. The later is adopted in our paper.

Note 4. As fairness model can only be applied to games with two players, but EWA learning model can be applied to multiplayer games($n \geq 2$), we only take $n=2$ into account to keep the consistence of both models, that is, there are only player i and j .

Note 5. See [20]: Rabin,Matthew(1993). Incorporating Fairness into game theory and economies. American Economic Review.83

Note 6. For Rabin's fairness model just applies two players only, I just choose the EWA learning model where $n=2$ to construct the fairness equilibrium learning model.

Table1. The strategy and its corresponding expected payoff

Strategy	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈	S ₉	S ₁₀
Division Proportion (%)	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
Average Payoff	95	85	75	65	55	45	35	25	15	5
Reject Proportion of Responders	0.95 ^a	0.78	0.57	0.35	0.07	0	0	0	0	0
Expected Payoff	4.75 ^b	18.7	32.6	42.25	51.15	45	35	25	15	5

Note: a: data come from the experiment of Camerer in 2003;

b: Expected Payoff = Average Payoff *(1-Reject Proportion of responders)

Table 2. Attractions and choice probabilities of strategies of EWA model in period one

Strategy S_k	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Attraction value $A^k(1)$	114.07	100.72	75.30	68.83	64.09	13.79	10.73	7.66	4.60	1.53
Choice probability $P^k(1)$	0.54	0.274	0.076	0.054	0.043	0.003	0.002	0.002	0.002	0.001

Table 3. The fair payoffs of strategies of the proposer

strategy S_k	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
fair income	4.28	18.36	32.19	41.99	51.16	45.52	35.25	25.19	15.5	12.25

Table 4. Attractions and choice probabilities of strategies of FGL model in period one

strategy S_k	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
$A^k(1)$	113.93	100.62	75.17	68.75	64.09	13.95	10.80	7.72	4.75	3.75
$P^k(1)$	0.539	0.275	0.076	0.055	0.043	0.003	0.002	0.002	0.002	0.002

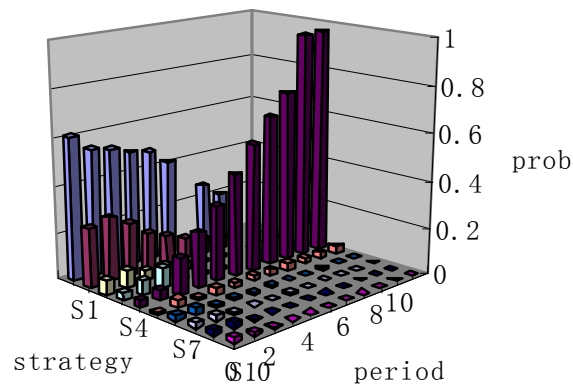


Figure 1. Strategies of proposers simulated by EWA learning model in ultimatum game

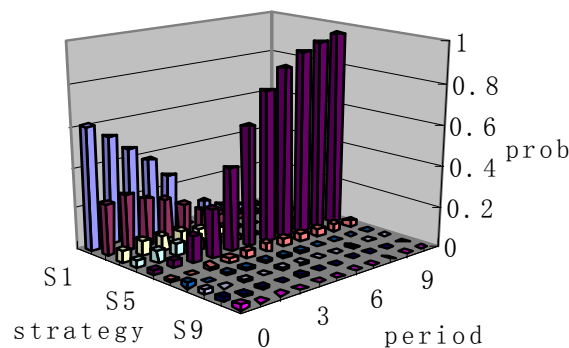


Figure 2. Strategies of proposers simulated by FGL model in ultimatum game