Honam Mathematical J. **35** (2013), No. 2, pp. 225–234
http://dx.doi.org/10.5831/HMJ.2013.35.2.225

AN APPROACH TO THE PROBLEM OF COMMON POOL RESOURCES THROUGH AN EXTENSION OF THE EQUILIBRIUM CONCEPT†

Jaegug Bae, Jongseok Kim*, and Eun Sook Kang

Abstract. Many studies of experimental economics have produced outcomes which contradict the predictions of Nash equilibrium, which relies heavily upon the premise of selfishness of an individual. In the games involving contexts of social conflicts represented by the prisoners' dilemma game, the experiments yields outcomes quite different from what are predicted by the conventional wisdom. In order to fill this gap between the conventional Nash Equilibrium and experimental outcomes, non-selfish (or other-regarding) motives of human behavior are introduced and then a new equilibrium concept, RAE-equilibrium is developed. It is also proved that an RAE-equilibrium exists under quite general conditions. Then it is applied to the prisoners' dilemma game that some of the experimental outcomes can be explained.

1. Introduction

The intention of a country or an institution that provides common pool resources is to manage and/or maintain ways to enable many people to use the commons effectively for a long period of time. However, it is a well-known fact that the use by unspecific multi-parties without any restriction causes the economic inefficiency of resources under the abuse in so-called 'the tragedy of the commons'. Existing studies on the issue of abusing commons have three analysis models: the tragedy of the commons, the prisoner's dilemma game and the logic of collective

Received March 25, 2013. Accepted June 5, 2013.
2010 Mathematics Subject Classification. Primary 91A10; Secondary 91A40.
Key words and phrases. non-selfishness, RAE-equilibrium, prisoners' dilemma game, social conflict or social dilemma.
†This work was supported by the National Research Foundation of Korea Grant funded by Korean Government (KRF-2009-32A-B00243).
*Corresponding author

action. And the critical part of these models is ultimately the free-rider problem ([9, Chap. 1], [10]). If there is no guidance or management, the game situation where each individual attempts to maximize its own profit, as shown in the prisoner's dilemma game, brings out an inefficient equilibrium selection in overall position. As the resolution plan of such a game situation, a number of scholars have proposed approaches through systematic studies of institutionalism such as social system, governance, self-regulated rules ([7], [8], [11]) and diverse deformations of repeated games, evolutionary stability of equilibrium, as well as the reflection on existing game theory based on reasonable selfishness ([1], [2], [3], [4], [6]). In this paper, such reflection is contemplated to extend the equilibrium concept of existing game theory in a way of presenting new alternative on the issue of commons.

2. An extension of the equilibrium concept

In the history of the game theory, the verification of the existence of the Nash equilibrium is rather monumental. Thereafter, all developments of equilibrium concepts has the basic premises of Nash equilibrium at all times. The Nash equilibrium is the equilibrium to enable each player to select a strategy to maximize its own interest once the strategy of other players are set under the reasonable selfishness of human. Obviously, it could not be said that all humans select only a strategy that maximize its own profit in accordance with the cultural environment that formulates the player's interest or depending on the game situation. Notwithstanding the foregoing, the purpose of the Nash equilibrium concept is clear. In other words, human has the tendency to maximize its economic value under any circumstance in general and on the other hand, it is difficult to make mathematical model for the scale of non-economic preferential value under the cultural bias of each individual. Therefore, it is inevitable to have restrictive limit in applying the equilibrium concept devoid of the consideration of cultural and innate preferential value of human to the actual game social environment. In fact, in several studies of the game situation for the survival of the fittest or social game situation under the natural ecosystem, such an innate limitation of the Nash equilibrium concept has been pointed out ([1], [2], [4], [5]), and in addition, studies have been undertaken to describe such phenomenon with reciprocal altruism theory, group selection theory, homo-reciprocan concept and others.

2.1 Righteousness and altruism

Human is basically born with the selfishness, but by intrinsically and by effect of sociocultural education, it is a being with the sensitivity on good and justice, desire to be righteous and desire for affection and consideration on others. Such a human attribute that the existing game theory overlooks is sought herein in the aspects of righteousness and altruism. Here, righteousness is the concept that embraces fairness and justice and it takes a look at the game situation only under the point of view of overall righteousness including itself by completely excluding the point of view in maximizing self interest. Altruism is the point of view based only on maximizing the interest of others.

Degrees of righteousness and altruism are to be written as λ and μ, respectively where $0 \leq \lambda,\ \mu \leq 1$. These are independent variables. However, under the game situation, manifestation of the righteousness could be impacted on the degree of altruism. The player with $\lambda = \mu = 0$ has the interest only on maximization of its own profit. In the repeated prisoner's dilemma game, the player of $\mu = 1$ would select 'Cooperation' at all times while the player of $\lambda = 1$ may use the Tit-for-Tat strategy. However, the player with $\lambda = \mu = 1$ also uses the strategy to select 'Cooperation' at all times. In other words, manifestation of righteousness is controlled by the degree of altruism without mutually conflicting.

2.2 Measurement of righteousness and altruism

Altruism may be relatively simply measured. For example, let's assume that each player is paid for certain amount of money equally and collect public fund. At this time, the amount dividing the contribution amount of each player with the paid amount can be considered as the degree of altruism of the player. Of course, the degree of altruism of each player has to be considered as differed depending on the game situation and player group. This is attributable to the fact that the degree of altruism of each will be different depending on whether the player set is family member, meeting of friends or meeting of unknown people. In the meantime, the measurement of degree of righteousness may be measured through public goods game, ultimate game and others but it requires certain care. The ultimate game is the game to propose to give certain amount to the player A and the player provides part of the amount to the player B. At this time, if the player B accepts the proposed amount, the player B keeps the amount and the rest of the amount will be kept by the player A. If the player B rejects the proposed amount, neither player may keep any money. By surveying the upper limit amount that the first player proposes, and critical value of the amount refusing the proposal, the degree of righteousness may be measured, but the reaction of players responding to such experiment may be influenced by the degree of altruism that a careful experiment design may be required. However, unlike the degree of altruism, the degree of righteousness is the expression of one's fair and righteous spirit that it would be reasonable to assume that it does not differ in the type of player group.

2.3 RAE-equilibrium

Throughout the paper, we denote the set of real numbers by the character $\mathbf{R}$.

Definition 2.1. Let $I = \{1, 2, \cdots, n\}$ be a set of players where the degree pair of righteousness and altruism of each player i is (λ_i, μ_i). Then $\Phi_i,\ S_i = \Delta(\Phi_i)$, represents the set of pure strategies and the set of mixed strategies of the player i, respectively and $S = \prod S_i$ is the set of mixed strategy combinations. When the payoff function is given as

$$g : S \to \mathbf{R}^n \qquad g(s) = (g_1(s), g_2(s), \cdots, g_n(s)),$$

for a mixed strategy combination $s = (s_1, s_2, \cdots, s_n) \in S$, we define the set $R_i(s)$ of righteous responses, the set $A_i(s)$ of altruistic responses and the set $E_i(s)$ of egocentric responses for the player i, respectively as followings:

$$R_i(s) = \left\{ \rho_i \in S_i \mid \forall \delta_i \in S_i,\ \sum_j g_j(\rho_i, s_{-i}) \geq \sum_j g_j(\delta_i, s_{-i}) \right\}$$

$$A_i(s) = \left\{ \alpha_i \in S_i \mid \forall \delta_i \in S_i,\ \sum_j g_j(\alpha_i, s_{-i}) \geq \sum_j g_j(\delta_i, s_{-i}) \right\}$$

$$E_i(s) = \left\{ \epsilon_i \in S_i \mid \forall \delta_i \in S_i,\ \sum_j g_j(\epsilon_i, s_{-i}) \geq \sum_j g_j(\delta_i, s_{-i}) \right\}.$$

Now, for a mixed strategy combination $s \in S$, let's consider the player i's best response (RAE-response) that counts his or her righteousness, altruism and selfishness together. A simple linear combination

$$(1 - \mu)\lambda\rho + \mu\alpha + (1 - \lambda - \mu + \lambda\mu)\epsilon$$

of responses ρ, α, ϵ (deleting subscripts) according to the degrees of righteousness and altruism might appropriately compromise the selections, but the payoff of the player on this strategy may completely not to do this way. Namely, even if the degrees of righteousness and altruism are very small, $(1-\mu)\lambda\rho + \mu\alpha$ may effect for critical influence on the payoff of the player. Therefore, the best response would be more reasonable to define with the strategy to maximize the newly defined preferential level

$$(1-\mu_i)\lambda_i \sum_j g_j(s) + \mu_i \sum_{j \neq i} g_j(s) + (1 - \lambda_i - \mu_i + \lambda_i\mu_i) g_i(s)$$

for the player i. However, the ingredient of degree of righteousness, $\sum_j g_j(s)$, contains only the utilitarian implication without containing the implication of fairness. Therefore, the standard deviation with the multiplication of $-\lambda\mu$ is added to reflect the element of fairness. In conclusion, the preference function of each player i is defined as follows:

$$\begin{aligned} f_i(s) =& (1-\mu_i)\lambda_i \sum_j g_j(s) - \lambda_i\mu_i \sqrt{Var(g(s))} \\ &+ \mu_i \sum_{j \neq i} g_j(s) + (1 - \lambda_i - \mu_i + \lambda_i\mu_i) g_i(s) \\ =& (1-\mu_i) g_i(s) + (\lambda_i + \mu_i - \lambda_i\mu_i) \sum_{j \neq i} g_j(s) - \lambda_i\mu_i \sqrt{Var(g(s))}. \end{aligned}$$

Here, $Var(g(s))$ is the variance of $g_1(s), g_2(s), \cdots, g_n(s)$, i.e.,

$$Var(g(s)) = \frac{1}{n} \sum_{j=1}^{n} \left(g_j(s) - E[g(s)]\right)^2, \quad E[g(s)] = \frac{1}{n} \sum_{j=1}^{n} g_j(s)$$

Now, the best response of the player counting its righteousness and altruism, namely, RAE-response is defined as the strategies to maximize the preference function.

Definition 2.2. Let the preference function of the player i be

$$\begin{aligned} f_i(s) =& (1-\mu_i) g_i(s) + (\lambda_i + \mu_i - \lambda_i\mu_i) \sum_{j \neq i} g_j(s) \\ & - \lambda_i\mu_i \sqrt{Var(g(s))}, \qquad s \in S. \end{aligned} \tag{1}$$

For a mixed strategy combination $s \in S$, an RAE-response of the player i is a mixed strategy $\sigma_i \in S_i$ to maximize $f_i(\sigma_i, s_{-i})$ whence the set of

RAE-responses of the player i is

$$RAE_i(s) = \{\sigma_i \in S_i \mid \forall \delta_i \in S_i,\ f_i(\sigma_i, s_{-i}) \geq f_i(\delta_i, s_{-i})\}.$$

When the strategy s_i^* of each player of the mixed strategy combination $s^* = (s_1^*, s_2^*, \cdots, s_n^*)$ satisfies the reflexive property on the RAE-response, i.e.,

$$\forall i \in I,\ s_i^* \in RAE_i(s^*)$$

s^* is said to be an RAE-equilibrium and for given game G, $RAEE(G)$ denotes the set of all RAE-equilibria.

Theorem 2.3. *Any finite n-person normal form game*

$$G = (\Phi_1, \Phi_2, \cdots, \Phi_n, g_1, g_2, \cdots, g_n)$$

possess at least one RAE-equilibrium.

Proof. Let $g(s) = (g_1(s), g_2(s), \cdots, g_n(s))$ be a payoff vector for $s \in S$. For $s_i, s_i' \in S_i$ and $t \in [0,1]$, $g_i : S \to \mathbf{R}$ satisfies

$$g_i(ts_i + (1-t)s_i', s_{-i}) = tg_i(s_i, s_{-i}) + (1-t)g_i(s_i', s_{-i}) \tag{2}$$

and is a continuous function on t. By definition, the RAE-equilibrium is the Nash equilibrium of the new game $G_f = (S_1, S_2, \cdots, S_n, f_1, f_2, \cdots, f_n)$ with the preference function of (1) for the payoff function. However, each f_i is not a linear combination of g_j's that relation (2) does not hold for f_i. Therefore, in order to verify the existence of an RAE-equilibrium, it requires the result of J. B. Rosen that extends the existence theorem of the Nash equilibrium. By Rosen's theorem ([12], [13, Theorem 2.1.1]), it suffices to show that each f_i is concave. Namely, it is to be shown that $\zeta''(t) \leq 0$ where

$$\zeta : [0,1] \to \mathbf{R}, \quad \zeta(t) = f_i(ts_i + (1-t)s_i', s_{-i})$$

for arbitrary fixed $s_i, s_i' \in S_i,\ s_{-i} \in S_{-i}$. For the notational convenience, let us write $g(t) = g(ts_i + (1-t)s_i', s_{-i}),\ g_j(t) = g_j(ts_i + (1-t)s_i', s_{-i})$. By (1) with the application of (2), we obtain obviously that $\zeta''(t) = -\lambda_i \mu_i \frac{d^2}{dt^2}\sqrt{Var(g(t))}$. Hence what we have to prove is that

$$\frac{d^2}{dt^2}\sqrt{Var(g(t))} \geq 0.$$

If we write $Var'(g(t)) = \frac{d}{dt}Var(g(t)),\ Var''(g(t)) = \frac{d^2}{dt^2}Var(g(t))$, then,

$$\frac{d^2}{dt^2}\sqrt{Var(g(t))} = \frac{Var(g(t))Var''(g(t)) - (1/2)(Var'(g(t)))^2}{2(Var(g(t)))^{3/2}}.$$

Thus it would be sufficient to prove that

$$\frac{1}{2}(Var'(g(t)))^2 \leq Var(g(t))Var''(g(t)). \tag{3}$$

Here, let X and Y be random variables on $\{1, 2, \cdots, n\}$ such that $X(j) = g_j(t),\ Y(j) = g_j'(t),\ 1 \leq j \leq n$. Then $Var(g(t)) = Var(X)$,

$$\begin{aligned}\frac{1}{2}Var'(g(t)) &= \frac{1}{n}\sum_{j=1}^{n}(g_j(t) - E[g(t)])(g_j'(t) - E[g'(t)]) \\ &= \sum_{j=1}^{n}\frac{g_j(t)g_j'(t)}{n} - E[g(t)]E[g'(t)] = E[XY] - E[X]E[Y]\end{aligned}$$

and

$$\frac{1}{2}Var''(g(t)) = \sum_{j=1}^{n}\frac{(g_j'(t))^2}{n} - (E[g'(t)])^2 = E[Y^2] - (E[Y])2 = Var(Y).$$

Therefore inequality (3) is equivalent to

$$\frac{(E[XY] - E[X]E[Y])^2}{Var(X)Var(Y)} \leq 1 \quad \text{or} \quad -1 \leq \frac{E[XY] - E[X]E[Y]}{\sqrt{Var(X)}\sqrt{Var(Y)}} \leq 1.$$

The middle term of the last inequality represents exactly the correlation coefficient of two random variables X and Y, whose absolute value known to be less than or equal to 1 at all times that inequality (3) is verified. □

3. Application of the RAE-equilibrium in the prisoner's dilemma game

In this section we apply new concept of RAE-equilibrium to the following specific prisoner's dilemma (bimatrix) game:

$$I = \{1, 2\}, \quad \Phi_1 = \Phi_2 = \{D, C\}, \quad \begin{bmatrix}(1,1) & (3,0)\\ (0,3) & (2,2)\end{bmatrix}.$$

For a mixed strategy combination $s = (s_1, s_2)$ with $s_1 = \alpha D + (1 - \alpha)C,\ s_2 = \beta D + (1 - \beta)C$, the payoff functions would be

$$\begin{aligned} g_1(s) &= \alpha\beta + 3\alpha(1-\beta) + 2(1-\alpha)(1-\beta) = \alpha - 2\beta + 2,\\ g_2(s) &= \alpha\beta + 3(1-\alpha)\beta + 2(1-\alpha)(1-\beta) = -2\alpha + \beta + 2.\end{aligned}$$

from whence the preference functions follow:

$$\begin{aligned}f_1(s) =&(1-2\lambda_1-3\mu_1+\lambda_1\mu_1)\alpha+(-2+\lambda_1+3\mu_1-2\lambda_1\mu_1)\beta\\&+2(1+\lambda_1+\lambda_1\mu_1)-(3/2)\lambda_1\mu_1|\alpha-\beta|,\\f_2(s) =&(-2+\lambda_2+3\mu_2-2\lambda_2\mu_2)\alpha+(1-2\lambda_2-3\mu_2+\lambda_2\mu_2)\beta\\&+2(1+\lambda_2+\lambda_2\mu_2)-(3/2)\lambda_2\mu_2|\alpha-\beta|.\end{aligned}$$

First, we assume that $\alpha \geq \beta$. Then

$$\frac{\partial f_1}{\partial \alpha}=1-2\lambda_1-3\mu_1-\frac{1}{2}\lambda_1\mu_1,\quad \frac{\partial f_2}{\partial \beta}=1-2\lambda_2-3\mu_2-\frac{5}{2}\lambda_2\mu_2.$$

Thus for given fixed β, the RAE-response of player 1 is $\alpha=1$ if $\frac{\partial f_1}{\partial \alpha}>0$ and $\alpha=\beta$ if $\frac{\partial f_1}{\partial \alpha}<0$ and any $\alpha\in[\beta,1]$ if $\frac{\partial f_1}{\partial \alpha}=0$. On the other hand, for given fixed α, the RAE-response of player 2 is $\beta=\alpha$ if $\frac{\partial f_2}{\partial \beta}>0$ and $\beta=0$ if $\frac{\partial f_2}{\partial \beta}<0$ and any $\beta\in[0,\alpha]$ if $\frac{\partial f_2}{\partial \beta}=0$. Note that $\frac{\partial f_1}{\partial \alpha}<0 \Leftrightarrow \mu_1>-4+\frac{26}{\lambda_1+6}$ and $\frac{\partial f_2}{\partial \beta}<0 \Leftrightarrow \mu_2>\frac{4}{5}+\frac{14}{5(5\lambda_2-6)}$. The discussion is very symmetric for $\alpha\leq\beta$, that the following theorem is obtained.

Theorem 3.1. *For the unit square $D=[0,1]\times[0,1]$, consider two subsets*

$$\begin{aligned}U_1&=\left\{(x,y)\in D\,|\,y>-4+\frac{26}{x+6}\right\},\\U_2&=\left\{(x,y)\in D\,|\,y>\frac{4}{5}+\frac{14}{5(5x-6)}\right\}.\end{aligned}$$

In the foregoing prisoner's dilemma game, if the degree pair of righteousness and altruism for a player, $(\lambda,\mu)\in U_1$ and that of the other player, $(\lambda,\mu)\in U_2$, then (C,C) is the unique RAE-equilibrium.

From the above, under the general game situations including the tragedy of commons or the prisoner's dilemma game, the equilibrium concept of game theory in consideration of righteousness and altruism is sought. As discussed in the prisoner's dilemma game, it obtains the result of conforming to the general common sense as meaningful to select mutual cooperation (C) if each player secures certain level of righteousness and altruism. As the RAE-equilibrium becomes exactly the Nash equilibrium when both degrees of righteousness and altruism are 0 that this equilibrium is the complete generalization of the Nash equilibrium. The intrinsic righteousness and altruism of human are the concepts devised to overcome the innate limitation of earlier mentioned Nash equilibrium that what we are confident is the fact that the RAE-equilibrium concept would be much more effective analysis tool in the analysis of social conflict, such as the issue of commons, or evolutionary stability of ecosystem. In order to qualitatively verify and support such fact, following research task required would be shown as follows.

- Review of diverse definition of righteousness
- Study to apply RAE-equilibrium concept in several types of existing game theory such as repeated game, evolutionary game theory and others
- Study to statistically analyze how effective to the description for social/ecologic phenomenon for the solution of game situation based on RAE-equilibrium concept
- Study on experimental design for the measurement of righteousness and altruism
- Comparative study of existing several equilibria and RAE-equilibrium through diverse game experiments

References

[1] J.-D. Choi, *The possibility of biological altruism (in Korean)*, Philosophical Study **64** (2004), 179-200.

[2] J.-K. Choi, *Game theory and Human nature (in Korean)*, Trends & Perspectives **41** (1999), 287-294.

[3] J.-K. Choi, *Trembles may support cooperation in a repeated prisoner's dilemma game*, Journal of Economic Behavior & Organization **63(3)** (2007), 384-393.

[4] J.-K. Choi, S. Bowles, *The coevolution of parochial altruism and war*, Science **318(5850)** (2007), 636-640.

[5] J.-K. Choi, *Appearance of Altruistic Human (in Korean)*, Roots & Leaves, Seoul, 2009.

[6] O. Gossner, T. Tomala, *Secret correlation in repeated games with imperfect monitoring*, Math. Oper. Res. **32(2)** (2007), 413-424.

[7] A. Lecours, *New institutionalism: Theory and Analysis*, University of Toronto Press, Canada, 2005.

[8] M. Lee, *Institution, common pool resource and governance (in Korean)*, Korean Journal of Public Admininstration **44(2)** (2006), 247-275.

[9] E. Ostrom, *Governing the Commons*, Cambridge University Press, New York, 1990.

[10] E. Ostrom, *A behavioral approach to the rational choice theory of collective action*, American Political Science Review **92(1)** (1998), 362-383.

[11] S. Robin, C. Staropoli, *Experimental methodology to inform new institutional economics issues (in New Institutional Economics, Chap. 7)*, Cambridge University Press, New York, 2008.

[12] J. B. Rosen, *Existence and uniqueness of equillibrium points for concave n-person games*, Econometrica **33** (1965), 520-534.

[13] E. Van Damme, *Stability and Perfection of Nash Equilibria*, Springer-Verlag, Berlin, 1991.

Jaegug Bae
Department of Applied Mathematics, Korea Maritime University,
#1 Dongsam-Dong, Yeongdo-Gu, Pusan 606-791, S. Korea.
E-mail: jgbae@hhu.ac.kr

Jongseok Kim
Department of International Trade and Economics,
Korea Maritime University,
#1 Dongsam-Dong, Yeongdo-Gu, Pusan 606-791, S. Korea.
E-mail: kimjs_jeje@naver.com

Eun Sook Kang
Department of Maritime Administration, Korea Maritime University,
#1 Dongsam-Dong, Yeongdo-Gu, Pusan 606-791, S. Korea.
E-mail: eskang@hhu.ac.kr