数理解析研究所講究録
第 1557 巻 2007 年 173-178

Reconsideration of the Definition of Commitment-Proof Agreements*

Takuya Masuzawa†

March 30, 2007

Abstract

We define a core-like solution concept of games in strategic form, called a commitment-proof agreement, to discuss free rider problems.

1 Introduction

The max-mini behavior, which can be track back to von Neumann and Morgenstern (1944), is regarded as fundamental in the theory of n-person cooperative games. This idea is extended to games without side-payments. In the α-coalitional game defined by Aumann and Peleg (1960), a coalition can take a payoff vector if they have a strategy combination that guarantees the payoff regardless of the others' strategy. A set of payoff vectors which any coalition cannot improve upon by itself in this meaning is called the α-core.

There have been, however, some critics against the max-min-behavior from economic point of view: Rosenthal (1971), Chander and Tulkens (1996), Crrarini and Marini (2004). In the theory of the α-core, we assume that a coalition objects to a proposal without exploiting any effort of the others. In a public good provision game, for example, they object if they can improve their payoff even without using any of public good the outsiders provide. On the other hand, many economists are interested in the problem arising from the relation between the non-excludability of public good and the exploitation of the others' effort, called "the free rider problem". The theory based on the max-mini behavior excludes any possibility to analyze such a problem while it is straightforward theoretically. Some alternatives have been, therefore, proposed.

We also try to define a core-like solution concept to analyze the free rider problem. Before distinguishing our theory from the previous ones, we need to

*The issues discussed here were given in the first half of the talk by the author in the workshop of the mathematical economics at Kyodai-kaikan (December 10, 2006). The authors' view on the issues discussed in the second half has been changed. It seems, thus, to be appropriate to discuss them in detail somewhere else.

†Faculty of Economics, Osaka University of Economics, 2-2-8, Osumi, Higashi-yodogawa-ku, Osaka-shi, Osaka-fu, Japan, E-mai: masuzawa(at)osaka-ue.ac.jp

clarify the interpretation of the core. In my point of view, it is not unique and sometimes, intentionally or not, treated ambiguously. We need, thus, to clarify what we intend by the core rather than to discuss what the core really means.

I think that, following Greenberg (1990, p.p.166–167), who initiates "theory of social situations", it is useful to introduce the difference between a prescriptive theory and a descriptive theory. I mean by a "prescriptive theory" a theory that provides possible ways to persuade the players for an active intervener. On the other hand by a descriptive theory I mean a theory that provides an expectation or an explanation of the result of a game for an observer detached from the players. In this paper, the core shall be regarded as a solution concept not for descriptive analyses but for prescriptive analyses. Note that what I mean by "prescriptive" is different from "normative". The word, "normative", comprises a judgment about good or evil in economics. We, by the core, discuss the possibility of intervene but says nothing about good or evil in most grave sense.

From this point of view, the emptiness or the largeness of the core is regarded to be not a pathological but a distinctive result of the theory. On the one hand, if the core of a game is empty then we find that any proposal to the players is rejected by some coalition and any intervene turns out a failure. On the contrary, the non-emptiness of the core implies that we can propose an outcome that cannot be rejected. Moreover, the largeness of a core implies that many outcomes can be proposed to the players and cannot be objected by any coalition.

Rosenthal (1971) argues that the core based on the max-mini behavior is not " intuitively stable" in the analyses of economic externality and proposes to assume that the complement coalition chooses a "coalition rational" or "individual rational" strategy after an improving coalition S chooses their strategy. Ichiishi (1993, p.p.66–67) argues that if the public good economy is discussed not by the α-core notion but by the strong equilibrium then we observe that we confront the free rider problem. Chander and Tulkens (1996) define another kind of coalitional game called the γ-coalitional game. They assume that a coalition can take a Nash equilibrium payoff vector of a game the player of which consists of all individual outsiders and the coalition treated as one player. Crrarini and Marini (2004) define another alternative by the sub-game perfect equilibrium of a game where first a coalition chooses their strategy and then outsiders choose their strategy individually and simultaneously.

Greenberg (1990, p.p.109–110) defines a solution concept through his initiated framework. He considers a situation where if an improving coalition, S, chooses a strategy then a new game the player of which is $N \setminus S$ arises. His solution is regarded as a possible way of successive intervention. First we confront original game and propose an outcome for players. If a coalition objects to it and commits itself to a strategy then the others confront a new game, called a "position", and propose an outcome of it. If, in the new game, we propose, to $N \setminus S$, a "stable" outcome such that S does not enjoy higher payoff than their objected agreement in the first game, then the objection of S fails and the initial proposal is stable in this sense. The stability is defined recursively.

A points of agreement of these studies is that the behavior of $N \setminus S$ is restricted, which makes it possible to reflect free rider problems. The benefit of Greenberg ' s solution is that his theory is consistent as a prescriptive theory and does not need a so strong assumption on the players' expectations of other's behavior. On the other hand, other solutions rely on the assumption that the players behave as Nash equilibrium describes. The behavior, "improving upon", is, however, not consistent with this assumption. The contradiction of such hybrid solutions has been ignored.

We must, however, point out that Greenberg's solution needs some modification. The first reason for this is that Greenberg's solution is not necessarily in the α-core, which is considered as a weak and fundamental condition for the stability of the core. The second reason is that an outcome is stable in Grennberg's sense whenever any sub-game has no stable outcome even if we restrict stable outcomes within the α-core. It is strange that having no counter objection gives a good strategy to the intervener. In this paper, we propose a modification of Greenberg's solution as to clear such problems and discuss an example.

2 Definition

We denote by $\Re$ the set of real numbers. Let $N = \{1, 2, \ldots, n\}$ be a set of players, any subset $S \subset N$ is called a coalition. By X^i we mean a set of strategies of available to $i \in N$ and by X^S we refer to a Cartesian product of X^i over S. Typical elements of X^S are denoted by w^S, x^S, y^S. A payoff function of i denoted by u^i is a function from X^N to $\Re$, and, for all $S \subset N$ we define u^S by $u^S := (u^i)_{i \in S}$. A strategic form game is specified by a list $G = (N, (X^i)_{i \in N}, (u^i)_{i \in N})$. Elements of $\Re^S$ are called payoff vectors and denoted by a^S, b^S, and c^S. We write that $a^S \geq b^S$ when, for all $i \in S$, $a^i \geq b^i$, and that $a^S \gg b^S$ when $a^i > b^i$ for all $i \in S$.

For $S \subset N$ and $x^S \in X^S$, we denote by $G|x^S$ a strategic form game such that a set of players are $N \setminus S$, a strategy of $i \in N \setminus S$ is X^i and the payoff of $i \in N \setminus S$ when $x^{N \setminus S}$ is chosen is defined by $u^i(x^S, x^{N \setminus S})$.

A payoff vector $a^S \in \Re^S$ is α-effective for S if there exists $x^S \in X^S$ such $a^S \leq u^S(x^S, y^{N \setminus S})$ for all $y^{N \setminus S} \in X^{N \setminus S}$. A coalition improves upon $a^N \in \Re^N$ via b^S if and only if b^S is α-effective and $b^S \gg a^S$. The α-core is a set of payoff vectors, a^S, which are α-effective for N and can not be improved upon by any coalition. We say that a strategy bundle $x^N \in X^N$ is an α-core strategy if and only if $u^N(x^N)$ is in the α-core.

The α-effectiveness gives a method to evaluate a strategy bundle of a coalition without any conjecture on the others action. A strategy is evaluated by payoff vectors guaranteed by it. This notion can be regarded as an extension of the max-min behavior to NTU game. Some economists criticize this idea and argue that some strategy of the outsiders' seems sometimes unrealistic and an evaluation by the α-effectiveness is too pessimistic. Such criticisms are, however, based only on some intuition and convention of people in the real world. I do not follow this criticism here because an intuition and a convention of people in an interactive decision environment should not be adopted as a principle without any restriction in game theory. On the contrary, they should be explained by the theory. It is only worthy of principle to assume no conjecture.

While the definition of the α-core is theoretically persuasive enough, we must admit that we need to define another core-like solution in order to discuss the effect of non-excludability of public good on the possibility of agreements.

A set of commitment-proof agreements, denoted by $\sigma(G)$, is defined recursively. For any one-person game, a strategy bundle is commitment-proof agreement if it maximizes the payoff of the unique player. For $N \geq 2$, y^N is a commitment-proof agreement of G if and only if there exists no $S \subset N$ and $x^S \in X^S$ such that either

1. $S = N$ and $u^N(x^N) \gg u^N(y^N)$, or
2. $S \neq N$,
   and, for all $x^{N\setminus S} \in \sigma(G|x^S)$, $u^S(y^N) \ll u^S(x^S, x^{N\setminus S})$.

The commitment-proof agreements can be easily interpreted from an intervener's point of view. If a coalition objects to a commitment-proof proposal and takes x^S then the intervener can choose a strategy, $x^{N\setminus S}$, which is not welcome to at least one member of S and is commitment-proof in a situation after x^S is chosen. The intervener can, then, persuade $N \setminus S$ to accept $x^{N\setminus S}$.

This definition is obtained by a modification of CSSB of coalition commitment situation by Greemberg (1990, p.p.109–110). Our definition become essentially equivalent if (2) is replaced by the following:

3. $S \neq N$, $\sigma(G|x^S) \neq \emptyset$,
   and, for all $x^{N\setminus S} \in \sigma(G|x^S)$, $u^S(y^N) \ll u^S(x^S, x^{N\setminus S})$.

We call the solution obtained by adding $\sigma(G|x^S) \neq \emptyset$ Greenberg' solution here. The difference is, however, not so trivial. In our definition, if $\sigma(G)$ is nonempty then for all subgame F, $\sigma(F)$ is also nonempty. On the other hand, for some subgame F,$\sigma(F)$ may be empty even if $\sigma(G)$ is nonempty. To see the adequateness of our definition, remember that we defined CPA from the intervener's point of view. It is not persuasive to regard an objection as ineffective only for the reason that the intervener cannot propose any stable outcome after the objection. In other words, we should consider the intervener to be able to propose a stable outcome only if he can propose a stable outcome for any subgame,

By simple observation, we see that the set of CPAs is a subset of α-core strategies but that of Greenberg's solutions is not necessarily so. In Masuzawa (2002), the author proposed another definition of the commitment-proofness. The author, first, thought that a commitment-proof agreement should be in the α-core, and the condition that it is in the α-core is simply added to the definition. The definition in Masuzawa (2002), however, does not exclude the possibility that an outcome is commitment-proof while a subgame has no commitment-proof agreement.

3 An Example

Consider the following example due to Shapley and Shubik (1969), called "the Lake". Every player has two possible strategies C, and D. For all player $i \in N$, the payoff function is defined by

$$u_i(x_1, \ldots, x_n) = \begin{cases} -d|\{j \in N : j \neq i, x_j = D\}| & \text{if } x_i = D \\ -c - d|\{j \in N : j \neq i, x_j = D\}| & \text{if } x_i = C, \end{cases}$$

where $0 < c, d$ such that c/d is not an integer. Let k be the least integer such that $c/d < k - 1$.

Then, three solutions, the α-core, Greenberg's solution, and CPA, give different outcomes.

Proposition 1 *In the Lake, the followings hold.*

1. *The α-core:*
   (a) *If $0 < |S| < k$, D^S is the unique α-core strategy of $G|x^{N \setminus S}$.*
   (b) *If $|S| \geq k$, a strategy bundle x^S is an α-core strategy of $G|x^{N \setminus S}$, if and only if $|\{i \in S | x_i = C\}| \geq k$, and $|\{i \in S | x_i = D\}| < k$.*
2. *Greenberg's Solution (G-CPA):*
   (a) *If $0 < |S| < k$, then D^S is the unique G-CPA of $G|x^{N \setminus S}$.*
   (b) *If $|S| = mk$ for some integer m, C^S is the unique G-CPA of $G|x^{N \setminus S}$.*
   (c) *If $mk < |S| < (m+1)k$ for some integer m, there exists no G-CPA of $G|x^{N \setminus S}$.*
3. *Commitment-proof agreement (CPA):*
   (a) *If $0 < |S| < k$, then D^S is the unique CPA of $G|x^{N \setminus S}$.*
   (b) *If $|S| = k$, C^S is the unique CPA of $G|x^{N \setminus S}$.*
   (c) *If $k < |S|$, there exists no CPA of $G|x^{N \setminus S}$.*

Strategy D is punishment-dominat over C. From the result of Masuzawa (2003), the α-coalitional game is ordinally convex and the α-core is large. Both a set of G-CPA and one of CPA is smaller than α-core in this game. If $m \geq 2$ and $|S| = mk$, $G|x^{N \setminus S}$ has a G-CPA while some subgames of it have no G-CPA. The method by Masuzawa (2002) also does not exclude this phenomenon because G-CPA of this game is in the α-core. On the contrary, if the number of the players is larger than k, the game has no CPA, which means that the intervener has no strategy to persuade the players to accept. This impossibility is regarded as a free rider problem.

Reference

Aumann, R,J., Peleg, B., (1960) Von Neumann-Morgenstern Solutions to Cooperative Games Without Side Payments *Bulletin of the American Mathematical Society.*

Chander, P., Tulkens, H., (1997) The core of an economy with multilateral environmental externalities. *International Journal of Game Theory* 26: 3 379–401

Currarini, S., Marini, M.,(2003) A Conjectural Cooperative Equilibrium for Strategic Form Games, in *Game Practice and the Environment*, Ed. Edgar Elgar, London, 2003.

Foley, D., (1970) Lindarl's Solution and the Core of an Economy with Public Goods. *Econometrica. 38.* 66–72.

Greenberg, J., (1990) *The Theory of Social Situations: A Game-Theoretic Approach,*Cambridge University Press, Cambridge.

Ichiishi, T., (1992) *The cooperative nature of the firm*, Cambridge University Press, Cambridge.

Masuzawa, T., (2002) *Commitment-Proof Agreements*, Keio University.

Masuzawa, T., (2003) Punishment Strategy Make the α-coalitional game ordinally convex and balanced. *International Journal of Game Theory 32* 479-483.

Rosenthal, R., (1971) External economies and cores. *Journal of Economic Theory* 3:182–188.

Rosenthal, R., (1972) Cooperative Games in Effectiveness Form. *Journal of Economic Theory 5.* 88-101.

Shapley, L., Shubik, M., (1969) On the Core of an Economic System with Externalities. *American Economic Review.59.* 678–684.

Von Neumann. J., Morgenstern. O., (1944) *Theory of Games and Economic Behavior* Princeton University Press, Princeton.