FARSIGHTED STABILITY IN PRISONER'S DILEMMA

Akihiro Suzuki
Yamagata University

Shigeo Muto
Tokyo Institute of Technology

Abstract   We study players' behavior in the prisoner's dilemma by using two stability notions: the stable set of von Neumann and Morgenstern with indirect domination and the largest consistent set defined by Chwe. Both notions assume possibility of sequential deviations and farsightedness of players. When players use only pure strategies, these two stability concepts provide us with the same outcome: { (Cooperation, Cooperation), (Defect, Defect)} when players behave independently; and { (Cooperation, Cooperation)} when players' joint, but not binding, moves are also considered. In the prisoner's dilemma with many alternatives such as the mixed extension of the prisoner's dilemma, the two notions produce completely different outcomes. In the stable set, every individually rational outcome could be stable in the former case; and only a Pareto efficient outcome is essentially stable in the latter case. The largest consistent set consists of all individually rational outcomes in either case.

1. Introduction

Recently, Chwe defined a new stability notion, called the largest consistent set (Chwe[1994]). The largest consistent set applies to situations in which players (or groups of players) act publicly with farsightedness. That is, players consider the possibility that once a player (or a group of players) shifts his/her strategy, another player (another group of players) reacts, a third player (or a third group of players) in turn shifts his/her strategy, and so on. A sequence of shifts might continue without limit. Hereafter shifts of strategies will be simply called "moves" or "deviations." Moves by groups of players are not binding; i.e., some players in a moving group may later deviate and make a new move with players within or even outside the group. A move by a group of players is called "binding" if no player can deviates after the move. The largest consistent set assumes that a deviation is deterred if a sequence of deviation ends with a stable outcome in which some players in the deviating group do not prefer. An outcome is stable if every possible deviation is deterred. Since whether an outcome is stable depends on whether other outcomes are stable, stability and instability of outcomes must be consistent as a whole. The largest consistent set provides a condition that guarantees the consistency.

The largest consistent set is, in the definition, very close to a classical solution concept in game theory, the stable set defined by von Neumann and Morgenstern [1953]. Though the original stable set of von Neumann and Morgenstern has been applied mainly to coalitional form games, Greenberg [1990] applied the concept to strategic form games. Though the stable set assumes myopic behavior of players, one may introduce farsighted behavior as done by Harsanyi [1974].

Then there remains only a small difference between the largest consistent set and the stable set with players' farsighted behavior. We hereafter call the latter set, the modified stable set. The difference is just whether deviating players are optimistic (the modified stable set) or pessimistic (the largest consistent set) concerning other players' deviations that might follow. In the modified stable set players deviate if there exists at least one sequence of ensuing

deviations in which they end up with better payoffs, while in the largest consistent set players deviate only when in every possible sequence of deviations they are better off in the end. The modified stable set is always contained in the largest consistent set.

Chwe presented interesting properties of the largest consistent set together with comparison with the modified stable set not only in abstract games but also in strategic and coalitional form games. To examine how useful the new stability concept, its applications to particular games are most desired, especially games that reflect conflicts often observed in our social/economic system. In this paper we pick up the prisoner's dilemma and examine what will come out through these stability concepts.

This paper thus has two purposes. The one is to see whether the two stability notions, the modified stable set and the largest consistent set, produce different outcomes in the prisoner's dilemma. The other is to examine whether the two stability notions give us new insights into the study of the prisoner's dilemma.

Principal findings are the following. When players use only pure strategies, they produce the same outcome: {(Cooperate, Cooperate), (Defect, Defect)} when only individual player's moves are assumed; and {(Cooperate, Cooperate)} when players' joint, but not binding, moves are also considered.

In the mixed extension of the game, however, the modified stable set and the largest consistent set provide us with completely different outcomes. When only individual moves are assumed, there exist infinitely many modified stable sets and any individually rational strategy pair is contained in at least one of them. A vector of two players' payoffs that gives them at least their maximin values is called individually rational; and a strategy pair is called individually rational if it gives an individually rational payoff vector. The largest consistent set is unique and consists of all individually rational strategy combinations. When joint, but not binding, moves are also assumed, modified stable sets are essentially singleton sets, each consisting of a Pareto efficient strategy combination, while the largest consistent set again consists of all individually rational strategy combinations.

We hereupon remark on the mixed extension. One may claim that it is not real to assume that mixed strategies are observed, since strategies actually taken are pure ones. We would say that the mixed extension is just an approximation of the prisoner's dilemma with many alternatives. Results similar to those obtained in the mixed extension hold in the prisoner's dilemma with finitely many alternatives.

Thus as to the paper's first purpose, we claim the following. In the mixed extension of the prisoner's dilemma or in the prisoner's dilemma with many alternatives, the largest consistent set does not provide a useful prediction concerning players' behavior. The modified stable set, however, gives us a very sharp and interesting prediction, i.e., only a Pareto efficient outcome results, in particular in the case where two players' joint, but not binding moves are assumed. Furthermore we claim from the stable set analysis that in the prisoner's dilemma two players' joint moves are indispensable to exclude non-Pareto optimal outcomes; but joint moves are not necessarily binding.

As to the paper's second purpose, analyses of the prisoner's dilemma have been done exclusively in terms of noncooperative game theory. It is well known that (Defect, Defect) is the unique Nash equilibrium if the prisoner's dilemma is played only once, though both players are better in (Cooperate, Cooperate). It holds that irrespective of whether players use only pure strategies or use also mixed strategies. Even if the prisoner's dilemma is played repeatedly, we did not go further if the number of repetition is finite. The unique equilibrium is a repetition of (Defect, Defect).

If the prisoner's dilemma is repeated infinitely many times, however, the Nash equilibrium gives us a completely different outcome. Repetition of (Cooperate, Cooperate)

could be equilibrium; but many other outcomes including repetition of (Defect, Defect) appear also as equilibrium outcomes. More precisely, all pairs of payoff vectors that are individually rational can be attained as equilibrium. This is called the Folk Theorem. Therefore the Nash equilibrium analysis succeeds in attaining a Pareto efficient outcome; but fails to give us a sharp prediction of players' behavior in the repetition of the prisoner's dilemma.

One way to eliminate repetition of (Defect, Defect) is to use the notion of renegotiation-proof equilibrium. The renegotiation-proof equilibrium is a refinement of Nash equilibrium; and allows players to communicate with each other at the beginning of each repetition and to avoid Pareto inferior Nash equilibrium. Players coordinate their choices not only in the current stage but also in future stages; but even if they reach an agreement on coordination, the agreement is not binding. That is, they freely deviate from the agreement and take other choices. Though van Damme [1991] succeeded in eliminating non-Pareto efficient outcomes by using one version, there exist many versions of the renegotiation-proof equilibrium, some of which fail to eliminate non-Pareto efficient outcomes. For example, see van Damme [1989].

Conceptually the stable set is completely different from the Nash equilibrium; but in the prisoner's dilemma the two notions give similar outcomes. If only individual player's moves are assumed, any individually rational strategy pair is contained in at least one modified stable set: the result is very similar to the Folk Theorem. When players' joint moves are allowed, the modified stable set gives us a sharper result than the notion of renegotiation-proof equilibrium: the modified stable set always gives us only Pareto efficient outcomes, while many versions of the renegotiation-proof equilibrium fail to do so.

The rest of the paper will be organized as follows. In the next section, definitions and notation will be given. In Section 3, we will describe the modified stable set and the largest consistent set in the prisoner's dilemma with pure strategies. In Section 4, these two sets will be studied in detail in the mixed extension. In Section 4, we will use a particular version of the prisoner's dilemma to make discussion as simple as possible. The two sets in a general version of prisoner's dilemma will be briefly discussed in the Appendix. The paper closes in Section 5 with some remarks.

2. Stable Set and Consistent Set

Let $G = (N, \{S_j\}_{j \in N}, \{u_j\}_{j \in N})$ be a strategic form game where $N = \{1, 2, \cdots, n\}$ is the set of players, S_i is the set of strategies of player i and u_i is this player's payoff function, $u_i : S = S_1 \times S_2 \times \cdots \times S_n \to R$ (the set of all reals). A subset of players $M \subseteq N$ is called a coalition.

For any two strategy combinations $s = (s_1, \cdots, s_n)$, $t = (t_1, \cdots, t_n) \in S$, we say s is induced from t via coalition $M \subseteq N$ if $t_i = s_i$ for all $i \in N \setminus M$; that is, the combination s is reached from t by moves only of players in M. It is easily seen from the definition that if s is induced from t via coalition M, t is induced from s via coalition M. Thus we denote this inducement relation by $t \leftrightarrow_M s$. We say s indirectly dominates t, denoted $s \succ t$, if there is a sequence of strategy combinations $t = s^0, s^1, \ldots, s^{p-1}, s^p = s$ and a sequence of sets of players $M^1, \cdots, M^p$, such that for all $j = 1, \cdots, p$, $s^{j-1} \leftrightarrow_{M^j} s^j$ and $u_i(s) > u_i(s^{j-1})$ for all $i \in M^j$. We sometimes say that s indirectly dominates t starting with M^1, denoted $s \succ_{M^1} t$, to specify the set of players which deviates first from t^1.

In particular, when every M^j is a singleton, we say s individually indirectly dominates t, denoted $s \rhd t$, and s individually indirectly dominates t starting with i,

---

[1] We may assume $M^{j-1} \neq M^j$ for all $j = 1, \cdots, m$ since consecutive moves by the same set of players can be combined into one move.

denoted $s \triangleright_i t$, in case we want to specify the player who first deviates from t. The term "individually" may be omitted unless any confusion results.

We hereupon remark that in the definition of indirect domination we implicitly assume that deviation by a group of players is not binding, i.e., some players in a deviating group may later make another deviation with players within or even outside the group. Similarly a single deviating player may later change his/her strategy again.

When $p = 1$ in the definition of indirect domination, we simply say that s directly dominates t. When we want to specify a deviating player (or a group of players), we say, for example, s directly dominates t via player i (or a group of players M). This direct domination was initially defined by von Neumann and Morgenstern [1953] in characteristic function form games, and later defined by Greenberg [1990] in strategic form games. Indirect dominations that are not direct ones will be called in the following indirect dominations with $p \geq 2$. It should be noted that the indirect domination defined above is borrowed from Chwe [1994]. Though Harsanyi [1974] first proposed the notion of indirect domination, his definition is slightly different.

A set of strategy combinations $K \subseteq S$ is a stable set w.r.t. $\triangleright$ (resp. $\succ$) if (1) for any $s, t \in K$, neither $s \triangleright t$ (resp. $s \succ t$) nor $t \triangleright s$ (resp. $t \succ s$), and (2) for any $t \in S \setminus K$, there exists $s \in K$ such that $s \triangleright t$ (resp. $s \succ t$). Property (1) and (2) are called internal and external stability, respectively. In what follows, stable sets w.r.t. $\triangleright$ or $\succ$ are sometimes simply called stable sets unless any confusion results.

The consistent set w.r.t. $\triangleright$ (resp. $\succ$) is a set $K \subseteq S$ satisfying the following two properties. (1) For any $s \in K$, any $r \in S$ and any $M \subseteq N$ such that $s \leftrightarrow_M r$, there exists $t \in K$ such that $[t = r$ or $t \triangleright r$ (resp. $t \succ r$)] and $u_i(s) \geq u_i(t)$ for at least one $i \in M$. (2) For any $t \in S \setminus K$, there exist $r \in S$ and $M \subseteq N$ with $t \leftrightarrow_M r$ such that for all $s \in K$ with $[s = r$ or $s \triangleright r$ (resp. $s \succ r$)], $u_i(s) > u_i(t)$ for all $i \in M$. We call property (1) and (2), internal and external consistency, respectively. The largest consistent set is a consistent set that contains all others [2].

It should be noted that in the definition of the consistent set w.r.t. $\triangleright$ (resp. $\succ$) if we replace "there exists" in (1) by "for all", and "for all $s \in K$" in (2) by "there exists $s \in K$", then we obtain the definition of the stable set w.r.t. $\triangleright$ (resp. $\succ$). Thus it can be shown that any stable set w.r.t. $\triangleright$ (resp. $\succ$) is contained in the largest consistent set w.r.t. $\triangleright$ (resp. $\succ$). See Chwe [1994, Proposition 3] for the proof. Moreover the largest consistent set is always unique, but stable sets are generally not unique.

We explain more in detail what external and internal stability (resp. external and internal consistency) imply in the definition of stable set (resp. consistent set).

First pick a stable set. Suppose players have common understanding that each strategy combination (for simplicity hereafter called outcome) inside the stable set is "stable" and that each outcome outside the stable set is "unstable". Here the "stability" means that no player or no group of players has an incentive to deviate from it, and the "instability" means that there is at least one player or one group of players that has an incentive to deviate from it. Then the external and internal stability guarantee that the common understanding is never disproved, and thus continues to prevail.

In fact, suppose the set is both internally and externally stable, and pick any outcome in the set. Then by the internal stability, no player (or no group of players) can be better off[3] by deviating from it even if ensuing deviations end up with his/her (or their) most favorable outcome inside the set. Thus no player deviates or no group of players reach an agreement to

---

[2] A union of any two consistent sets is a consistent set. See Chwe [1994, Proof of Proposition 1].

[3] As for a group of players, this means that there is at least one player who is not better off.

deviate; and thus each outcome in the set remains stable. A deviating player (or deviating players) may be better off if ensuing deviations end up with an outcome outside the set; but outcomes outside the set are commonly considered to be unstable. Thus a deviating player(s) can never expect that ensuing deviations end up with such an outcome.

Next pick any outcome outside the set. Then by the external stability, there exists at least one player who can become better if he/she deviates from it and ensuing deviations of other players end up with his/her most favorable outcome inside the set. As for the case of a group of players, all members in the group can be better off. The final outcome is considered to be stable since it is in the set. Hence a player or a group of players will deviate. Hence each outcome outside the set remains unstable.

It is to be noted that, in the internal and external stability of the stable set, a deviating player(s) expects that ensuing deviations will be the most favorable one to him/her (or them).

The consistent set has similar implication. The external and internal consistencyguarantee that the common understanding that each outcome inside the set is stable and each outcome outside the set is unstable is never disproved. The difference is whether a deviating player(s) has an optimistic view (the stable set) or a pessimistic view (the consistent set) concerning other players' deviations that may ensue. In the consistent set, a deviating player(s) expects that ensuing deviations will end up with the least favorable outcome to him/her (or them).

3. Stability in Prisoner's Dilemma

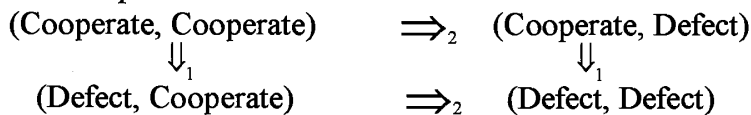The following payoff matrix describes the prisoner's dilemma.

1 \\ 2	Cooperate	Defect
Cooperate	a, a	0, b
Defect	b, 0	c, c

where a, b, c are reals satisfying $b > a > c > 0$.

We first present the stable set and the largest consistent set when two players use only pure strategies. Thus the set of strategy combinations is $S = \{$(Cooperate, Cooperate), (Cooperate, Defect), (Defect, Cooperate), (Defect, Defect)$\}$ where in each combination the former (resp. the latter) is player 1's (resp. 2's) strategy.

3.1. Stability without coalitional moves

If only individual moves are assumed, one may easily see that there is no indirect domination with $p \geq 2$. Direct dominance relations are summarized as follows.

$$
\begin{array}{ccc}
\text{(Cooperate, Cooperate)} & \Rightarrow_2 & \text{(Cooperate, Defect)} \\
\Downarrow_1 & & \Downarrow_1 \\
\text{(Defect, Cooperate)} & \Rightarrow_2 & \text{(Defect, Defect)}
\end{array}
$$

Here $\Rightarrow$ denotes dominance relations. For example, (Cooperate, Cooperate) $\Rightarrow_2$ (Cooperate, Defect) implies that (Cooperate, Defect) directly dominates (Cooperate, Cooperate) via player 2.

It is easily seen that the (unique) stable set w.r.t. $\triangleright$ and the largest consistent set w.r.t. $\triangleright$ are the same and given by the set $\{$(Cooperate, Cooperate), (Defect, Defect)$\}$.

3.2. Stability with coalitional moves

If coalitional moves are taken into consideration, (Cooperate, Cooperate) directly dominates (Defect, Defect) via coalition $\{1,2\}$; and thus we have the following chart of direct dominance relations. Since (Cooperate, Cooperate) $\succ$ (Cooperate, Defect), (Defect, Cooperate) and there is no other indirect domination with $p \geq 2$, the unique stable set w.r.t. $\succ$ and the largest consistent set w.r.t. $\succ$ are the same and given by the singleton set $\{$(Cooperate, Cooperate)$\}$.

(Cooperate, Cooperate) $\Rightarrow_2$ (Cooperate, Defect)

$\Downarrow_1$ $\Downarrow_1$

(Defect, Cooperate) $\Rightarrow_2$ (Defect, Defect)
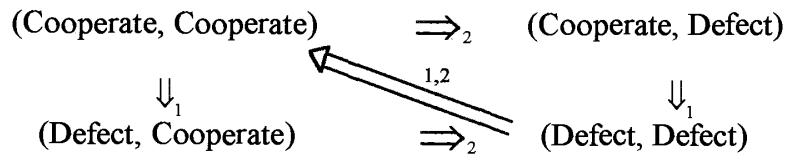
Hence if two players' coalitional, but not binding, moves are assumed and players are farsighted, the stable set and the largest consistent set both succeed in generating cooperation of players as the unique stable outcome. In the mixed extension, however, these two stability concepts produce completely different outcomes as will be shown in the next section.

4. Mixed Extension of Prisoner's Dilemma

We now study stable outcomes in the mixed extension of the prisoner's dilemma[4]. To make discussion as clear as possible, we will focus on the following particular version of the prisoner's dilemma. Similar results hold in the general prisoner's dilemma given in the previous section. Refer to the Appendix for detailed discussions.

2 1	Cooperate	Defect
Cooperate	4, 4	0, 5
Defect	5, 0	1, 1

In the mixed extension, we have $N = \{1,2\}$, $S_1 = S_2 = [0,1]$ where $s_1 \in S_1$ (resp. $s_2 \in S_2$) is the probability that player 1 (resp. 2) takes "Cooperate"; and thus through simple calculations payoff functions are given as $u_1(s_1,s_2) = 1 - s_1 + 4s_2$ and $u_2(s_1,s_2) = 1 - s_2 + 4s_1$. It is easily seen that the maximin payoffs to players 1 and 2 are both 1 in this game. We hereafter call a strategy combination that gives both players at least (resp. more than) their maximin payoffs, an individually rational (resp. a strictly individually rational) strategy combination.

In the next subsection, we study stable outcomes when only individual moves are assumed.

4.1. Stability without coalitional moves

Since all dominations are done via a single player, the term "individually" will be omitted in dominance relations throughout this subsection. For $t \in S$, let $\triangleright_i (t) = \{s \in S : t \triangleright_i s\}$ for $i = 1,2$. $\triangleright_i (t)$ is the set of strategy combinations that t indirectly dominates when player i deviates first. Then we have the following lemma concerning $\triangleright_1 (t)$ and $\triangleright_2 (t)$.

Lemma 4.1: Take any strategy combination $t = (t_1, t_2)$. Let $O = (0,0)$, $A = (1,0)$, $B = (0,1)$, $C = (1,1)$, $D = (1, t_2)$, $E = (1, 1/4 - t_1/4 + t_2)$, $F = (1/4 - t_2/4 + t_1, 1)$, $G = (t_1, 1)$, $H = (0, -t_1/4 + t_2)$, $I = (-t_2/4 + t_1, 0)$ and $G' = (t_1, 0)$. The point E (resp. H) is the intersection of player 1's isoprofit line with $u_1 = u_1(t)$ and with the vertical line $s_1 = 1$ (resp. $s_1 = 0$). Similarly the point F (resp. I) is the intersection of player 2's isoprofit line with $u_2 = u_2(t)$ and with the horizontal line $s_2 = 1$ (resp. $s_2 = 0$). Then $\triangleright_1 (t) = $ (triangle tDE) \ (line tE) and $\triangleright_2 (t) = $ (triangle tFG) \ (line tF)[5]. See Figure 1.

Proof: First it should be noted that in the region above (resp. to the right of) player 1's (resp. 2's) isoprofit line passing through t, player 1 (resp. 2) gains more than what he gains in t.

---

[4] As stated in Introduction, the mixed extension is used for an approximation of the prisoner's dilemma with many alternatives.

[5] In what follows, (triangle tDE) \ (line tE), (triangle tFG) \ (line tF tF), etc. will be simply written as tDE \ tE, tFG \ tF, etc. unless any confusion results.

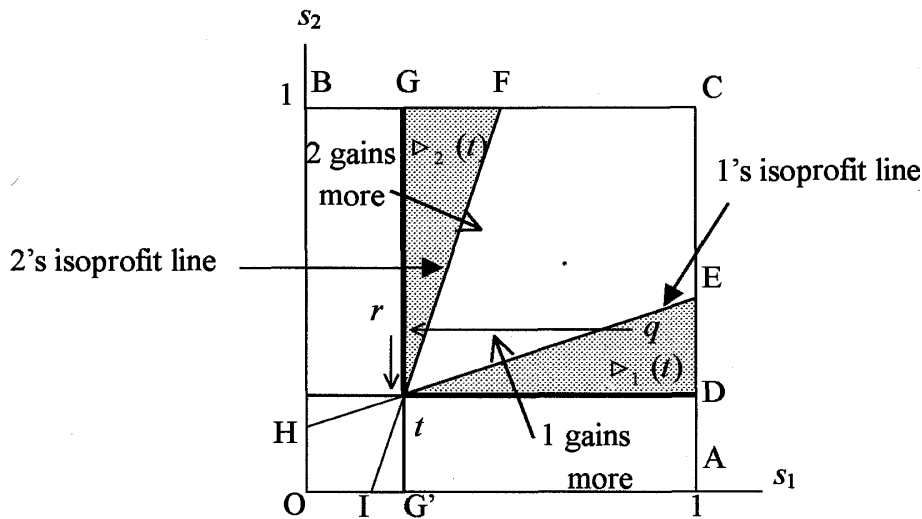Hence player 1's (resp. 2's) payoff increases as we move towards northwest (resp. southeast)

Figure 1 $\rhd_1 (t)$ and $\rhd_2 (t)$

corner. On a line parallel to the horizontal axis (resp. vertical axis), player 1's (resp. 2's) payoff increases as we move to left (resp. downward).

We will prove $\rhd_1 (t) = t DE \setminus t E$. A similar proof applies to $\rhd_2 (t) = t FG \setminus t F$. Take any point $q = (q_1, q_2)$ in $t DE \setminus t E$. Let $r = (t_1, q_2)$. If q is in (line $t D$) $\setminus$ (point t) (simply denoted $t D \setminus \{t\}$), then t directly dominates q via player 1. If otherwise, $u_1(t) > u_1(q)$ and $u_2(t) > u_2(r)$; and thus $q \in \rhd_1 (t)$. We next show that no point outside $t DE \setminus t E$ is in $\rhd_1 (t)$. No point above or on the line HE is in $\rhd_1 (t)$ because the payoff to player 1 is never below his/her payoff in t. Thus take any point q locating below HtD and suppose $t \rhd_1 q$. Since q is below HtD, q is not directly dominated by t. Thus there exist a sequence of strategy combinations $q = s^0, s^1, \ldots, s^{m-1}, s^m = t$ $(m \geq 2)$, and a sequence of players $1 = i^1, i^2, \ldots, i^m$ such that $s^{j-1} \leftrightarrow_{i^j} s^j$ and $u_{i^j}(t) > u_{i^j}(s^{j-1})$ for all $j = 1, \ldots, m$. Suppose first $i^m = 1$; then s^{m-1} must be on $t D \setminus \{t\}$ since $u_1(t) > u_1(s^{m-1})$. Hence s^{m-2} must be in $t FG \setminus (t F \cup t G)$ since $u_2(t) > u_2(s^{m-2})$. Similarly s^{m-3} must be in $t DE \setminus (t E \cup t D)$. Repeat this procedure; then we obtain that all of $q = s^0, s^1, \ldots, s^m = t$ are in $t FG \setminus (t F \cup t G)$ or $t DE \setminus (t E \cup t D)$. This contradicts that q is taken from the region below HtD. The same contradiction follows when $i^m = 2$. Thus we have shown that q is not in $\rhd_1 (t)$ and thus $\rhd_1 (t) = t DE \setminus t E$. Q.E.D.

Theorem 4.2: (1) Let K^1 be a continuous curve connecting the origin O and a point on the edge B'C or A'C on which each player's payoff varies monotonically where OA' (resp. OB') is player 1's (resp. 2's) isoprofit line passing through O. See Figure 2. Then K^1 is a stable set w.r.t. $\rhd$. (2) The origin O is contained in all stable sets w.r.t. $\rhd$, but no point in regions OAA' \ OA' or OBB' \ OB' is contained in any stable set w.r.t. $\rhd$.

Proof: (1) Internal and external stability follow from Lemma 4.1. (2) We note that by Lemma 4.1 the origin O is not dominated by any other point, and thus by the external stability of the stable set it must be contained in any stable set. Since all points in OAA' \ OA' and OBB' \ OB' are indirectly dominated by O again by Lemma 4.1, any stable set must be in the region OA'CB'. Q.E.D.

It is to be noted any strategy combination in the region OA'CB' is individually rational. Hence Theorem 4.2 shows that if only individual moves are assumed and players are
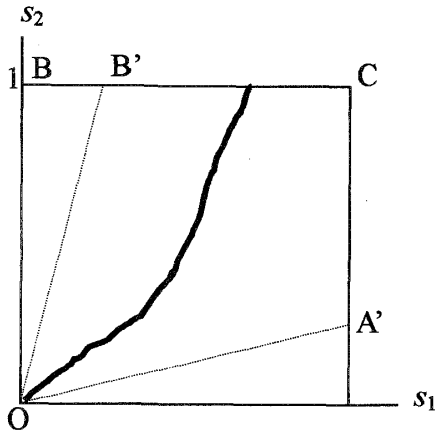
Figure 2   A stable set w.r.t. $\triangleright$

farsighted, then any individually rational strategy combination is contained in at least one stable set, and thus could be stable. Recall the remark on stability of strategy combinations given in Section 2 just below the definitions of the stable set and the consistent set. Further it has been shown that (Defect, Defect) is always stable, and that any strategy combination that is not individually rational is never stable.

The next theorem shows that the largest consistent set consists of all individually rational strategy combinations.

Theorem 4.3: Let L^I be the set of individually rational strategy combinations, i.e.,
$$L^I = \{s = (s_1, s_2) \in S : u_1(s), u_2(s) \geq 1\}. \text{ See Figure 3.}$$
Then L^I is the largest consistent set w.r.t. $\triangleright$.

Proof: Internal consistency: Take any point $t = (t_1, t_2)$ in L^I and any $q = (q_1, q_2)$ in S and any $i \in N$ such that $t \leftrightarrow_i q$. Without loss of generality let $i = 1$ and $q_2 = t_2$. If $q_1 > t_1$, then $u_1(t) > u_1(q)$; thus the internal consistency holds if q is in L^I. If q is not in L^I, then take t again; then since t directly dominates q, the internal consistency holds. Suppose $q_1 < t_1$. Then we can take a point $r = (r_1, r_2) \in L^I$ such that $r_1 = q_1$, $r_2 < q_2$ (thus $q \leftrightarrow_2 r$ and $u_2(r) > u_2(q)$) and $u_1(t) \geq u_1(r)$. For example, we may take r on the line OA'. Thus the internal consistency holds. See Figure 3.
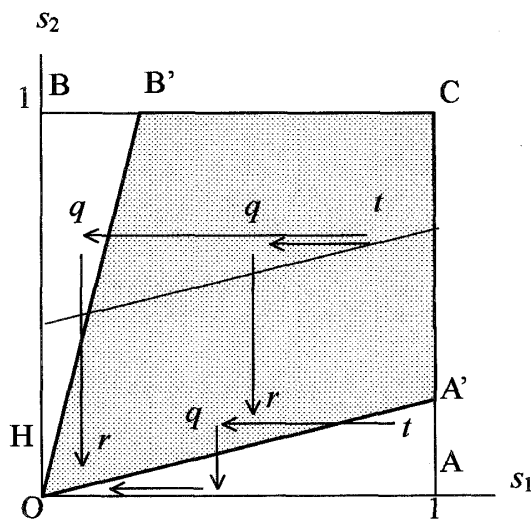

Figure 3   The largest consistent set without coalitional moves

External consistency: Take any $t = (t_1, t_2)$ outside L^I. Without loss of generality, suppose $u_1(t) < 1$. Then we can take a point $q = (q_1, q_2) \in L^I$ with $q_1 < t_1$ and $q_2 = t_2$; thus q directly dominates t via player 1. Note that in any point in L^I player 1 gains more than in t. Hence even if player 2 deviates from q and a sequence of deviation follows, player 1 is better off in the end point of the sequence if it is in L^I. See Figure 3. Thus the external consistency holds.

Finally we show that L^I is the largest consistent set. Suppose there exist a consistent set $L^{\prime I}$ that is larger than L^I. Take a point $t = (t_1, t_2)$ in $L^{\prime I} \setminus L^I$. Without loss of generality we suppose t is in the region OAA'$\setminus$ OA'; thus $u_1(t) < 1$. Pick $q = (0, t_2)$. Note that $t \leftrightarrow_i q$. Since $L^{\prime I}$ is a consistent set, by the internal consistency there exists r in $L^{\prime I}$ such that [$r = q$ or $r \triangleright q$] and $u_1(r) \leq u_1(t)$. Since $u_1(r) > u_1(t)$, neither $r = q$ nor $r \triangleright_1 q$ holds. Thus $r \triangleright_2 q$. Let $q = q^0, q^1, ..., q^p = r$ be the sequence of strategy combinations and $i^1, i^2, ..., i^p$ be the sequence of players which appear in $r \triangleright_2 q$. Thus $i^1 = 2$ and $i^2 = 1$. Since $i^1 = 2$, q^1

must be on OB and $u_1(q^1) \geq 1$. Thus $u_1(q^1) > u_1(t)$. Since $r \rhd q$ and $i^2 = 1$, we obtain $u_1(q^1) > u_1(t)$; and thus $r \rhd t$, which contradicts the fact that both r and t in L^π. Q.E.D.

In summary, when only individual moves are assumed, though the stable set w.r.t. indirect domination and the largest consistent set are different in their shapes, these sets provide similar implications concerning stability of strategy combinations: all individually rational strategy combinations could be stable. More precisely, every individually rational strategy combination is included in at least one stable set. Thus any of them could be stable if a stable set containing it would be realized. On the other hand, the largest consistent set consists of all individually rational strategy combinations; and thus all of them are stable.

This similarity, however, disappears when coalitional, but not binding, moves are taken into consideration.

4.2. Stability with coalitional moves

For $t \in S$, let $\succ_M (t) = \{s \in S : t \succ_M s\}$, $M = \{1\},\{2\},\{1,2\}$. The symbol $\succ_M (t)$ denotes the set of strategy combinations which t indirectly dominates when M deviates first. Let $\succ (t) = \succ_{\{1\}} (t) \cup \succ_{\{2\}} (t) \cup \succ_{\{1,2\}} (t)$. Then we have the following lemma concerning $\succ (t)$.

Lemma 4.4: (1) Take any strictly individually rational strategy combination $t = (t_1, t_2)$ (i.e., in t, both players' payoffs are strictly greater than the maximin payoff, 1). Let A, B, C, D, E, F, G, H, I be as in Figure 1. See Figure 4. Then $\succ (t) = t\text{EAOBF} \setminus (t\text{E} \cup t\text{F})$ (the shaded area in Figure 4).

(2) Take a strategy combination $t = (t_1, t_2)$ with $u_1(t) \leq 1$, i.e., player 1's payoff is not strictly individually rational. Let A, B, C, E and I be as in Figure 1, and let $F = (1, 4 + t_2 - 4t_1)$, $H = (t_1 - 4t_2, 0)$ and $J = (t_1 - 4t_2, 1)$. The point F (resp. I) is the intersection of player 2's isoprofit line passing through t and the vertical line AC(resp. horizontal line OA); and the point E (resp. H) is the intersection of player 1's isoprofit line passing through t and the line AC (resp. line OA). The line HJ is parallel to the vertical axis. See Figure 5. Then $\succ (t) = t\text{FCJHAE} \setminus (t\text{E} \cup t\text{F} \cup \text{HJ})$ (the shaded area in Figure 5).

(3) Take a strategy combination $t = (t_1, t_2)$ with $u_2(t) \leq 1$, i.e., player 2's payoff is not strictly individually rational. Let A, B, C, F and H be as in Figure 1, and let $E = (4 + t_1 - 4t_2, 1)$, $I = (0, t_2 - 4t_1)$ and $J = (1, t_2 - 4t_1)$. The point F (resp. I) is the intersection of player 2's isoprofit line passing through t and the horizontal line BC(resp. vertical line OB); and the point E (resp. H) is the intersection of player 1's isoprofit line passing through t and the line BC (resp. line OB). The line IJ is parallell to the horizontal axis. See Figure 6. Then $\succ (t) = t\text{FBIJCE} \setminus (t\text{E} \cup t\text{F} \cup \text{IJ})$ (the shaded area in Figure 6).
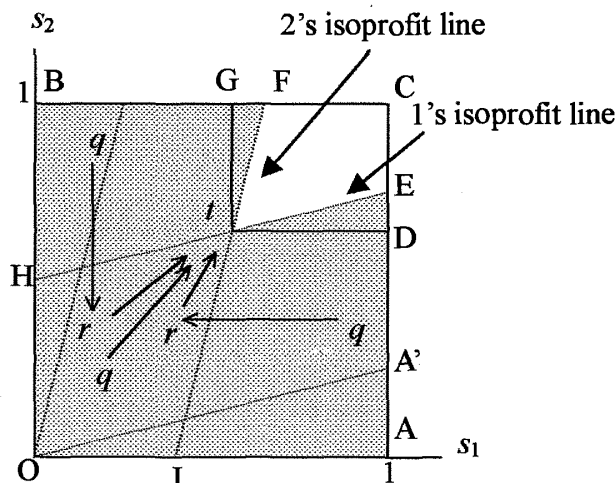

Figure 4 $\succ (t)$ of Lemma 4.
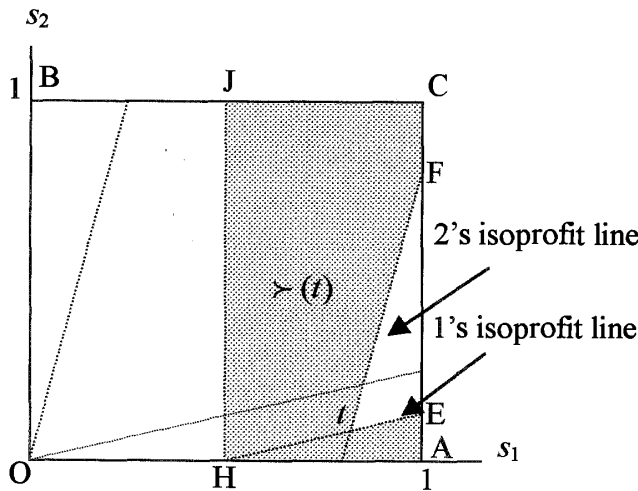
Figure 5 $\succ (t)$ of Lemma 4.4 (2)
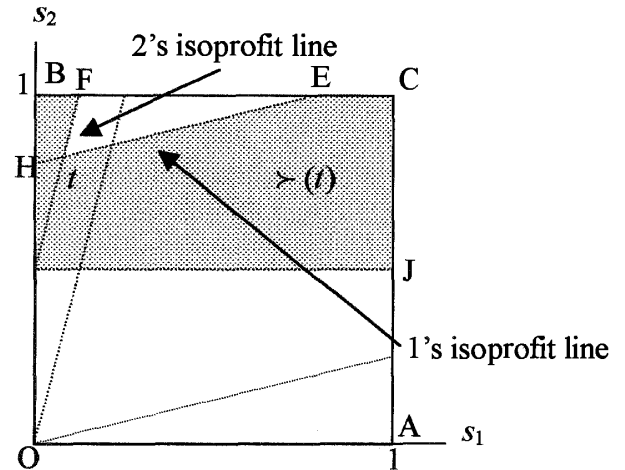
Figure 6 $\succ (t)$ of Lemma 4.4

Proof: (1) $\succ (t) \supseteq t\text{EAOBF} \setminus (t\text{E} \cup t\text{F})$: Take a point $q = (q_1, q_2)$ in $t\text{EAOBF} \setminus (t\text{E} \cup t\text{F})$. If q is in $t\text{ED} \setminus t\text{E}$ or $t\text{FG} \setminus t\text{F}$, then Lemma 4.1 shows that q is in $\succ_{\{1\}} (t)$ or $\succ_{\{2\}} (t)$; and thus q is in $\succ (t)$. Suppose next q is in $t\text{IOH} \setminus (t\text{H} \cup t\text{I})$. Then $u_1(t) > u_1(q)$ and $u_2(t) > u_2(q)$; thus q is directly dominated by t via coalition $\{1,2\}$ and thus in $\succ (t)$. Next take q in $t\text{DAI} \setminus t\text{D}$. Take a point $r = (r_1, r_2)$ in $t\text{IOH} \setminus (t\text{I} \cup t\text{H})$ with $r_2 = q_2$. See Figure 4. Then $t \succ q$ since $q \leftrightarrow_1 r$, $r \leftrightarrow_{12} t$, $u_1(t) > u_1(q)$, $u_1(t) > u_1(r)$ and $u_2(t) > u_2(r)$; thus q is in $\succ (t)$. A similar proof applies to q in $t\text{GBH} \setminus t\text{G}$. See Figure 4.

Conversely take a point q in $t\text{ECF}$. Then $u_1(q) \geq u_1(r)$ and $u_2(q) \geq u_2(r)$; and thus q is not in $\succ (t)$. Thus $\succ (t) = t\text{EAOBF} \setminus (t\text{E} \cup t\text{F})$ holds.

(2) The first half of the proof is the same as above. As for the second half, the fact that any point in $t\text{EF}$ is not in $\succ (t)$ is shown in the same manner as above. Take any point q in JBOH. Then $u_1(q) \geq u_1(t)$; so if q is in $\succ (t)$, the first deviator in the indirect domination must be player 2. But even if player 2 deviates from q, the new strategy pair is still in JBOH; and thus again player 1's payoff is greater than or equal to his/her payoff in t. Hence such an indirect domination is impossible. Thus q is not in $\succ (t)$.

(3) Similar to (2). Q.E.D.

Theorem 4.5: (1) Let $K^1 = \{t\}$, $t = (t_1, t_2)$, $1/4 < t_1 \leq 1$, $t_2 = 1$, or $t_1 = 1$, $1/4 < t_2 \leq 1$. That is, t is a point on the edge A'C or B'C in Figure 7: points A' and B' are excluded. Let $K^2 = \{(0,0), (1,1/4)\}$ and $K^3 = \{(0,0), (1/4,1)\}$. That is, K^2 consists of O and B' and K^3 consists of O and A'. Then K^1, K^2 and K^3 are stable sets w.r.t. $\succ$.

(2) Any point not contained in the three stable sets above is never be an element of any stable set w.r.t. $\succ$; and thus no stable set w.r.t. $\succ$ exists other than K^1, K^2 and K^3.

Proof: (1) K^1: Easily follows from Lemma 4.4(1). K^2: The following hold by Lemma 4.4(2): All points not on the line OB are in $\succ (A')$; all points on OB other than O are directly dominated by O via player 2; and A' is not in $\succ (O)$. Thus K^2 is a stable set w.r.t. $\succ$. K^3: A proof similar to K^2 applies. Use Lemma 4.4(3) instead of 4.4(2).

(2) Take a point $x = (x_1, x_2)$ not included in any of K^1, K^2, K^3 and suppose x is in some stable set K. First suppose x is in the interior of the quadrangle OA'CB'. See Figure 8. Then by Lemma 4.4(1) and by the external stability of the stable set, there must exist at least one point of K belonging to the region xC'CC'' \ $\{x\}$. By Lemma 4.4(1) every point in this region dominates x which contradicts the internal stability of the stable set.

Next suppose $x = (x_1, x_2)$ is in OAA' \ ($\{O\} \cup \{A'\}$). See Figure 8. A similar proof applies to the case where x is in OBB' \ ($\{O\} \cup \{B'\}$). By Lemma 4.4(2), any point in regions
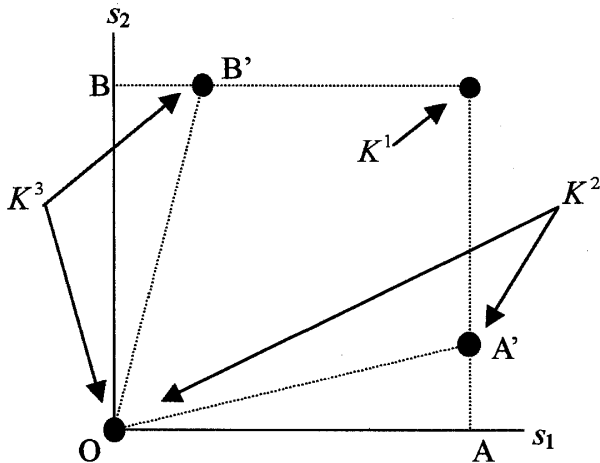
Figure 7 Stable sets K^1, K^2, K^3 w.r.t. $\succ$

xA''A''' and OGG'B is not indirectly dominated by x. Here the point A'' (resp. A''') is the intersection of player 2's (resp. player 1's) isoprofit line passing through x and the line AC, and the point G is the intersection of player 1's isoprofit line passing through x and the line OA. The line GG' is parallel to the vertical line OB. By the external stability of the stable set there must exist at least one point $y = (y_1, y_2)$ of the stable set K in the region xA''A''' $\cup$ OGG'B. By the same reason as above, y is not in xA''A'''. Further y is not in OGG'B \ (O'B''B $\cup$ $\{G\}$) where $O' = (0, x_2)$ and B'' is the intersection of

player 2's isoprofit line passing through O' and the line BC. In fact suppose y is in this region. Then if $y_2 > x_2$, then we can take a point $z = (z_1, z_2)$ such that $z_2 = x_2$, $u_1(z) < u_1(y)$ and $u_2(z) < u_2(y)$ as shown in Figure 8. Thus y indirectly dominates x since $u_1(y) > u_1(x)$. If $y_2 = x_2$, y directly dominates x via player 1 since $y_1 < x_1$. If $y_2 < x_2$, then we can take a point $z' = (z_1', z_2')$ such that $z_2' = x_2$ and $z_1' = y_1$. See Figure 8. Then since $u_1(y) > u_1(x)$ and $u_2(y) > u_2(z')$, y indirectly dominates x. Thus y must be in O'B''B $\cup$ $\{G\}$. Then by Lemma 4.4(2),(3), any point in region OGG''O' \ $\{G\}$ is indirectly dominated neither by x nor by y where G'' is the intersection of O'x and GG'. Therefore the external stability of K fails. Q.E.D.
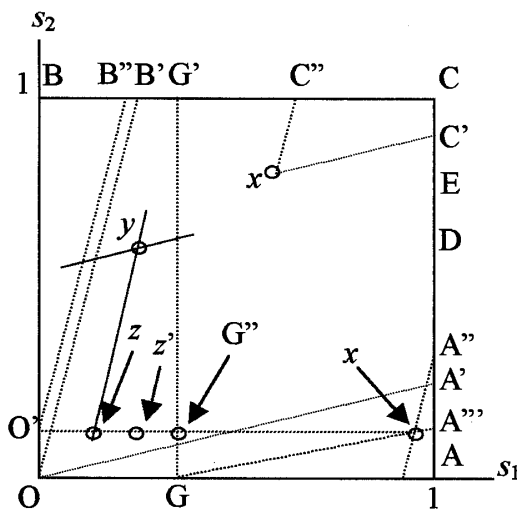

Figure 8 Illustration of x in Proof of Theorem 4.5

This theorem shows that when players' coalitional, but not binding, moves are assumed, essentially a single Pareto efficient point results as a stable outcome. We, however, have an exceptional case in which (Defect, Defect) could be stable together with one Pareto efficient point in which one player gains the same payoff as in (Defect, Defect).

The next theorem shows that the largest consistent set consists of all individually rational strategy combinations: the same as in the case without coalitional moves.

Theorem 4.6: Let L be the set of individually rational strategy combinations, i.e., $L = \{s = (s_1, s_2) \in S : u_1(s), u_2(s) \geq 1\}$. See Figure 9. Then L is the largest consistent set.

Proof: Internal consistency: Take any point $t = (t_1, t_2)$ in L and any $q = (q_1, q_2)$ in S and any $M \subseteq N = \{1,2\}$. When $M = \{1\}$ or $\{2\}$, the proof is the same as Theorem 4.3. Let $M = \{1,2\}$. Suppose q is not in L. Suppose q is in OBB' \ OB'; a similar proof applies when q is in OAA' \ OA'. See Figure 9. Then as shown in Figure 9, we can take $r = (r_1, r_2) \in L$ with $r_1 = q_1$, $r_2 < q_2$, and thus $q \leftrightarrow_2 r$, $u_2(r) > u_2(q)$ and $u_1(t) \geq u_1(r)$. For example, we can take r on the line OA'. If q is in L, then we may suppose $u_1(q) > u_1(t)$ and $u_2(q) > u_2(t)$; if otherwise, at least one player in M is worse off in q and thus the internal consistency directly follows. Then in a manner similar to the above, we can take a point $r = (r_1, r_2) \in L$ such that $q \leftrightarrow_2 r$, $u_2(r) > u_2(q)$ and $u_1(t) \geq u_1(r)$. Thus the internal consistency holds.

External consistency: Take any $t = (t_1, t_2)$ outside L. Without loss of generality, suppose $u_1(t) < 1$. Then as shown in Figure 9, we can take a point $q = (q_1, q_2) \in L$ with $q_1 < t_1$ and

Figure 9 The largest consistent set and its consistency

$q_2 = t_2$; thus q directly dominates t via player 1. Note that in any point in L player 1 gains more than in t. Hence even if player 2 deviates from q and a sequence of deviation follows, player 1 is better off in the end point of the sequence if it is in L. Thus the external consistency holds.

Finally we show that L is the largest consistent set. Suppose there exist a consistent set L' that is larger than L. Take a point t in $L' \setminus L$. Without loss of generality we suppose $t = (t_1, t_2)$ is in the region OAA' \ OA'; thus $u_1(t) < 1$. Pick $q = (0, t_2)$. See Figure 9. Note that $t \leftrightarrow_1 q$. Since L' is a consistent set, by the internal consistency there exists r in L' such that $[r = q$ or $r \succ q]$ and $u_1(r) \leq u_1(t)$. Since $u_1(q) > u_1(t)$, neither $r = q$ nor $r \succ_M q$, $M = \{1\}, \{1,2\}$, holds. Thus $r \succ_{\{2\}} q$. Let $q = q^0, q^1, \ldots, q^p = r$ be a sequence of strategy combinations and $M^1, M^2, \ldots, M^p$ be a sequence of players 1,2 or the player set $\{1,2\}$ which appear in $r \succ_2 q$. Thus $M^1 = \{2\}$ and $M^2 = \{1\}$ or $\{1,2\}$. Since $M^1 = \{2\}$, q^1 must be on OB and $u_1(q^1) \geq 1$. Thus $u_1(q^1) > u_1(t)$, which contradicts $1 \in M^2$. Q.E.D.

5. Concluding Remarks

We have shown in the prisoner's dilemma that the stable set with indirect domination and the largest consistent set produce the same result in pure strategies, but that in the mixed extension and thus in the prisoner's dilemma with many alternatives, outcomes are completely different. Though Chwe[1994] claimed an advantage of the largest consistent set over the stable set showing examples in which the former exists though the latter does not exist. As shown in this paper, however, in games like the prisoner's dilemma the stable set gives us much sharper prediction on players' behavior than the largest consistent set. Furthermore the

stable set analysis reveals that players joint move (though not binding) is indispensable for excluding inefficient outcomes in the prisoner's dilemma.

As we noticed in Introduction, similarity exists between the outcomes obtained in this paper and the results obtained through equilibrium analysis in noncooperative game setup. It must be interesting to study more in detail relations between the stable set with indirect domination (also the largest consistent set) and equilibrium in two-person strategic form games and their dynamic extensions. To proceed this analysis, an alternating move dynamic game, for example, those proposed by Maskin and Tirole [1988] and Bhaskar [1989] could be a good setting since indirect domination assumes consecutive moves of players or groups of players. The analysis will be done in a future paper.

Another interesting extension of the research is the study of a multi-person prisoner's dilemma, or sometimes called a social dilemma. This is also studied in a future paper.

In concluding the paper, we remark that the stable set with only direct domination does not provide us with interesting prediction of players' behavior. When only individual moves are assumed, a typical stable set is given in Figure 10 which is similar to the one given in Theorem 4.2: see also Figure 2. This stable set could, however, locate in regions OAA' and OBB'. Hence outcomes that are not individually rational could be stable. When coalitional moves are also taken into consideration, several types of stable sets exist. Figure 11 depicts one of them. This stable set cannot go into region OA'CB' of individually rational outcomes because of domination via coalition {1,2}. In fact, any point in the interior of OA'CB' directly dominates the origin (0,0) via {1,2}. Hence no outcome belonging to this stable set except the origin (0,0) is individually rational. There may exist a stable set which contains strategy combination (Cooperate, Cooperate); and moreover there exists another stable set containing even (Cooperate, Defect), (Defect, Cooperate).
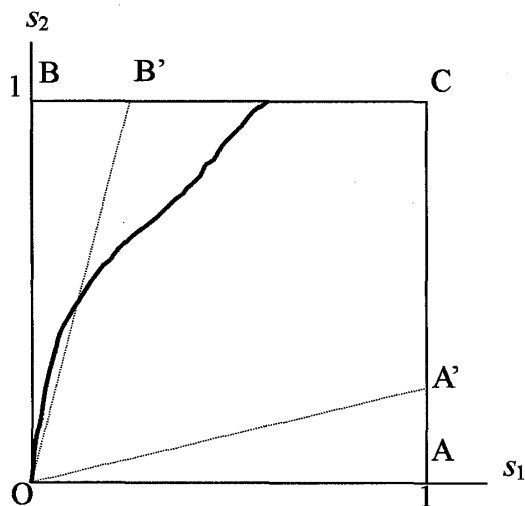

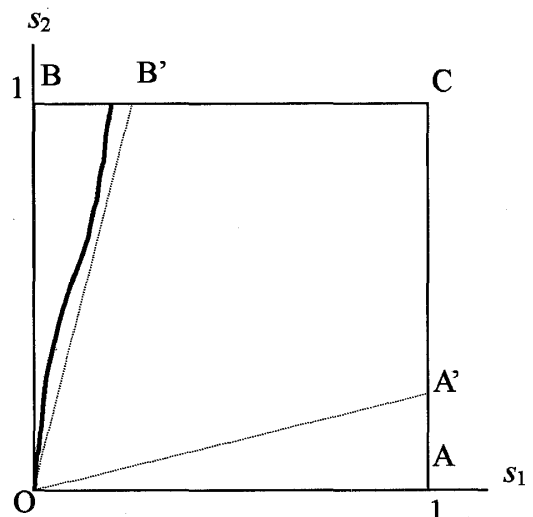**Figure 10** A stable set w.r.t. direct domination whereonly individual moves are assumed

Figure 11 A stable set w.r.t. direct domination where coalitional moves are also considered

References

[1] V. Bhaskar: Quick responses in duopoly ensure monopoly pricing. *Econ. Letters*, **29** (1989) 103-107.

[2] M. S.-Y. Chwe: Farsighted coalitional stability. *Journal of Economic Theory*, **63** (1994) 299-325.

[3] J. Greenberg: *The Theory of Social Situations: An Alternative Game-Theoretic Approach* (Cambridge Univ. Press, Cambridge, 1990).

[4] J.C. Harsanyi: An equilibrium-point interpretation of stable sets and a proposed alternative definition. *Management Science*, **20** (1974) 1472-1495.

[5] E. Maskin and J. Tirole: A theory of dynamic oligopoly, I: overview and quantity competition with large fixed costs. *Econometrica*, **56** (1988) 549-569.

[6] E. van Damme: Renegotiation-proof equilibria in repeated prisoners' dilemma. *Journal of Economic Theory*, **47** (1989) 206-217.

[7] E. van Damme: *Stability and Perfection of Nash Equilibria. 2nd ed.* (Springer-Verlag, Berlin, 1991).

[8] J. von Neumann and O. Morgenstern: *Theory of Games and Economic Behavior, 3rd ed.* (Princeton Univ. Press, Princeton, 1953).

Appendix

In this appendix, we will present stable sets and the largest consistent set in the general prisoner's dilemma. Recall the payoff matrix of the general prisoner's dilemma given in Section 3:

1	2 Cooperate	Defect
Cooperate	a, a	0, b
Defect	b, 0	c, c

Since $b > a > c > 0$, player 1 can increase his/her payoff by increasing the probability of using "defect", or decreasing the probability of using "cooperation", for each mixed strategy of player 2. Similarly for player 2.

Though in the particular example studied in the text, straight lines give players' isoprofit curves, it is not generally true. Consider a space of mixed strategy pairs of two players as given in the text. Recall the horizontal axis (resp, the vertical axis) is player 1's (resp, 2's) probability of taking "cooperation." Let s_1 (resp. s_2) be the probability that player 1 (resp. player 2) takes "cooperate."

Take the player 1's isoprofit line that gives him/her a fixed payoff level u. Then

$$u_1(s_1, s_2) = c - cs_1 + (b-c)s_2 + (a-b+c)s_1 s_2 = u, \text{ i.e., } s_2 = \frac{u - c + cs_1}{b - c + (a - b + c)s_1}.$$

Since $\dfrac{\partial u_1}{\partial s_1} = -c + (a - b + c)s_2$, $\dfrac{\partial u_1}{\partial s_2} = b - c + (a - b + c)s_1$ and $\dfrac{\partial u_1}{\partial s_1 \partial s_2} = \dfrac{\partial u_1}{\partial s_2 \partial s_1} = a - b + c$, we obtain

$$\frac{ds_2}{ds_1} = -\frac{-c + (a - b + c)s_2}{b - c + (a - b + c)s_1}, \quad \frac{d^2 s_2}{ds_1^2} = -\frac{2(a - b + c)(-c + (a - b + c)s_2)}{(b - c + (a - b + c)s_1)^2}.$$

Since $b > a > c > 0$ and $0 \le s_1, s_2 \le 1$, we obtain that $\dfrac{ds_2}{ds_1} > 0$, i.e., s_2 is increasing in s_1, and that

$$a - b + c > 0 \text{ iff } \frac{d^2 s_2}{ds_1^2} < 0, \quad a - b + c < 0 \text{ iff } \frac{d^2 s_2}{ds_1^2} > 0 \quad (\text{thus } a - b + c = 0 \text{ iff } \frac{d^2 s_2}{ds_1^2} = 0)$$

where iff implies "if and only if."

The isoprofit line passing through the origin $(s_1, s_2) = (0,0)$ is given by

$$s_2 = \frac{c s_1}{(a - b + c)s_1 + b - c}$$

since $u_1(0,0) = c$. Thus when $s_1 = 1$, we have $s_2 = c/a$ and thus $0 < s_2 < 1$ since $a > c > 0$. Hence the isoprofit line passing through the origin intersects $s_1 = 1$ at s_2 with $0 < s_2 < 1$.

Player 2's isoprofit curves have similar properties since players 1 and 2 are symmetric in the sense that $u_1(s_1, s_2) = u_2(s_2, s_1)$ for all (s_1, s_2).

Next, we examine slopes of two players' isoprofit curves. As shown above, the slope of player 1's isoprofit curve is given by

$$\frac{ds_2}{ds_1} = \frac{-c + (a - b + c)s_2}{b - c + (a - b + c)s_1}.$$

In a similar manner, we obtain the slope of player 2's isoprofit curve,

$$\frac{ds_2}{ds_1} = -\frac{b - c + (a - b + c)s_2}{-c + (a - b + c)s_1}.$$

Hence if $a - b + c = 0$, then

$$-\frac{-c + (a - b + c)s_2}{b - c + (a - b + c)s_1} = \frac{c}{a} < \frac{a}{c} = -\frac{b - c + (a - b + c)s_2}{-c + (a - b + c)s_1}$$

since $a > c$. Thus (slope of player 1's isoprofit curve) $\leq$ (slope of player 2's isoprofit curve). If $a - b + c \neq 0$, then (slope of player 1's isoprofit curve) $\leq$ (slope of player 2's isoprofit curve) if and only if

$$s_2 \geq -s_1 + \frac{2c - b}{a - b + c}.$$

(1) $a - b + c = 0$: In this case, we have $u_1(s_1, s_2) = c - cs_1 + (b - c)s_2$. Therefore player 1's isoprofit curve passing through the origin is a line: the line OA' in Figure 12 is a typical one. Similarly player 2's isoprofit curve passing through the origin is like OB'. Thus the case is just the one we studied in the text. Note that (slope of player 1's isoprofit curve) $<$ (slope of player 2's isoprofit curve). Stable sets are similar to those given in Figures 2 and 7.
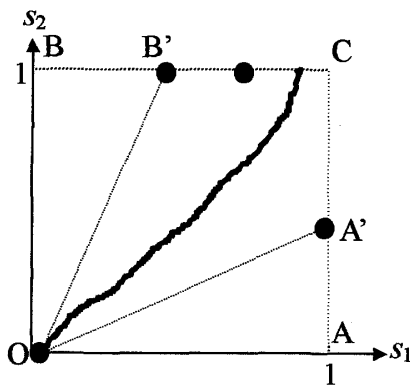

Figure 12   Stable sets in (1)

See Figure 12. When only individual moves are assumed, any monotone curve connecting the origin and a point on the edge B'C or A'C is a stable set. When coalitional moves are assumed in addition, three types of stable sets exist. The one is a singleton set consisting of a point on the edge B'C (excluding B') or the edge A'C (excluding A'). The others are two point sets $\{B', O\}$ and $\{A', O\}$. The largest consistent set is the quadrangle OA'CB'.

(2) $a - b + c > 0$: In this case the isoprofit curve is concave since $d^2 s_2 / ds_1^2 < 0$. Player 1's isoprofit curve passing through the origin intersects the 45° line only once at $s_1 = 0$ when $b \geq 2c$ and twice at $s_1 = 0$ and $s_1 = (2c - b)/(a - b + c)$ when $b < 2c$. See Figures 13 and 14. In the former case, $(2c - b)/(a - b + c) < 0$, and thus

$s_2 \geq -s_1 + (2c - b)/(a - b + c)$ holds. Hence (slope of player 1's isoprofit curve) $\leq$ (slope of player 2's isoprofit curve), and outcomes are similar to the case of $a - b + c = 0$. Stable sets are similar to those given in Figure 12. See Figure 13. The largest consistent set is OA'CB'.

In the latter case, $s_2 \geq -s_1 + (2c - b)/(a - b + c)$ holds and thus (slope of player 1's isoprofit curve)$\leq$ (slope of player 2's isoprofit curve) when $s_1 \geq (2c - b)/(a - b + c)$. Stable sets are similar to those given in Figure 14. Differences are: (1) Monotone curves start from the point O' instead of the origin O.; (2) Stable sets containing B' or A' with coalitional moves are not two point but three point sets; the point O' newly joined. The largest consistent set is O'A'CB' $\cup \{O\}$. Strategy pairs outside this region are all indirectly dominated by the point O', i.e.,

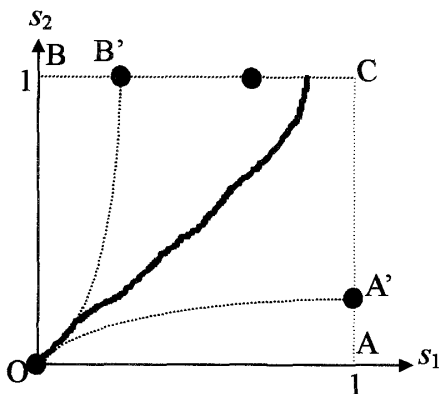$$(s_1, s_2) = ((2c - b)/(a - b + c), (2c - b)/(a - b + c)) \text{ or by the origin O.}$$

Figure 13   Stable sets in (2): the first case

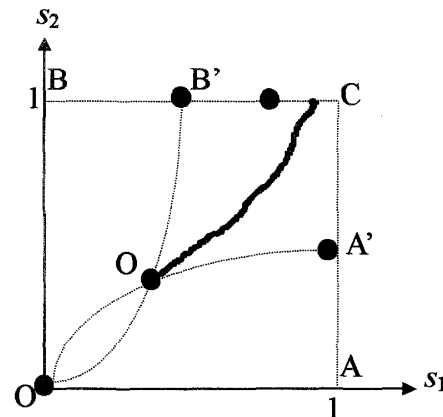**Figure 14**   Stable sets in (2): the second case

(3) $a - b + c < 0$: In this case the isoprofit curve is convex since $d^2 s_2 / ds_1^2 > 0$. We divide this case into two subclasses depending on whether the line $s_2 = -s_1 + (2c - b)/(a - b + c)$ intersects the square OACB. If the line does not intersect the

Figure 15   Stable sets in (3): the first case

square, thus $(2c - b)/(a - b + c) > 2$, or $b > 2a$, then $s_2 \geq -s_1 + (2c - b)/(a - b + c)$ holds in every point in the square. Hence (slope of player 1's isoprofit curve)$\leq$(slope of player 2's isoprofit curve) in the square, and outcomes are similar to those in the case of $a - b + c = 0$. Stable sets are similar to those given in Figure 12. See Figure 15. The largest consistent set is OA'CB'.
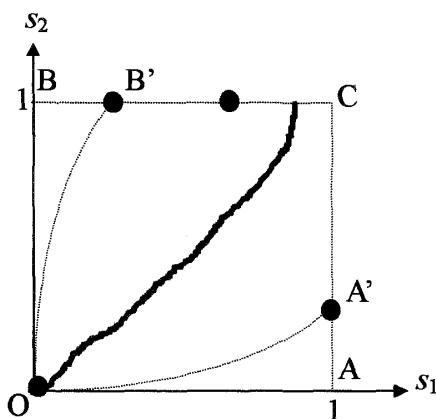
If $b \geq 2a$, then (slope of player 1's isoprofit curve)$\leq$(slope of player 2's isoprofit curve) below the line $s_2 = -s_1 + (2c - b)/(a - b + c)$, and (slope of player 1's isoprofit curve)$\geq$(slope of player 2's

isoprofit curve) above the line $s_2 = -s_1 + (2c - b)/(a - b + c)$. Thus any point on the line indirectly dominates all points above the line. Below the line $s_2 = -s_1 + (2c - b)/(a - b + c)$, outcomes are similar to the cases above. Stable sets are given in Figures 16 and 17. Figure 16 (resp. Figure 17) depicts the case where the line $s_2 = -s_1 + (2c - b)/(a - b + c)$ does not intersect (resp. intersects) the isoprofit curves through the origin. In both figures, stable sets are similar to those in previous figures. Differences are the following. (1) Monotone curves ends with the line $s_2 = -s_1 + (2c - b)/(a - b + c)$. (2) Singleton stable sets are on the line

$s_2 = -s_1 + (2c - b)/(a - b + c)$. (3) In Figure 17, two point stable sets consist of the origin O and A''' or B''', intersections of the line $s_2 = -s_1 + (2c - b)/(a - b + c)$ and isoprofit curves through the origin. The largest consistent set is OA'CB'.
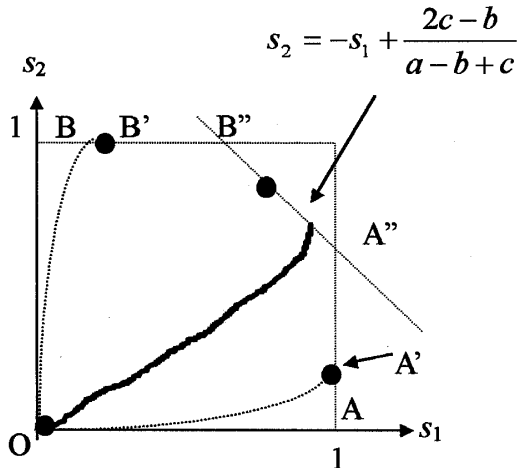
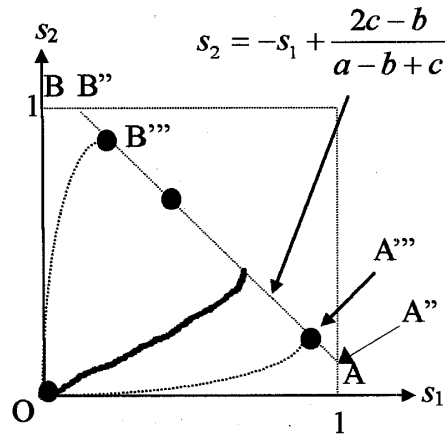

Figure 16 Stable sets in (3): the second case

Figure 17 Stable sets in (3): the third case

Shigeo Muto
Department of Value and Decision Science,
Graduate School of Decision Science and Technology,
Tokyo Institute of Technology
2-12-1 Oh-okayama, Meguro-ku, Tokyo 152-8552, Japan
E-mail: muto@valdes.titech.ac.jp