A Noncooperative Approach to General n-Person Cooperative Games

Akira Okada
Graduate School of Economics, Hitotsubashi University
2-1 Naka, Kunitachi, Tokyo 186-8601 JAPAN
Phone: +81 42 580 8599
Fax: +81 42 580 8748
E-mail: aokada@econ.hit-u.ac.jp

December, 2004

Abstract. We present a noncooperative foundation of the asymmetric Nash bargaining solution for a general n-person cooperative game in strategic form with coalition formation. In this setup, the Nash bargaining solution should be immune to any coalitional deviations. Unlike the classic approach of von Neumann and Morgenstern, our noncooperative approach shows that a threat by the complementary coalition should be consistent with the Nash bargaining solution. This result leads to a new concept of the core, called the *Nash core*, for a cooperative game in which any deviating coalition anticipates the Nash bargaining solution behavior of the complementary coalition. The main theorem shows that, in the limit that the probability of negotiation failure goes to zero, a (totally) efficient stationary subgame perfect equilibrium payoff of the bargaining model is equal to the Nash bargaining solution of the cooperative game, provided that it belongs to the Nash core. The weights of players in the asymmetric Nash bargaining solution are endogenously determined by a probability distribution to select a proposer.

1 Introduction

We consider a general cooperative situation among n individuals where they can communicate and form coalitions, which are enforceable, and cooperation within a coalition may have external effects on the utility of individuals outside the coalition. It covers a wide range of multilateral cooperation problems: exchange markets with externality, cartels among oligopolistic firms and international alliance among countries, and so on. These cooperative situations can be described by an n-person game in strategic form.

In game theory, there have been two different approaches to the general n-person cooperative situation described above. One is the *cooperative game approach* initiated by the classic work of von Neumann and Morgenstern (1944). von Neumann and Morgenstern reduce the n-person game in strategic form to its coalitional form (also called the characteristic function form), by assuming that individuals' utilities are transferable, that is, what is achievable for a coalition depends only on the sum of members' utilities, and moreover that the members of the coalition jointly choose their actions with expectation that the complementary coalition reacts by damaging them in the worst way. The first assumption of transferable utility is not critical to the von Neumann-Morgenstern theory. In fact, their theory has been extended to the case of non-transferable utility by Aumann (1961 and 1967), Aumann and Peleg (1960) and others. The second assumption on coalitional behavior, however, has been criticised on the grounds that it allows incredible threats by the complementary coalition (Scarf 1971, for example). The von Neumann-Morgenstern theory of a general cooperative game is regarded as a two-stage procedure. First, by using the minimax solution of a zero-sum two-person game between a coalition and its complementary coalition, one defines the value of the coalition. The function assigning to every coalition its value is called the characteristic function. Secondly, one investigates players' behavior only based on the characteristic function by applying a certain cooperative solution such as the stable set, core, Shapley value and others. The characteristic function has been used as a very convenient tool to analyse cooperative games. However, at the same time, a difficulty in the cooperative game approach is in the two-stage procedure itself of using the characteristic function.[1] It is not clear how coalitional behavior assumed in two stages can be justified in a consistent manner by rational behavior of individual players. The approach lacks a unified framework of analysis.

The other is the *noncooperative game approach* initiated by Nash (1951 and 1953). In his approach (called the Nash program), Nash proposed to study cooperative games based on reduction to noncooperative games by modelling pre-play ne-

[1] von Neumann and Morgenstern themselves point out this difficulty. They write, "Now it would seem that the weakness of our present theory lies in the necessity to proceed in two stages: To produce a solution of the zero-sum two-person game first and then, by using this solution, to define a characteristic function in order to be able to produce a solution of the general n-person game, based on the characteristic function (von Neumann and Morgenstern 1947, p. 608).

gotiations as moves in a noncooperative bargaining game. Analysing an equilibrium point of the noncooperative bargaining game, one can explain coalitional behavior as the result of individual players' payoff maximization. Nash (1953) presented a noncooperative foundation of his bargaining solution of a two-person cooperative game obtained by a set of axioms in his initial work (Nash 1950). An obvious restriction of Nash's work is that it covers only two-person general cooperative games. Recently, the noncooperative game approach to n-person cooperative games has received widespread research interests: Selten (1981), Chatterjee et al. (1993), Perry and Reny (1994), Moldovanu and Winter (1995), Okada (1996) among others. Most works, however, are based on an n-person coalitional form game and exclude the externality of coalitional behavior.[2]

In this paper, we attempt to extend a noncooperative foundation of the (asymmetric) Nash bargaining solution to an n-person general cooperative game in strategic form. Our bargaining game is based on the random-proposer model (Okada 1996) which is a generalization of the Rubinstein's (1982) alternating-offers model. In the model, a proposer is selected according to some probability distribution among active players. A proposal is a pair of a coalition and a jointly mixed action for members. The proposal is agreed by the unanimous consent among the members. When a coalition is formed, all remaining players continue their negotiations. The agreement by the coalition has an external effect on other players' negotiations. If a proposal is rejected, then negotiations may end with a small probability. When negotiations end, all players, except those who have already bound to some coalitional strategies, select their actions independently.

The purpose of our analysis is to characterize a stationary subgame perfect equilibrium (SSPE)[3] when the probability of negotiation failure is sufficiently small. In particular, since we are mainly interested in the noncooperative foundation of the Nash bargaining solution in a general n-person cooperative game, our analysis focuses an SSPE with the efficiency property that all active players cooperate both *on* and *off* equilibrium path. Such an equilibrium is called *totally efficient*.

The main results of the paper are summarized as follows. First, we will prove that if all players form the largest coalition in an SSPE, their agreement should be equal to the asymmetric Nash bargaining solution where the weights of players are determined by the probability distribution selecting proposers and the disagreement point of the bargaining solution is given by a Nash equilibrium of the game. Secondly, we will prove that, when one subcoalition is formed (off equilibrium path) in a totally efficient SSPE, the complementary coalition reacts by the Nash bar-

[2]Exceptions are Bloch (1996) and Montero (2000) whose bargaining models are based on an n-person game in partition function form where the value of a coalition depends on what coalitions are formed among other players.

[3]The stationarity here means that every player's equilibrium strategy depends only on payoff-relevant state variables in the model. More precisely, this is a subgame perfect equilibrium satisfying subgame consistency proposed by Harsanyi and Selten (1988). Our equilibrium concept is equivalent to that of Markov-perfect equilibrim defined by Maskin and Tirole(?).

gaining solution for its own negotiation problem. This means that, unlike the von Neumann-Morgenstern theory, players outside the coalition do not react to damage it by employing the minimax strategy. With this result, we will prove that in order for the Nash bargaining solution to be sustained as a totally efficient SSPE in our noncooperative bargaining model, it should be immune to any coalitional deviation anticipating that the complementary coalition will react according to the Nash bargaining solution theory. In other words, the Nash bargaining solution must be in a form of the *core* of the underlying cooperative game in the sense that no coalitional deviation can improve upon it with the expectation of the Nash bargaining solution behavior of the complementary coalition. We will call this new type of the core for a cooperative game in strategic form the *Nash core*. Finally, we will prove that, in the limit that the probability of negotiation failure goes to zero, a totally efficient SSPE (uniquely) exists if and only if the Nash bargaining solution is in the Nash core of the game.[4]

The notion of the Nash core can be supported by an argument of the consistency of a cooperative solution (the Nash bargaining solution) as follows. Suppose that a cooperative solution is accepted as the standard of behavior in a game. Since any coalition of players can be freely formed, the cooperative solution should be stable against any coalitional deviation. When some coalition deviates from the solution, the behavior of the complementary coalition should be governed by the same standard of behavior. This consistency argument naturally leads to the condition that the Nash bargaining solution should belong to the Nash core.

The paper is organised as follows. Section 2 provides definitions and notations. Section 3 presents a noncooperative bargaining model for an n-person cooperative game in strategic-form. Section 4 states the main theorems. Section 5 proves the results. Section 6 concludes the paper.

2 Definitions and Notations

We start with several notations. For a finite set N with n elements, let R^N denote the n-dimensional Euclidean space with coordinates indexed by the elements of N. Any point in R^N is denoted by $x = (x_i)_{i \in N}$, and also by $x = (x_1, x_2, \cdots, x_n)$ when N is indexed as $\{1, 2, \cdots, n\}$. For $i \in N$ and $x = (x_i)_{i \in N} \in R^n$, x_{-i} denotes the $(n-1)$-dimensional vector constructed from x by deleting the i-th coordinate x_i in x. The point x is sometimes written as (x_i, x_{-i}). For $S \subset N$, R^S denotes the subspace of R^N spanned by the axes corresponding to elements in S. For a finite set T, the notation $\Delta(T)$ denotes the set of all probability distributions on T.

An *n-person cooperative game in strategic form* is defined by a triplet $G = (N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N})$ where $N = \{1, 2, \cdots, n\}$ is the set of players and each A_i

[4] The if-part is proved under a technical condition.

$(i \in N)$ is a finite set of player i's actions.[5] The Cartesian product $A = \Pi_{i \in N} A_i$ is the set of all action profiles $a = (a_1, \cdots, a_n)$ for n players. Player i's payoff function u_i is a real-valued function on A. A probability distribution on A_i is called a mixed action for player i. A subset S of N is called a *coalition* of players. For a coalition S, let $A_S = \Pi_{i \in S} A_i$ be the set of action profiles $a_S = (a_i)_{i \in S}$ for all members in S. A *correlated action* p_S of the coalition S is a probability distribution on A_S. The set of all correlated actions for the coalition S is given by $\Delta(A_S)$. The idea of a correlated action is that all members in a coalition choose their actions jointly according to the corresponding probability distribution. In the cooperative game G, it is assumed that any coalition S can make an enforceable agreement to employ any correlated action if all members agree to it.

A *coalition structure* $\pi = [S_1, \cdots, S_m]$ on N is defined by a partition of N, a class of subsets of N satisfying that $N = S_1 \cup \cdots \cup S_m$ and every two S_i and S_j are disjoint. For a coalition structure $\pi = [S_1, \cdots, S_m]$ on N, an element $p^\pi = (p_{S_1}, \cdots, p_{S_m})$ in $\Pi_{j=1}^m \Delta(A_{S_j})$ is called a *correlated action profile* for the coalition structure π. When a correlated action profile p^π for $\pi = (S_1, \cdots, S_m)$ is employed, each player $i \in N$ obtains the expected payoff

$$u_i(p^\pi) = \sum_{a_{S_1} \in A_{S_1}} \cdots \sum_{a_{S_m} \in A_{S_m}} \Pi_{j=1}^m p_{S_j}(a_{S_j}) \cdot u_i(a_{S_1}, \cdots, a_{S_m}) \tag{2.1}$$

where $p_{S_j}(a_{S_j})$ $(j = 1, \cdots, m)$ is the probability that the correlated action p_{S_j} of coalition S_j assigns to an action profile $a_{S_j} \in A_{S_j}$. Given a coalition structure $\pi = (S_1, \ldots, S_m)$ on N, we define

$$F(G, \pi) = \{(u_1(p^\pi), \cdots, u_n(p^\pi)) \in R^N \mid p^\pi \in \Pi_{j=1}^m \Delta(A_{S_j})\}. \tag{2.2}$$

$F(G, \pi)$ represents the set of all expected payoff vectors for n players attained by correlated action profiles for π. When π consists only of the grand coalition N, that is, $\pi = [N]$, $F(G, [N])$ is simply denoted by $F(G)$. We call $F(G)$ the *feasible set* of the cooperative game G. The feasible set $F(G)$ represents the set of all expected payoff vectors of n players when they form the grand coalition N. The set $F(G)$ is a polyhedral compact convex subset of R^N and $F(G) \supset F(G, \pi)$ for every coalition structure π on N. We remark that the set $F(G, \pi)$ is not necessarily convex.

The upper-right boundary H of the feasible set $F(G)$ is defined as the set of points in $F(G)$ undominated (in a weak sense of Pareto) by any point in $F(G)$. With abuse of notation, we denote the equation of H as

$$H(x_1, \cdots, x_n) = 0$$

where H is a function on the feasible set $F(G)$. With no loss of generality, we assume that $H(x) \geq 0$ for all $x \in F(G)$. Also, for simplicity of the analysis, we assume:

Assumption 2.1

[5] In this paper, we will distinguish "action" and "strategy" since we consider a sequential bargaining game in extensive form based on the game G in strategic form.

(i) H is a concave and differentiable function and the first derivatives of H with respect to $x_1, \cdots, x_n$ satisfy

$$\frac{\partial H}{\partial x_1} \leq 0, \cdots, \frac{\partial H}{\partial x_n} \leq 0$$

(the equality may hold at most at the end points of the upper-right boundary H).

(ii) $F(G)$ has the full dimension n.

(iii) The strategic-form game G has a Nash equilibrium (in mixed strategies) whose payoff vector $d = (d_1, \cdots, d_n) \in F(G)$ has the property that the boundary of the set $F_d(G) \equiv \{x \in F(G) \mid x_i \geq d_i \text{ for all } i \in N\}$, other than n hyperplanes $x_i = d_i$ $(i \in N)$, is a subset of H.

The differentiability assumption (i) causes no loss of generality to our results. We can easily extend our results to the non-differentiable case since the piecewise linear function of the upper-right boundary H can be made to be the limit of differentiable functions. The same sign of all the first derivatives $\frac{\partial H}{\partial x_i}$ implies that the variables x_i and x_j are mutually strictly decreasing functions of each other on the upper-right boundary H. For each $i \in N$, let $F_{-i}(G)$ denote the projection of $F(G)$ over $R^{N-\{i\}}$. For every $x_{-i} \in F_{-i}(G)$, we define $h_i(x_{-i}) = \max\{x_i \mid (x_i, x_{-i}) \in F(G)\}$. By the assumption (i) and the convexity of $F(G)$, h_i is a differential concave function over $F_{-i}(G)$. $h_i(x_{-i})$ is the maximum payoff that player i can receive in the feasible set $F(G)$ while all other players' payoffs are fixed at x_{-i}. Assumptions (ii) and (iii) are technical. The assumption (iii) guarantees that for all $x \in F_d(G)$ the point $(h_i(x_{-i}), x_{-i})$ is located on the upper-right boundary H of $F(G)$.

In the rest of this section, we introduce several notions in cooperative game theory. Since the classic work of von Neumann and Morgenstern (1944), the *characteristic function* approach has been employed in cooperative game theory to consider the problem of coalition formation and payoff distributions. The characteristic function of a cooperative game assigns to each coalition the set of payoff vectors that the coalition can "assure" its members. Regarding a strategic-form game, the following two kinds of characteristic functions have been primarily studied in the literature (see Aumann 1961 and 1967). A coalition S is said to be *α-effective* for a payoff vector $x \in R^N$ if there is a $p_S \in \Delta(A_S)$ such that for any $p_{N-S} \in \Delta(A_{N-S})$, we have $u_i(p_S, p_{N-S}) \geq x_i$ for all $i \in S$. Let $v^\alpha(S)$ be the set of all payoff vectors for which S is α-effective. A coalition S is said to be *β-effective* for $x \in R^N$ if for any $p_{N-S} \in \Delta(A_{N-S})$ there exists a $p_S \in \Delta(A_S)$ such that $u_i(p_S, p_{N-S}) \geq x_i$ for all $i \in S$. Similarly to $v^\alpha(S)$, let $v^\beta(S)$ be the set of all payoff vectors for which S is β-effective. It is easily shown that $v^\alpha(S) \subset v^\beta(S)$ for every $S \subset N$. The functions

v^α and v^β are called the α-characteristic function and the β-characteristic function, respectively.

Intuitively, $v^\alpha(S)$ is the set of all payoff vectors $x \in R^S$ such that coaltion S can guarantee all members at least the payoff x, independently of what the members of the complementary coalition $N - S$ choose to do. On the other hand, $v^\beta(S)$ is the set of all payoff vectors $x \in R^S$ such that $N - S$ cannot prevent S from getting at least x. In general, these two sets are different, but for a two-person game and for an n-person game with transferable utility where side-payments are allowed, they coincide by the minimax theorem.

The noncooperative approach in this paper yields a payoff distribution closely related to two standard cooperative solution concepts, the core and the Nash bargaining solution.

Definition 2.1.

(1) Let $v = v^\alpha$ or v^β. A payoff vector $x \in R^N$ is said to *dominate* a payoff vector $y \in R^N$ with respect to v if there exists some coalition S of N such that $x \in v(S)$ and $x_i > y_i$ for all $i \in S$

(2) The *α-core* of a cooperative game G is the set of payoff vectors $x \in F(G)$ which are not dominated by any other payoff vector in $F(G)$ with respect to v^α. The *β-core* of G is the set of payoff vectors $x \in F(G)$ which are not dominated by any other payoff vector in $F(G)$ with respect to v^β.

Definition 2.2. Let $\theta^N = (\theta_i^N)_{i \in N} \in \Delta(N)$, and $d^N = (d_i^N)_{i \in N} \in F(G)$. A correlated action $b^* \in \Delta(A_N)$ of N is called the *(asymmetric) Nash bargaining solution* of G if b^* is an optimal solution of the maximization problem

$$\begin{array}{ll} \max & \sum_{i=1}^{n} \theta_i^N \cdot \log[u_i(p) - d_i^N] \\ \text{subject to} & (1)\ p \in \Delta(A_N) \\ & (2)\ u_i(p) \geq d_i^N \quad \text{for all} \quad i = 1, \cdots, n. \end{array}$$

Here, θ^N is called the *weight vector* of players, and d^N the *disagreement point.* The Nash bargaining solution b^* of G with the weight vector θ^N and the disagreement point d^N is denoted by $b^*(G, \theta^N, d^N)$ whenever the dependency on G, θ^N and d^N should be emphasized. The payoff $u(b^*) = (u_i(b^*))_{i \in N}$ of players generated by the Nash bargaining solution b^* is called the *Nash bargaining solution payoff.*

In negotiations, the grand coalition N is not always formed. If the members of a coalition $S \subset N$ agree to choose a correlated action $p_S \in \Delta(A_S)$, all remaining players may continue their negotiations, given the agreement of the correlated action p_S by S. The following game describes negotiations after some coalition is formed.

Definition 2.3. Let G be an n-person cooperative game in strategic form. For every coalition S and every correlated action $p_S \in \Delta(A_S)$ of S, a *subgame* $G(p_S)$ of G is defined to be the same game as G except that all players in S are bound to follow the correlated action p_S.[6]

The feasible set $F(G(p_S))$ of a subgame $G(p_S)$ can be defined in the same manner as the feasible set $F(G)$ of G. Note that the set of "active" players is $N - S$ in the subgame $G(p_S)$. The model of a subgame $G(p_S)$ of G can describe a general situation where more than one coalitions form. Suppose that several disjoint coalitions $S_1, \cdots, S_k$ have been formed and that members in each coalition S_j have agreed to employ some correlated action $p^j \in \Delta(A_{S_j})$. Negotiations among the remaining players in $N - S$ can be described by the subgame $G(p_S)$ where $S = S_1 \bigcup \cdots \bigcup S_k$ and p_S is the correlated action of S generated by $p^1, \cdots, p^k$.

Our cooperative solution for a strategic-form game G does not simply specify a feasible payoff (or a correlated action) for the grand coalition N. Rather, it is a *payoff configuration*, which specifies for *every* coalition S of N a feasible payoff of S.[7] Since the feasible payoff for the coalition S depends on a correlated action of the complementary coalition $N - S$ in our set-up of a strategic-form game G, a payoff configuration specifies for every coalition S and every correlated action p^{N-S} of the complementary coalition $N - S$ a feasible payoff of S. Formally, a payoff configuration of G is defined as a function ϕ which assigns for every coalition T and every correlated action p_T of T an element $\phi(p_T)$ in the feasible set $F(G(p_T))$ of the subgame $G(p_T)$ (Put $S = N - T$ in the discussion above). In the next section, we will see that a payoff configuration of G can be naturally derived by a strategy profile for a noncooperative bargaining model in G.

We extend the Nash bargaining solution of G to a solution configuration of G. Let θ be a function assigning to each $S \subset N$ a weight vector $\theta^S \in \Delta(S)$ of members in S. We call θ the *weight configuration* of N. Let d be a function assigning to every correlated action $p_S \in \Delta(A_S)$ of every coalition S a point $d(p_S)$ in the feasible set $F(G(p_S))$ of the subgame $G(p_S)$. The point $d(p_S)$ is interpreted as a diagreement point for negotiations among members in the complementary coalition $N - S$, given that the coalition S employs the correlated action p_S. We call d the *disagreement configuration* of G.

Definition 2.4. The *Nash bargaining solution configuration* b^* of G with a weight configuration θ of N and a disagreement configuration d is a function which assigns

[6]Here we should not confuse a subgame of G with the standard notion of a subgame in an extensive-form game, although it turns out that every subgame of G naturally corresponds to a subgame of a noncooperative bargaining model in extensive form introduced in Section 3.

[7]The formulation of a cooperative solution as a payoff configuration is employed in the axiomatization of the Harsanyi value for a cooperative game with non-transferable utility by Hart (1985).

to every correlated action p_S of every coalition S the Nash bargaining solution $b^*(p_S) = b^*(G(p_S), \theta^{N-S}, d^{N-S})$ of the subgame $G(p_S)$.[8] The payoff configuration of G generated by b^* is called the *Nash bargaining solution payoff configuration.*

The characteristic function, which prescribes what a coalition can achieve by itself, has played a central role in cooperative game theory since von Neumann and Morgenstern (1944). A characteristic function of a strategic-form game assumes a certain behavior of a coalition S and the complementary coalition $N - S$. In the α-characteristic function, a coalition S, in attempting to improve its position, must take into account all strategic possibility open to the complemetary coalition $N - S$. In the literature (Scarf 1971, for example), it has been critisized that a coalition S excessively considers threats by the members in $N - S$ which may be harmful to themselves. Alternatively, one can argue that a counter-action of the complementary coalition $N - S$ should be consistent with the members' utility maximizing behavior. From this point of view, by using the Nash bargaining solution configuration, we define a new notion of effectiveness for a cooperative game in strategic form, which is weaker than the α-effectiveness.

Definition 2.5. Let b^* be the Nash bargaining solution configuration of a cooperative game G with a weight configuration θ and a disagreement configuration d.

(i) A coalition $S \subset N$ is said to be *Nash-effective* for a payoff vector $x \in R^n$ if there exists some $p_S \in \Delta(A^S)$ such that

$$u_i(p_S, b^*(p_S)) \geq x_i \quad \text{for all} \quad i \in S \tag{2.3}$$

where $b^*(p_S) = b^*(G(p_S), \theta^{N-S}, d^{N-S})$ is the Nash bargaining solution of the subgame $G(p_S)$ of G assigned by b^* under θ and d.

(ii) The *Nash characteristic function* v^{Nash} of G is a function that assigns to each coalition $S \subset N$ the set, denoted by $v^{Nash}(S)$, of all payoff vectors in R^n for which S is Nash-effective.

(iii) The *Nash core* of G is the core of G with respect to the Nash characteristic function v^{Nash}.

The Nash-effectiveness is based on the following idea. When a coalition S chooses a correlated action p_S, it should consider a counter-action of the complementary coalition $N - S$ which is consistent with its members' payoff maximization. In other words, the coalition S should consider only credible threats by the complementary coalition $N - S$. A question remains: what is the outcome of the payoff-maximizing

[8] For notational simplicity, we use the same symbol b^* for the Nash bargaining solution configulation as the Nash bargaining solution of a subgame $G(pS)$.

behavior of $N - S$? Since the members in $N - S$ can negotiate about their correlated action, it is reasonable to assume that the members in the complementary coalition $N - S$ agree to choose the Nash bargaining solution of their own negotiation problem described by the subgame $G(p_S)$, given that the coalition S chooses the correlated action p_S. It is easily seen that the Nash-effectiveness is weaker than the α-effectiveness. The β-effectiveness assumes coalitional behavior different from those in the α-effectiveness and the Nash-effectiveness. It is as if the coalition S forces the complementary coalition $N - S$ to move first, and then responds (Scarf 1971).

Without specifying a noncooperative bargaining model, it is hard to decide which notion of effectiveness is appropriate for the analysis of cooperation in a strategic-form game. In the next section, we will present a noncooperative sequential bargaining model and will show that the Nash-effectiveness can be justified by a subgame perfect equilibrium of the bargaining model.

3 A Noncooperative Bargaining Model

The bargaining model of an n-person cooperative game G in strategic form is divided into two phases, (i) negotiations for coalition formation and (ii) choosing actions. The negotiation phase consists of a (possibly) infinite sequence of bargaining rounds. After the negotiation phase, the model has the phase of choosing actions in which all members in coalitions are bound to follow their agreed-upon correlated actions and the remaining players outside coalitions choose their individual (mixed) strategies independently. Let θ be a weight configuration of N. The precise rule of the bargaining model is given below.

(I) negotiation phase:

The negotiation phase has a (possibly) infinite bargaining rounds t $(= 1, 2, \cdots)$. Let N_t be the set of all "active" players who do not belong to any coalitions in round t. In the initial round, we put $N_1 = N$. The sequence of moves is as follows.

(1) In the beginning of each round t, every player $i \in N_t$ is randomly selected as a proposer according to the probability distribution $\theta(N_t) \in \Delta(N_t)$ which the weight configuration θ assigns to N_t.

(2) The selected player i proposes a coalition S with $i \in S \subset N_t$ and a correlated action $p_S \in \Delta(A_S)$ of S.

(3) All other members in S either accept or reject the proposal sequentially according to a predetermined order over N_t. The order of responders do not affect the result in any critical way.

(4) If all responders accept the proposal (S, p_S), then it is agreed upon and becomes binding. Then, negotiation goes to the next round $t+1$ with $N^{t+1} = N^t - S$. The same process as above is repeated in round $t+1$ with the probability distribution $\theta^{N_{t+1}} \in \Delta(N_{t+1})$.

(5) If any one responder rejects the proposal, then two events may happen. With probability $1 - \varepsilon$ ($\varepsilon > 0$), negotiations continue in the next round $t+1$ with $N^{t+1} = N^t$ under the same rule as in round t. With probability ε, negotitations break down and the game goes to the next phase of choosing actions.

(6) The negotiation process ends when every player in N joins some coalition, and the game goes to the phase of choosing actions.

(II) choosing action phase:

In this phase, all players in N choose their own actions, depending on the outcome of the negotiation phase. There are three possible cases.

(1) When agreements $(S_1, p_{S_1}), \cdots, (S_m, p_{S_m})$ with $S_1 \cup \cdots \cup S_m = N$ and $p_{S_i} \in \Delta(S_i)$ $(i = 1, \cdots, m)$ are reached, the agreed-upon correlated actions $p_{S_1}, \cdots, p_{S_m}$ are played.

(2) When negotiations do not stop, all players i who do not join any coalitions choose their individual (mixed) strategies $p_i \in \Delta(A_i)$ independently.[9] The other players, who join some coalitions, play their agreed-upon correlated actions.

(3) When negotiations break down in some round after a proposal is rejected, the same rule as (2) is applied.

The bargaining model above is denoted by $\Gamma^{\varepsilon,\theta}$. Formally, $\Gamma^{\varepsilon,\theta}$ is represented as an infinite-length extensive game with perfect information, that is, all players know all past actions of the game when they make their choices. We also use a notation Γ^{θ} to describe the bargaining model where the probability ε of negotiation failure converges to zero.

A *(behavior) strategy* for player i in $\Gamma^{\varepsilon,\theta}$ is defined according to the standard theory of extensive games. Let h_i^t be a history of the game $\Gamma^{\varepsilon,\theta}$ when player i has a turn to move in round t of the negotiation phase. The history h_i^t is represented by a sequence of all past actions in $\Gamma^{\varepsilon,\theta}$ before playre i's move in round t. Specifically, it describes who were proposers in all past rounds and how players responded to all past proposals.[10] Similarly, let $\bar{h}$ be a whole history of the negotiation phase when the action phase starts. Roughly, a strategy s_i of player i in $\Gamma^{\varepsilon,\theta}$ is a function which

[9] Note that the probability of this event is zero as long as the probability ε that negotiations break down after a proposal is rejected is positive.

[10] When player i is a responder in round t, the proposer and all responses before player i in round t are included in h_i^t.

assigns her action $s_i(h)$ to every possible history $h = h_i^t$ or $\bar{h}$. Specifically, player i's action $s_i(h)$, $h = h_i^t$ or $\bar{h}$, is given as follows.
(i) When player i is a proposer in round t, $s_i(h_i^t)$ is a probability distribution (with a finite support[11]) on the set of all possible proposals (S, p_S) with $i \in S \subset N_t$ and $p_S \in \Delta(A_S)$,
(ii) When player i is a responder in round t, $s_i(h_i^t)$ is a probability distribution over $\{$accept, reject$\}$,
(iii) When the action phase starts and player i does not belong to any coalition in the negotiation phase, $s_i(\bar{h})$ is player i's mixed action in $\Delta(A_i)$. When player i belongs to some coalition, she follows the agreed-upon correlated action in the action phase.

Let P denote the set of all correlated action profiles p^π for all coalition structures π of N. For a strategy profile $s = (s_1, \cdots, s_n)$ of players in $\Gamma^{\varepsilon,\theta}$, a probability distribution μ on P (with a finite support) is determined. Then, player i's expected payoff for a strategy profile s is given by

$$Eu_i(s) = \int_P u_i(p^\pi) d\mu \tag{3.4}$$

where $u_i(p^\pi)$ is the expected payoff of player i for a correlated action profile p^π defined by (2.1). In what follows, the expected payoff $Eu_i(s)$ is denoted by $u_i(s)$ with abuse of notations, and expected payoff is simply called payoff, whenever no confusion arises. We remark that the expected payoff vector $(u_1(s), \cdots, u_n(s))$ for every strategy profile s in $\Gamma^{\varepsilon,\theta}$ belongs to the feasible set $F(G)$ of G.

For every correlated action p_S of every coalition S, let $\Gamma^{\varepsilon,\theta}(p_S)$ be the subgame of the extensive game $\Gamma^{\varepsilon,\theta}$ which starts after the agreement (S, p_S) has been reached. For notational convenience, we set $\Gamma^{\varepsilon,\theta}(p_\emptyset) = \Gamma^{\varepsilon,\theta}$ when S is an empty set $\emptyset$. In the same way as (3.4), a strategy profile $s = (s_1, \cdots, s_n)$ of players in $\Gamma^{\varepsilon,\theta}$ generates the expected payoff vector for players in the subgame $\Gamma^{\varepsilon,\theta}(p_S)$, which is an element of the feasible set $F(G(p_S))$ of the game $G(p_S)$. In this way, a strategy profile $s = (s_1, \cdots, s_n)$ in $\Gamma^{\varepsilon,\theta}$ naturally generates a payoff configuration of the cooperative game G.

The solution concept that we apply to the bargaining model $\Gamma^{\varepsilon,\theta}$ is a stationary subgame perfect equilibrium.

Definition 3.1. A strategy combination $s^* = (s_1^*, \cdots, s_n^*)$ of the game $\Gamma^{\varepsilon,\theta}$ is called a *stationary subgame perfect equilibrium (SSPE)* if s^* is a subgame perfect equilibrium of $\Gamma^{\varepsilon,\theta}$ where every player i's strategy s_i^* is *stationary* to satisfy the property that the action $s_i^*(h)$ prescribed by s_i^* to any history h depends only on the collection of agreements, $(S_1, p_{S_1}), \cdots, (S_m, p_{S_m})$ which have been reached on h.[12] The payoff

[11] The assumption of the finite support does not affect the result at all since any probability mixture, with finite or infinite support, of correlated actions p_S of S can be reduced to a single correlated action of S.

[12] Precisely speaking, when player i is a responder, his response surely depends on a current

(configuration) generated by an SSPE is called an SSPE payoff (configuration).

Agreements by coalitions compose a payoff-relevant history of negotiations in the sense that they determine the payoff structure in the future negotiations among the players outside coalitions. The SSPE requires that every player's action should depend only on such a payoff-relevant history. It, however, should be emphasized that deviations from the equilibrium are allowed to be non-stationary. In the context of negotiations, it implies forgiveness - "let bygones be bygones." Players do not treat one another unfavorably even if they were treated so in past rounds of negotiations.

It is well-known that in a broad class of Rubinstein-type sequential multilateral bargaining games including our model $\Gamma^{\varepsilon,\theta}$, there is a large multiplicity of subgame perfect equilibria when the discount rate of future payoffs or the probability of breakdown in negotiations is very small (see Sutton 1986 and Osborne and Rubinstein 1990 for this result). The multiplicity of subgame perfect equilibria holds even in the n-person pure bargaining game where no subcoalitions are allowed. Mainly, by this reason, the concept of an SSPE is employed in almost every literature of noncooperative multilateral bargaining model (see Baron and Ferejohn 1989, Perry and Reny 1994, Chatterjee et al. 1993, Okada 1996 and 2000, Okada and Winter 2002, Winter 1996 among others). One possible justification for an SSPE is a focal-point argument. It is the simplest type of subgame perfect equilibrium and thus it may be easier for players to coordinate their mutual expectations on it (see Baron and Kalai 1993 and Chatterjee and Sabourian 2000 on this line of research). The SSPE is a natural reference point of the analysis in multilateral bargaining models.

In the literature of the equilibrium selection in noncooperative games, the SSPE is equivalent to the subgame perfect equilibrium satisfying *subgame consistency* introduced by Harsanyi and Selten (1989). The subgame consistency in general extensive games requires that every player should behave in the same way across "isomorphic" subgames. In the context of our bargaining game $\Gamma^{\varepsilon,\theta}$, all subgames starting from the beginning of all rounds can be considered isomorphic as long as the same collections of agreements have been reached before, since they have identical game trees in such a case. Also, an SSPE can be reformulated as a Markov-perfect equilibrium (Fudenberg and Tirole 1991) of $\Gamma^{\varepsilon,\theta}$ by taking the collection of agreements reached in past negotiations as a payoff-relevant state variable at each round.

The bargaining game $\Gamma^{\varepsilon,\theta}$ may suffer from two kinds of inefficiency. The first kind of inefficiency is that a proposal is rejected and negotiations break down with a positive probability. The breakdown of negotiations typically results in an inefficient outcome. The second kind of inefficiency is the failure of the grand coalition N. It is known that the first kind of inefficiency may occur in the noncooperative multilateral bargaining game where an initial proposer is determined according to a fixed order over the player set and the first rejector becomes the next proposer just like the Rubinstein's two-person alternating-offers model (Chatterjee et al. 1993).

proposal and may depend on who a proposer is and on how responders preceding to him have behaved in the same round.

When utility is transferable, Okada (1996) proves that this is not the case in the non-cooperative bargaining game where proposers are chosen randomly in every round. In the next section, it will be shown that this result can be extended to the case of non-transferable utility. Specifically, we will prove that in every SSPE of $\Gamma^{\varepsilon,\theta}$, every player's proposal is accepted in the first round. This enables us to focus the problem of inefficiency caused by the formation of subcoalitions.

Definition 3.2.

(i) An SSPE s of $\Gamma^{\varepsilon,\theta}$ is called *efficient* if the grand coalition N is formed in the initial round of the negotiation phase, independent of a proposer.

(ii) An SSPE s of $\Gamma^{\varepsilon,\theta}$ is called *totally efficient* if the coalition of all active players (if any) are formed in every round of the negotiation phase, independent of history.

(iii) A *limit efficient* SSPE of Γ^{θ} is defined to be a limit of efficient SSPEs of $\Gamma^{\varepsilon,\theta}$ as ε goes to zero. A *limit totally efficient* SSPE of Γ^{θ} is defined to be a limit of totally efficient SSPEs of $\Gamma^{\varepsilon,\theta}$ as ε goes to zero.

In an efficient SSPE, the grand coalition N is formed in the initial round of negotiations on equilibrium path. A totally efficient SSPE has a stronger property that the coalition of all active players is formed not only *on* equilibrium path but also *off* equilibrium path. In other words, the totally efficient SSPE of $\Gamma^{\varepsilon,\theta}$ induces an efficient SSPE on a subgame $\Gamma^{\varepsilon,\theta}(p_S)$ of $\Gamma^{\varepsilon,\theta}$ for every correlated action p_S of every coalition S, independent of whether it is reached by the equilibrium path or not. Obviously, a totally efficient SSPE of $\Gamma^{\varepsilon,\theta}$ is an efficient SSPE.

4 Theorems

The aim of our analysis is to characterize a limit totally efficient SSPE in the bargaining game Γ^{θ}. In this section, we will state the main theorems. All proofs are given in the next section. The following proposition is useful to our analysis.

Proposition 4.1. (No delay) Let s^* be an SSPE of $\Gamma^{\varepsilon,\theta}$. Then, for every $i \in N$, player i's proposal is accepted in the initial round of the negotiation phase in s^*.

The proposition shows that there is no delay of agreement in the bargaining game $\Gamma^{\varepsilon,\theta}$. That is, some agreement of coalition is reached immediately on equilibrium path. The bargaining rule of $\Gamma^{\varepsilon,\theta}$ that a proposer is selected randomly in every round is critical to this result. Under the other rule that the first rejector becomes the next proposer, the proposition does not hold (see Chatterjee et al. 1993 and Okada 1996). We, however, remark that the agreed-upon coalition is not necessarily efficient.

We are now ready to state the main theorems in the paper.

Theorem 4.1. Let $v = (v_1, \cdots, v_n)$ be a limit efficient SSPE payoff of Γ^θ. Let θ^N be the weight vector for N assigned by the weight configuration θ. Then, v is the Nash bargaining solution payoff of the cooperative game G with the weight vector θ^N and a disagreement point $d = (d_1, \cdots, d_n)$ which is a Nash equilibrium payoff of G.

The theoorem shows that when the probability ε of negotiation failure is sufficiently small, players agree to the Nash bargaining solution in an efficient SSPE of $\Gamma^{\varepsilon,\theta}$. Two remarks are in order. First, the disagreement point of the Nash bargaining solution is given by a Nash equilibrium in the strategic-form game G. Unlike Nash's (1953) optimal threat model, our bargaining model $\Gamma^{\varepsilon,\theta}$ (and Γ^θ) does not allow players to commit themselves to incredible threats which will be implemented when negotiations fail. The SSPE of $\Gamma^{\varepsilon,\theta}$ prescribes that players should play a Nash equilibrium of G when negotiations break down. Secondly, the theorem shows that the weights of players for the Nash bargaining solution is endogenously determined by the probability distribution to select a proposer in the bargaining game. The more likely a player is selected as a proposer, the greater bargaining power she obtains.

With help of Theorem 4.1, we are going to characterize a limit totally efficient SSPE of Γ^θ. By definition, a totally efficient SSPE of Γ^θ induces a totally efficient SSPE of every subgame $\Gamma^\theta(p_S)$ of it which starts after a coalition S agree to play a correlated action p_S. In other words, the members of the coalition S should anticipate the totally efficient SSPE behavior of the complementary coalition. This observation naturally leads to the notions of the Nash effectiveness and thus of the Nash core (Definition 2.5). A limit totally efficient SSPE payoff of Γ^θ is in the Nash core. If not, there exists some coalition S of which members can improve upon their SSPE payoffs by employing some correlated action. Every member of S has an incentive to propose such a coalitional deviation (when selected as a proposer) since all other members of S accept it. This contradicts the SSPE property. The Nash bargaining solution configuration which defines the Nash core has the disagreement configuration d satisfying the following property:

(A) For every correlated action $p_S \in \Delta(A_S)$ of every coalition S, the disagreement configuration d of G assigns a Nash equilibrium payoff of the subgame $G(p_S)$ of G.

Theorem 4.2. Let ϕ^* be the payoff configuration generated by a limit totally efficient SSPE s^* of Γ^θ. Then

(i) ϕ^* is the Nash bargaining solution payoff configuration which has the weight configuration θ and a disgareement point configuration d satisfyng (A), and
(ii) for every $S \subset N$ and every $p_S \in \Delta(A_S)$, the payoff $\phi^*(p_S) \in F(G(p_S))$ assigned by ϕ^* belongs to the Nash core of the subgame $G(p_S)$ defined by the Nash bargaining solution configuration with θ and d.

It follows from Theorem 4.2 that when the breakdown probability ε of negotiation is very small, a totally efficient SSPE payoff is the Nash bargaining solution payoff with the weight vector θ^N and moreover that it belongs to the Nash core of G. Since the totally efficient SSPE of $\Gamma^{\varepsilon,\theta}$ has the *subgame property* that it induces a totally efficient SSPE on *every* subgame $\Gamma^{\varepsilon,\theta}(p_S)$ of $\Gamma^{\varepsilon,\theta}$, the property above of the totally efficient SSPE payoff should be true on every subgame $\Gamma^{\varepsilon,\theta}(p_S)$.

To understand the condition (ii) of Theorem 4.2, we discuss what the condition means in the special case of a transferable utility game (N, v) in characteristic function form where the characteristic function v assigns a real value $v(S)$ to every coalition S of N. For a coalition S, a restriction of v on S is denoted by v_S. Our notion of a subgame $G(p_{N-S})$ of G corresponds to a transferable utility game (S, v_S) with player set S. The (symmetric) Nash bargaining solution of (S, v_S) with the disagreement point $v(\{i\}) = 0$ for all $i \in S$ is given by the equal payoff vector $(1/|S|, \cdots, 1/|S|)$ where $|S|$ denotes the number of members in S. Since the value $v(S)$ of coalition S is independent of the action by the complementary coalition $N - S$, the Nash core of the game (S, v_S) is equal to the usual core. Therefore, the Nash bargaining solution $(1/|S|, \cdots, 1/|S|)$ belongs to the core of (S, v_S) if and only if $v(S)/|S| \geq v(T)/|T|$ for all subcoalitions T of S. For the equal weights, the condition (ii) of Theorem 4.2 is reduced to a simple condition in the transferable utility game (N, v): $v(S)/|S| \geq v(T)/|T|$ for all two coalitions S and T of N with $T \subset S$. We proved in Okada (1996, Theorem 3) that the two conditions in Theorem 4.2 are equivalent to the existence of a limit totally efficient SSPE of Γ^θ for a transferable utility game in characteristic function form. The last theorem shows that the converse of Theorem 4.2 also holds true for a general cooperative game G in strategic form (under some technical condition).

Theorem 4.3. Let b^* be the Nash bargaining solution payoff configuration of a cooperative game G with a weight configuration θ and a disgareement configuration d satisfying (A). If b^* satisfies

(B) for every $S \subset N$ and every $p_S \in \Delta(A_S)$, the payoff $b^*(p_S) \in F(G(p_S)$ assigned by b^* belongs to the interior of the strict Nash core[13] of subgame $G(p_S)$ relative to the upper-right boundary of the feasible set $F(G(p_S))$,

then b^* is a payoff configuration generated by a limit totally efficient SSPE of Γ^θ.

[13]The strict core is defined by the same manner as the core except that the domination requires that any member of a coalition is never worse-off with at least one member being better-off. When utility is transferable, the core and the strict core conincide.

Before we prove the theorems in the next section, we discuss an example of a three-person game to illustrate our results.

Example 4.1 (a three-person prisoner's dilemma)

Consider a three-person game G in strategic form given in Figure 4.1. The game can be interpreted as the prisoner's dilemma. Every player i (=1, 2, 3) has two actions, C_i (cooperate) and D_i (defect). If all players cooperate, all receive payoff 2. If any one of them unilaterally deviate, she receives the highest payoff 6, while the two other players receive zero payoffs. If any two players jointly deviate, then they receive payof 3 and the other player receives payoff -2. If all players defect, they receive payoff 1. Every player i has the dominant action D_i, and thus the game has a unique Nash equilibrium (D_1, D_2, D_3). It can be seen that the action profile (C_1, C_2, C_3) is the (symmetric) Nash bargaining solution of the cooperative game G with the disagreement point (D_1, D_2, D_3).

C_3:

	C_2	D_2
C_1	2, 2, 2	0, 6, 0
D_1	6, 0, 0	3, 3, -2

D_3:

	C_2	D_2
C_1	0, 0, 6	-2, 3, 3
D_1	3, -2, 3	1, 1, 1

Figure 4.1 A three-person prisoner's dilemma game

We will show that the Nash bargaining solution (C_1, C_2, C_3) is in the Nash core of the cooperative game G. Before we construct the Nash characteristic function, we explain the basic idea behind the Nash core. Suppose that a single player i, say $i = 3$, deviates from the Nash bargaining solution to defect. Then, players 1 and 2 negotiate about how to react to player 3's deviation. Their strategic possibility is described by the two-person game $G_{\{1,2\}}$ in Figure 4.2. In the game $G_{\{1,2\}}$, (D_1, D_2) is the dominant equilibrium, and thus is a unique disgreement point. Since (D_1, D_2) is Pareto efficient in the game $G_{\{1,2\}}$, it is trivially the Nash bargaining solution of $G_{\{1,2\}}$. That is, players 1 and 2 agree to react to player 3's deviation by (D_1, D_2). Then, player 3's payoff decreases from 3 to 1. Player 3 is worse-off by deviation. Next, suppose that any two players, say 1 and 2, defect jointly. Then, player 3 reacts to this coalitional deviation by defecting herself since D_3 is her optimal action to (D_1, D_2). Then, the payoff of both players 1 and 2 decrease from 2 to 1. Players

1 and 2 are worse-off by the joint deviation. Since no coalition can improve upon the Nash bargaining solution (C_1, C_2, C_3), it belongs to the Nash core. In this case, our result shows that the Nash bargaining solution (C_1, C_2, C_3) can be supported by a totally efficient SSPE of the bargaining model $\Gamma^{\varepsilon,\theta}$ when the probability ε of negotiation failure is sufficiently small.

The Nash characteristic function of the cooperative game G is constructed as follows. Suppose that player 3 employs any mixed action $p_3 = (p, 1-p)$ where p $(0 \le p \le 1)$ is the probability to select C_3. Then, by the same argument as above, players 2 and 3 react to player 1 by employing the Nash bargaining solution (D_2, D_3) of their own bargaining problem. Therefore, the set of the Nash characteristic function $V^{Nash}(\{3\})$ for player 3 is given by

$$V^{Nash}(\{3\}) = \{w_3 \in R| \ \ w_3 \le 1 - 3p \ \text{ for some } p, \ 0 \le p \le 1\}.$$

Since $2 > 1 - 3p$ for any p $(0 \le p \le 1)$, player 3 can not improve upon the Nash bargaining solution (C_1, C_2, C_3). The same result holds for $i = 1, 2$.

	C_2	D_2
C_1	0, 0	-2, 3
D_1	3, -2	1, 1

Figure 4.2 A two-person game $G_{\{1,2\}}$ between players 1 and 2 when player 3 defects.

Next, suppose that players 1 and 2 jointly employ any correlated action $p^{12} = (p, q, r, 1-p-q-r)$ where p is the probability assigned to an action profile (C_1, C_2), q the probability assigned to an action profile (C_1, D_2), and r the probability assigned to an action profile (D_1, C_2). Since player 3 chooses the dominant action D_3, the set of the Nash characteristic function $V^{Nash}(\{1,2\})$ for players 1 and 2 is given by

$$V^{Nash}(\{1,2\}) = \{(w_1, w_2) \in R^2| \ \ w_1 \le 1 - p + 2q - 3r, w_2 \le 1 - p - 3q + 2r \\ \text{for some } p, q, r \text{ with } 0 \le p, q, r \le 1, \ 0 \le p + q + r \le 1\}.$$

It is impossible that both inequalities $1 - p + 2q - 3r > 2$ and $1 - p - 3q + 2r > 2$ simultaneously hold for some p, q and r with $0 \le p + q + r \le 1$. Therefore, coalition {1, 2} can not improve upon the Nash bargaining solution (C_1, C_2, C_3). The same result holds for any other two-person coalition.

Finally, we remark that the Nash bargaining solution (C_1, C_2, C_3) does not belong to the Nash core if the payoff vector for the action profile $(D_1.D_2, D_3)$ is changed from (1, 1, 1) to (-1, -1, -1) in the game G. In the new game, if player 3 defects, then players 1 and 2 agree to react by the action profile (C_1, C_2), which is the Nash bargaining solution of their own negotiation problem with the disagreement point (-1, -1). Then, player 3 obtains the higher payoff 6. This means that player 3 can improve upon the Nash bargaining solution (C_1, C_2, C_3) of G. In the new game, the Nash bargaining solution (C_1, C_2, C_3) can not be supported by a limit totally efficient SSPE of the bargaining model.

5 Proofs

In this section, we will prove the results with help of several lemmas.

Lemma 5.1. Let $s^* = (s_1^*, \cdots, s_n^*)$ be an SSPE of $\Gamma^{\varepsilon,\theta}$, and let $q^* = (q_1^*, \cdots, q_n^*)$ be a mixed action profile of G which is played by s^* in the choosing action phase when no agreements have been reached in the negotiation phase. Then q^* must be a Nash equilibrium of G.

Proof. When no agreements have been reached in the negotiation phase, all n players select their actions independently in the choosing action phase, and thereafter the whole bargaining process of $\Gamma^{\varepsilon,\theta}$ ends. This rule of $\Gamma^{\varepsilon,\theta}$ implies that the subgame perfect equilibrium s^* of $\Gamma^{\varepsilon,\theta}$ must prescribe a Nash equilibrium of G in the choosing action phase when no agreements have been reached. Q.E.D.

In what follows, we fix the Nash equilibrium $q^* = (q_1^*, \cdots, q_n^*)$ of G given by an SSPE s^* in case of no agreements, and assume that q^* satisfies Assumption 2.1.(iii). We denote the expected payoffs of players for q^* by $d = (d_1, \cdots, d_n)$. It will be shown that $d = (d_1, \cdots, d_n)$ becomes the disagreement point of the Nash bargaining solution when all players are active in negotiations. When the grand coalition N is formed in s^*, it holds that the SSPE payoff $v = (v_1, \cdots, v_n)$ of s^* satisfies $v_i \geq d_i$ for all $i \in N$ (if $v_i < d_i$ for some i, i will obtain the expected payoff $(1 - \varepsilon)v_i + \varepsilon d_i$ higher than v_i by rejecting the proposal). If d is a Pareto-efficient point of $F(G)$, then $v = d$ must hold. In this case, the efficient SSPE of $\Gamma^{\varepsilon,\theta}$ is characterised trivially such that every player obtains the disagreement payoff d on the equilibrium play of s^*, independent of whether or not an agreement is reached. Therefore, without loss of generality, we can assume:

Assumption 5.1. The disagreement payoff d in an SSPE s^* of $\Gamma^{\varepsilon,\theta}$ is Pareto-inefficient in the feasible set $F(G)$ of G.

The following lemma proves Proposition 4.1 which shows no delay of agreement in every SSPE of $\Gamma^{\varepsilon,\theta}$.

Lemma 5.2. In every SSPE $s^* = (s_1^*, \cdots, s_n^*)$ of $\Gamma^{\varepsilon,\theta}$, every player's proposal is accepted in the initial round of the negotiation phase.

Proof. Let $v = (v_1, \cdots, v_n)$ be the expected payoffs of players for s^*, and let $F(G)$ be the feasible set of G. We note that $v \in F(G)$ since $F(G)$ is convex and v is a convex combination of a finite number of points in $F(G)$. By Assumption 5.1, there exists some $y = (y_1, \cdots, y_n)$ in $F(G)$ such that $y_i > d_i$ for all $i \in N$. Since y and v are in the convex set $F(G)$, it holds $(1-\varepsilon)v + \varepsilon y \in F(G)$ for any ε with $0 < \varepsilon < 1$. Then, select $p^N \in \Delta(A^N)$ such that $u_j(p^N) = (1-\varepsilon)v_j + \varepsilon y_j$ for all $j \in N$. Since $y_j > d_j$ for any j, we have

$$u_j(p^N) > (1-\varepsilon)v_j + \varepsilon d_j \quad \text{for all} \quad j \in N. \tag{5.5}$$

Suppose that every player i proposes (N, p^N). Since s^* is an SSPE of $\Gamma^{\varepsilon,\theta}$, the right-hand side of (5.5) is the expected payoff that player $j (\neq i)$ can obtain by rejecting the proposal (N, p^N). (5.5) implies that every player i's proposal (N, p^N) is accepted by all other players. This fact implies that player i's equilibrium proposal (not necessarily equal to (N, p^N)) must be accepted on equilibrium play of the SSPE s^*. Q.E.D.

Lemma 5.3. Let $s^* = (s_1^*, \cdots, s_n^*)$ be an efficient SSPE of $\Gamma^{\varepsilon,\theta}$, $v = (v_1, \cdots, v_n)$ the expected payoffs of players for s^*, and $d = (d_1, \cdots, d_n)$ the disagreement payoff of s^*. In s^*, every player $i \in N$ initially proposes a pair $(N,\, p_i)$ where $p^i \in \Delta(A^N)$ is the optimal solution of the maximization problem

$$\begin{aligned} \max \quad & u_i(p) \\ \text{subject to} \quad & (1)\ p \in \Delta(A^N) \\ & (2)\ u_j(p) \geq (1-\varepsilon)v_j + \varepsilon d_j \quad \text{for all } j \in N,\ j \neq i. \end{aligned} \tag{5.6}$$

Moreover, the proposal $(N,\, p^i)$ is accepted.

Proof. Let $c_j^\varepsilon \equiv (1-\varepsilon)v_j + \varepsilon d_j$ denote the RHS of the second constraint in (5.6). If responder j is offered more than c_j^ε, then it is optimal for her to accept the proposal. (5.6) can be reformulated as

$$\begin{aligned} \max \quad & h_i(x_{-i}) \\ \text{subject to} \quad & (1)\ x_{-i} \in F_{-i}(G) \\ & (2)\ x_j \geq c_j^\varepsilon \quad \text{for all } j \in N,\ j \neq i \end{aligned}$$

Recall that $h_i(x_{-i}) = \max\{x_i \mid (x_i, x_{-i}) \in F(G)\}$. The function h_i is continuous from Assumption 2.1(i). Let $x_{-i}^* \in R^{N-\{i\}}$ be the optimal solution of the problem above. It must hold from Assumptions 2.1.(ii) and (iii) that $x_j^* = c_j^\varepsilon$ for all $j \neq i$. For any $\varepsilon > 0$, $(c_j^\varepsilon)_{j \in N}$ is an interior point of the feasible set $F(G)$ (note that $v_j \geq d_j$

for all $j \in N$). Then, it holds from the continuity of h_i that for any sufficiently small $\delta_i > 0$, there exists $\delta_j > 0$ for all $j \neq i$ such that $h_i(x^*_{-i} + \delta_{-i}) \geq h_i(x^*_{-i}) - \delta_i$ where $\delta_{-i} = (\delta_j)_{j \neq i}$. This inequality means that if player i proposes the grand coalition N and the correlated action attaining payoffs $(x^*_{-i} + \delta_{-i}, h_i(x^*_{-i} + \delta_{-i}))$, this proposal is accepted and thus player i can obtain more than $h_i(x^*_{-i}) - \delta_i$. Since $\delta_i > 0$ can be chosen arbitrarily small, we can show that player i proposes the optimal solution p^i of (4.6) in the efficient SSPE s^* of $\Gamma^{\varepsilon,\theta}$. Lemma 5.1 shows that the proposal is accepted. Q.E.D.

Lemmata 5.2 and 5.3 characterize the equilibrium proposal of every player in an efficient SSPE of $\Gamma^{\varepsilon,\theta}$. We note that the optimal solution of the maximization problem in Lemma 5.3 gives only a necessary condition for the efficient SSPE proposal for every player i since the optimality of proposing the grand coalition N is not examined. Since player i can propose any subcoalition S of N, we must guarantee that the grand coalition N is actually the optimal proposal. This will be done in Theorem 4.2 where the Nash core plays an important role. Before going to the proof of Theorem 4.2, we will prove that the maximization problem in Lemma 5.3 characterizes the asymmetric Nash bargaining solution of G as the probability ε of negotiation failure goes to zero.

Lemma 5.4. Let $v = (v_1, \cdots, v_n)$ be a limit of efficient SSPE payoffs $v^\varepsilon = (v_1^\varepsilon, \cdots, v_n^\varepsilon)$ of $\Gamma^{\varepsilon,\theta}$ as ε goes to zero. Then,

$$\frac{v_1 - d_1}{\theta_1} \cdot \frac{\partial H}{\partial x_1}(v) = \cdots = \frac{v_n - d_n}{\theta_n} \cdot \frac{\partial H}{\partial x_n}(v) \tag{5.7}$$

$$H(v) = 0 \tag{5.8}$$

where $\theta = (\theta_1, \cdots, \theta_n)$ is the probability distribution which selects a proposer from the player set N, and $d = (d_1, \cdots, d_n)$ is the disagreement payoff of an efficient SSPE in $\Gamma^{\varepsilon,\theta}$ (independent of ε).

Proof. Let x_i^ε denote the payoff that every player $i \in N$ demands for herself in the initial round of the negotiation phase when the efficient SSPE of $\Gamma^{\varepsilon,\theta}$ is played. By Lemma 5.3, we can show that for every $i \in N$

$$H((1-\varepsilon)v_1^\varepsilon + \varepsilon d_1, \cdots, x_i^\varepsilon, \cdots, (1-\varepsilon)v_n^\varepsilon + \varepsilon d_n) = 0. \tag{5.9}$$

Also, by Lemma 5.3 and the definition of v^ε, we can obtain

$$v_i^\varepsilon = \theta_i x_i^\varepsilon + (1-\theta_i)[(1-\varepsilon)v_i^\varepsilon + \varepsilon d_i], \quad \text{for all} \quad i = 1, \cdots, n. \tag{5.10}$$

For each $i \in N$, define $z^{\varepsilon,i} \in F(G)$ as

$$z^{\varepsilon,i} = ((1-\varepsilon)v_1^\varepsilon + \varepsilon d_1, \cdots, x_i^\varepsilon, \cdots, (1-\varepsilon)v_n^\varepsilon + \varepsilon d_n). \tag{5.11}$$

$z^{\varepsilon,i}$ is the payoff vector proposed by player i in the initial round of the negotiation phase in the efficient SSPE s^ε of $\Gamma^{\varepsilon,\theta}$. For any $i,\ j \in N (i \neq j)$, we have from (5.9)

$$H(z^{\varepsilon,i}) - H(z^{\varepsilon,j}) = 0.$$

By Taylor's theorem, there exsits some $\lambda,\ 0 < \lambda < 1$, such that

$$\begin{aligned} 0 &= H(z^{\varepsilon,i}) - H(z^{\varepsilon,j}) \\ &= [x_i^\varepsilon - (1-\varepsilon)v_i^\varepsilon - \varepsilon d_i] \cdot \frac{\partial H}{\partial x_i}(\lambda z^{\varepsilon,i} + (1-\lambda) z^{\varepsilon,j}) \\ &\quad + [(1-\varepsilon)v_j^\varepsilon + \varepsilon d_j - x_j^\varepsilon] \cdot \frac{\partial H}{\partial x_j}(\lambda z^{\varepsilon,i} + (1-\lambda) z^{\varepsilon,j}). \end{aligned} \tag{5.12}$$

(5.10) yields

$$x_i^\varepsilon - (1-\varepsilon)v_i^\varepsilon - \varepsilon d_i = \frac{1}{\theta_i}[v_i^\varepsilon - (1-\varepsilon)v_i^\varepsilon - \varepsilon d_i] = \frac{\varepsilon}{\theta_i}(v_i^\varepsilon - d_i). \tag{5.13}$$

By substituting (5.13) into (5.12), we can prove

$$\frac{v_i^\varepsilon - d_i}{\theta_i}\frac{\partial H}{\partial x_i}(\lambda z^{\varepsilon,i} + (1-\lambda) z^{\varepsilon,j}) = \frac{v_j^\varepsilon - d_j}{\theta_j}\frac{\partial H}{\partial x_j}(\lambda z^{\varepsilon,i} + (1-\lambda) z^{\varepsilon,j}). \tag{5.14}$$

By assumption, we have $\lim_{\varepsilon\to 0} v^\varepsilon = v$, which implies from (5.10) that $\lim_{\varepsilon \to 0} x_i^\varepsilon = v_i$ for all i . Thus, it follows from (5.11) that

$$\lim_{\varepsilon\to 0} v^\varepsilon = \lim_{\varepsilon\to 0} z^{\varepsilon,1} = \cdots = \lim_{\varepsilon\to 0} z^{\varepsilon,n} = v. \tag{5.15}$$

We can prove (5.7) from (5.14) and (5.15), and can prove (5.8) from (5.9) and (5.15). Q.E.D.

In view of (5.9) and (5.10), the efficient SSPE payoffs $v^\varepsilon = (v_1^\varepsilon, \cdots, v_n^\varepsilon)$ of $\Gamma^{\varepsilon,\theta}$ is characterized as a solution of

$$v_i^\varepsilon = \theta_i \cdot h_i((1-\varepsilon)v_{-i}^\varepsilon + \varepsilon d_{-i}) + (1-\theta_i)\cdot\{(1-\varepsilon)v_i^\varepsilon + \varepsilon d_i\} \qquad \text{for all } i \in N \tag{5.16}$$

(5.16) is called the *equilibrium equation* of the efficient SSPE payoffs of $\Gamma^{\varepsilon,\theta}$.

We are now ready to prove Theorems 4.1 and 4.2.

Proof of Theorem 4.1. The maximization problem in Definition 2.2 is reformulated as

$$\begin{aligned} \max_x \textstyle\sum_{i=1}^n \quad & \theta_i \cdot \log(x_i - d_i) \\ \text{subject to} \quad & (1)\ H(x_1, \cdots, x_n) \geq 0 \\ & (2)\ x_i \geq d_i \quad \text{for all} \quad i \in N. \end{aligned}$$

By Assumption 5.1, the optimal solution $x^* = (x_1^*, \cdots, x_n^*) \in R^N$ satisfies $H(x_1^*, \cdots, x_n^*) = 0$ and $x_i^* > d_i$ for all $i \in N$. Therefore, the Kuhn-Tucker condition gives

$$\frac{\theta_i}{x_i^* - d_i} - \lambda \frac{\partial H}{\partial x_i}(x^*) = 0, \quad i = 1, \cdots, n$$
$$H(x^*) = 0$$

where λ is the Lagrange multiplier. From the concavity of $H(x_1, \cdots, x_n)$ and Assumption 5.1, x^* is the optimal solution of the maximization problem if and only if x^* satisfies the Kuhn-Tucker condition. Together with this fact, Lemma 5.4 proves the theorem. Q.E.D.

Proof of Theorem 4.2. Let ϕ^* be the payoff configuration of Γ^θ generated by a limit totally efficient SSPE $s^* = (s_1^*, \cdots, s_n^*)$. Let $s^\varepsilon = (s_1^\varepsilon, \cdots, s_n^\varepsilon)$ be totally efficient SSPEs of $\Gamma^{\varepsilon,\theta}$ which converges to $s^* = (s_1^*, \cdots, s_n^*)$ as ε goes to zero. By the same proof as Lemma 5.1, we can show that for every correlated action $p_S \in \Delta(A^S)$ of every coalition S, s^ε induces a Nash equilibrium of the subgame $G(p_S)$ of G when negotiations break down among all players in $N - S$. Let $d(p_S)$ denote the payoffs of such a Nash equilibrium, and let d denote the disagreement configuration of G which assigns $d(p_S)$ to every subgame $G(p_S)$ of G. Let $\Gamma^{\varepsilon,\theta}(p_S)$ denote a subgame of $\Gamma^{\varepsilon,\theta}$ which starts after agreement (S, p_S) is reached. By applying Theorem 4.1 to every subgame $\Gamma^{\varepsilon,\theta}(p_S)$, we can show that the payoff configuration ψ^* satisfies (i).

We will next prove (ii). Let $x^* = (x_1^*, \cdots, x_n^*) \in R^N$ be the payoff vector which the payoff configuration ϕ^* assignes to the game G. For notational simplicity, we will prove only that x^* belongs to the Nash core of G defined by the Nash bargaining solution configuration b^* with θ and d. The same proof can be easily applied to the payoff vector $\phi^*(p_S)$ which the payoff configuration ϕ^* assignes to every correlated action p_S of every S. Suppose that x^* does not belong to the Nash core of G. By the definition of the Nash core, there exists some coalition $T \subset N$ and some payoff vector $y \in v^{Nash}(T)$ such that

$$y_i > x_i^* \quad \text{for all} \quad i \in T, \tag{5.17}$$

where v^{Nash} is the Nash characteristic function (see Definition 2.5). By the definition of v^{Nash}, the fact that $y \in v^{Nash}(T)$ means that there exists some correlated action $p_T \in \Delta(A^T)$ of T such that

$$u_i(p_T, b^*(p_T)) \geq y_i \quad \text{for all} \quad i \in T \tag{5.18}$$

where $b^*(p_T)$ is the Nash bargaining solution of the subgame $G(p_T)$. Let $\phi^\varepsilon(p_T)$ be the correlated action employed by the complementary coalition $N - T$ in the totally efficient SSPE s^ε of $\Gamma^{\varepsilon,\theta}$ after p_T is agreed by the coalition T. By Theorem 4.1, we can show that

$$\lim_{\varepsilon \to 0} \phi^\varepsilon(p_S) = b^*(p_S). \tag{5.19}$$

Let $x^\varepsilon = (x_1^\varepsilon, \cdots, x_n^\varepsilon)$ be the payoff vector of the totally efficient SSPE s^ε. Then,

$$\lim_{\varepsilon \to 0} x^\varepsilon = x^*. \tag{5.20}$$

In view of (5.17), (5.18), (5.19) and (5.20), it holds that for sufficiently small $\varepsilon > 0$

$$u_i(p_T, \phi^\varepsilon(p_T)) > x_i^\varepsilon \quad \text{for all} \quad i \in T. \tag{5.21}$$

Now, suppose that player $i \in T$ deviates from s^ε and proposes (T, p_T). If it is agreed upon, then all responders j in T receive the payoff $u_j(p_S, \phi^\varepsilon(p_T))$ since thereafter the complementay coalition $N - S$ reacts to S by choosing $\phi^\varepsilon(p_T)$. If the proposal (S, p_S) is rejected, they receive the continuation payoff $(1-\varepsilon)x_j^\varepsilon + \varepsilon d_j$, which is smaller than x_j^ε (note that $x_j^\varepsilon > d_j$). From (5.21), it is optimal for all responders in T to accept (T, p_T). Therefore, on the equilibrium play of s^ε, the proposal (T, p_T) is agreed and the proposer i is better-off. This contradicts that s^ε is an SSPE of $\Gamma^{\varepsilon,\theta}$. Q.E.D.

To prove Theorem 4.3, we first establish that there exists a solution for the equilibrium equation (5.16) of the efficient SSPE of $\Gamma^{\varepsilon,\theta}$ by the Brouwer's fixed point theorem.

Lemma 5.5 Let $v = (v_1, \cdots, v_n)$ be the Nash bargaining solution payoffs of G with the weight vector $\theta = (\theta_1, \cdots, \theta_n)$ and the disagreement point $d = (d_1, \cdots, d_n)$. For any sufficiently small $\varepsilon > 0$, there exists a solution $v^\varepsilon = (v_i^\varepsilon)_{i \in N} \in F(G)$ of (5.16) such that v^ε converges to v as ε goes to zero.

Proof. Let $F^* = \{x \in F(G) \mid x_i \geq d_i \text{ for all } i \in N\}$. For every $x \in F^*$ and every $i \in N$, define

$$g_i^\varepsilon(x) = \theta_i \cdot h_i((1-\varepsilon)x_{-i} + \varepsilon d_{-i}) + (1-\theta_i) \cdot \{(1-\varepsilon)x_i + \varepsilon d_i\}. \tag{5.22}$$

It can be proved that $g^\varepsilon(x) = (g_1^\varepsilon(x), \cdots, g_n^\varepsilon(x))$ is a continuous function from the compact convex subset F^* of R^n to itself. Then, by Brouwer's fixed point theorem, there exists a fixed point $v^\varepsilon \in F^*$ of g^ε satisfying (5.16). Since F^* is a compact set, there exists some converging subsequence of $\{v^\varepsilon\}$. Take any such subsequence of $\{v^\varepsilon\}$. Let $\bar{v}$ denote its limit. By the same proof as in Theorem 4.1 (and Lemma 5.4), we can prove $\bar{v} = v$. This implies that the sequence $\{v^\varepsilon\}$ itself has the limit $\bar{v}$. Q.E.D.

Let b^* be the Nash bargaining solution payoff configulation of G. By applying the same proof as Lemma 5.5 to every subgame $G(p_S)$ of G, we can show that there exists a solution for the equilibrium equation of an efficient SSPE of the subgame $\Gamma^{\varepsilon,\theta}(p_S)$ of $\Gamma^{\varepsilon,\theta}$. Let $v^\varepsilon(p_S)$ denote the solution. Lemma 5.5 also shows that $v^\varepsilon(p_S)$ converges to the Nash bargaining solution payoff $b^*(p_S)$ of $G(p_S)$.

Proof of Theorem 4.3 Let d be a disagreement configuration of G satisfying (A). For every correlated action $p_S \in \Delta(A^S)$ of every coalition S, let $d(p_S) \in F(G(p_S))$ denote the disagreement point which the configuration d assigns to the subgame $G(p_S)$ of G. With abuse of notation, we also denote by $d = (d_1, \cdots, d_n)$ the disagreement point in G assigned by the disagreement configuration d.

Define every player i's strategy s_i^ε in $\Gamma^{\varepsilon,\theta}$ as follows.

(1) When no coalition forms,

(i) propose the grand coalition N and the correlated action yielding the payoff vector in (5.11)

$$z^{\varepsilon,i} = (h_i((1-\varepsilon)v_{-i}^\varepsilon + \varepsilon d_{-i}), (1-\varepsilon)v_{-i}^\varepsilon + \varepsilon d_{-i}),$$

where $v^\varepsilon = (v_i^\varepsilon)_{i \in N} \in F(G)$ is a solution of the equilibrium equation (4.16) (of which existence is proved in Lemma 5.5).

(ii) accept any proposal yielding a payoff not less than $(1-\varepsilon)v_i^\varepsilon + \varepsilon d_i$,

(iii) employ the Nash equilibrium of G given by the disagreement configuration d when negotiations break down.

(2) When some coalition S forms and some correlated action $p_S \in \Delta(A^S)$ of S is agreed, the strategy s_i^ε is defined by the same way as above except that N and v^ε are replaced with $N - S$ and $v^\varepsilon(p_S)$, respectively. When more than one coalition form, s_i^ε is defined in a similar way by taking S as the union of coalitions.

Let ϕ^ε be the payoff configuration generated by the strategy profile $s^\varepsilon = (s_1^\varepsilon, \cdots, s_n^\varepsilon)$ constructed above. Since $v^\varepsilon(p_S)$ is a solution for the equilibrium equation of an efficient SSPE of $\Gamma^{\varepsilon,\theta}(p_S)$ for every $p_S \in \Delta(A_S)$, we can show that $\phi^\varepsilon(p_S) = v^\varepsilon(p_S)$, and that $\phi^\varepsilon(p_S)$ converges to the Nash bargaining solution payoff $b^*(p_S)$ of $G(p_S)$ with θ and d when ε goes to zero.

What remains to be proved is that the strategy profile $s^\varepsilon = (s_1^\varepsilon, \cdots, s_n^\varepsilon)$ is an SSPE of $\Gamma^{\varepsilon,\theta}(p_S)$. For this purpose, it is sufficient to prove that player i's proposal $z^{\varepsilon,i}$ is optimal given s^ε. For each $j \in N$, let $z_j^{\varepsilon,i}$ denote the j-th vector of player i's proposal $z^{\varepsilon,i}$, that is,

$$z_i^{\varepsilon,i} = h_i((1-\varepsilon)v_{-i}^\varepsilon + \varepsilon d_{-i}), \quad z_j^{i,\varepsilon} = (1-\varepsilon)v_j^\varepsilon + \varepsilon d_j, \ j \neq i.$$

Since the disagreement point $d = (d_1, \cdots, d_n)$ of G is an interior point of $F(G)$ from Assumption 5.1, $(1-\varepsilon)v^\varepsilon + \varepsilon d$ is also an interior point of $F(G)$ (note that $F(G)$ is a convex set of R^N). This implies that $h_i((1-\varepsilon)v_{-i}^\varepsilon + \varepsilon d_{-i}) > (1-\varepsilon)v_i^\varepsilon + \varepsilon d_i$ for every $i \in N$. Then, it follows from (5.16) that $z_i^{\varepsilon,j} < v_i^\varepsilon < z_i^{\varepsilon,i}$ for any $j \neq i$. Also, we can see from Lemma 5.5 that v^ε and every $z^{\varepsilon,i}$ converge to the Nash bargaining solution payoffs v of G with the weights $\theta = (\theta_1, \cdots, \theta_n)$ and the disagreement point

$d = (d_1, \cdots, d_n)$ as ε goes to zero. Since v belongs to the interior (relative to the upper-right boundary H of the feasible set $F(G)$) of the strict Nash core of G and $z^{\varepsilon,i}$ belongs to the boundary H, we can see that $z^{\varepsilon,i}$ also belongs to the (relative) interior of the strict Nash core for any sufficiently small ε. Take any coalition S and any correlated action p_S of S. By definition, the payoff vector $u = (u_j(p_S, b^*(p_S))_{j \in N}$ is Nash-effective for S, that is, $u \in v^{Nash}(S)$. The fact that $z^{\varepsilon,i}$ is in the strict Nash core implies that if $u_j(p_S, b^*(p_S)) \geq z_j^{\varepsilon,i}$ for all $j \in S, j \neq i$, then $z_i^{\varepsilon,i} \geq u_i(p_S, b^*(p_S))$. Otherwise, u dominates $z^{\varepsilon,i}$ via S in the strict sense with respect to v^{Nash}. Therefore, $z_i^{\varepsilon,i}$ is the optimal value (attained by $S = N$) of the maximization problem

$$\begin{aligned} \max \ & u_i(p_S, b^*(p_S)) \\ \text{subject to} \quad & (1)\ S \subset N,\ p_S \in \Delta(A^S) \\ & (2)\ u_j(p_S, b^*(p_S)) \geq z_j^{\varepsilon,i} \quad \text{for all} \quad j \in S,\ j \neq i. \end{aligned}$$

This means that the strategy s_i^ε prescribes the optimal proposal of player i. It is clear that s_i^ε prescribes the optimal action for responders. By applying the same proof to all subgames of $\Gamma^{\varepsilon,\theta}$ starting after some agreement has been reached, we can prove that $s^\varepsilon = (s_1^\varepsilon, \cdots, s_n^\varepsilon)$ is a totally efficient SSPE of $\Gamma^{\varepsilon,\theta}$, and for every correlated action $p_S \in \Delta(A^S)$ of S, the expected payoff $v^\varepsilon(p_S)$ converges to the Nash bargaining solution payoffs $b^*(p_S)$ of $G(p_S)$ as ε goes to zero. Q.E.D.

6 Concluding Remarks

We have presented a noncooperative foundation of the asymmetric Nash bargaining solution for a general cooperative game where players can form coalitions and their payoffs depend on what other players do outside coalitions. In this general cooperative situation, a strategic interaction between one coalition and its complementary coalition plays a critical role in determining a final outcome of the game. Unlike the classic theory of von Neumann and Morgenstern, our noncooperative approach requires that any coalitional behavior be consistent with members' payoff maximization. Our analysis has focused an efficient equilibrium where all active players in negotiations form the largest (efficient) coalition, independent of history. The main theorem shows that the Nash bargaining solution can be supported by the efficient equilibrium of the bargaining model where the probability of negotiation failure is very small, if and only if the Nash bargaining solution belongs to the Nash core of the game. The Nash core is defined by the standard core concept under the supposition that a threat by the complementary coalition should be consistent with the Nash bargaining solution theory.

We conclude the paper with a few remarks. First, in our result, the disagreement point of the Nash bargaining solution is determined by a Nash equilibrium of a strategic-form game which is the primitive model of our cooperative situations.

Obviously, in order to derive a unique outcome of rational behavior in a general cooperative game, we need an equilibrium selection theory of a game. Secondly, the weights of the asymmetric Nash bargaining solution is given by the probability distribution to select a proposer. A natural question is how such a probability distribution is determined in a real situation. In our view, this question is truly empirical, beyond the scope of this paper. Some social and political factors may determine the probability distribution. For example, in local communities, a seniority rule (older persons propose more often than younger ones) tends to prevail. In international negotiations, countries with larger populations (or GDPs) may be given more opportunities to make proposals. Thirdly, our result shows that a totally efficient equilibrium does not necessarily exist. In a game without the totally efficient equilibrium, more than one coalition form. Then, an issue of renegotiations should be studied. Players may want to renegotiate their on-going agreements to attain a Pareto-improving payoff allocation. In Okada (2000), we considered the problem of renegotiations in coalitional bargaining in a transferable utility game in characteristic function form. It is shown that the possibility of successive renegotiations necessarily leads to an efficient allocation when the prevailing agreement is considered as the threat point of renegotiation. The possibility of renegotiation, however, has a negative effect in distorting the equity of a final allocation by inducing the first-mover rent. Finally, we remark that our bargaining model has a sequential structure of moves where a coalition and its strategy are simultaneously agreed. The model can capture one of essential aspects in coalitional bargaining that a coalition is reacted by the complementary coalition. An alternative model has a different structure of moves that a coalition structure is determined in the first stage and an action profile of coalitions is determined in the second stage. It is an interesting project to compare different models of coalitional bargaining.

References

[1] Aumann, R.J. (1967) A Survey of Cooperative Games without Side Payments. *Essays in Mathematical Economics in Hornor of Oskar Morgenstern*, ed. M. Shubik, Princeton University Press, Princeton, 3-27.

[2] Aumann, R.J. and Peleg, B. (1960) von Neumann-Morgenstern Solutions to Cooperative Games without Side Payments. *Bulletin of the American Mathematical Society* 66, 173-179.

[3] Baron, D. and J. Ferejohn (1989) Bargaining in Legislatures. *American Political Science Review* 83, 1181-1206.

[4] Bloch, F. (1996) Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division. *Games and Economic Behavior* 14, 90-123.

[5] Chatterjee, K., B. Dutta, D. Ray, and K. Sengupta (1993) A Noncooperative Theory of Coalitional Bargaining. *Review of Economic Studies* 60, 463-477.

[6] Chatterjee, K. and H. Sabourian (2000) Multiperson Bargaining and Strategic Complexity. *Econometrica* 68, 1491-1509.

[7] Gul, F. (1989) Bargaining Foundations of Shapley Value. *Econometrica* 57, 81-95.

[8] Harsanyi, J.C. (1959) A Bargaining Model for the Cooperative n-Person Game. *Contributions to the Theory of Games IV*. ed. by A.W. Tucker and R.D. Luce. Princeton University, Princeton, 325-355.

[9] Harsanyi, J.C. (1963) A Simplified Bargaining Model for the n-Person Cooperative Games. *International Economic Review* 4, 194-220.

[10] Harsanyi, J.C. and R. Selten (1988) *A General Theory of Equilibrium Selection in Games*, The MIT Press, Cambridge.

[11] Hart, S. and A. Mas-Colell (1996) Bargaining and Value. *Econometrica* 64, 357-380.

[12] Moldovanu, B. and E. Winter (1995) Order Independent Equilibria. *Games and Economic Behavior* 9, 1995, 21-34.

[13] Montero, M. (2000) Endogenous Coalition Formation and Bargaining. *Dissertation Series* 71, Center for Economic Research, Tilburg University.

[14] Moulin, H. (1984) Implementing the Kalai-Smorodinsky Bargaining Solution. *Journal of Economic Theory* 33, 32-45.

[15] Nash, J.F. (1950) The Bargaining Problem. *Econometrica* 18, 155-162.

[16] Nash, J.F. (1951) Non-Cooperative Games. *Annals of Mathematics* 54, 286-295.

[17] Nash, J.F. (1953) Two-Person Cooperative Games. *Econometrica* 21, 128-140.

[18] Okada, A. (1996) A Noncooperative Coalitional Bargaining Game with Random Proposers. *Games and Economic Behavior* 16, 97-108.

[19] Okada, A. (2000) The Efficiency Principle in Non-cooperative Coalitional Bargaining. *Japanese Economic Review* 51, 34-50.

[20] Perry, M. and P.J. Reny (1994) A Non-cooperative View of Coalition Formation and the Core. *Econometrica* 62, 795-817.

[21] Rubinstein, A. (1982) Perfect Equilibrium in a Bargaining Model. *Econometrica* 50, 97-109.

[22] Scarf, H.E. (1971) On the Existence of a Cooperative Solution Gor a General Class of N-Person Games. *Journal of Economic Theory* 3, 169-181.

[23] Selten, R. (1981) A Noncooperative Model of Characteristic-Function Bargaining. In V. Boehm and H. Nachthanp (eds.), *Essays in Game Theory and Mathematical Economics in Honor of Oscar Morgenstern*, Bibliographisches Institut Mannheim, 139-151.