GAMES AND ECONOMIC BEHAVIOR **20**, 201–237 (1997)
ARTICLE NO. GA970567

Stable Coalition Structures with Externalities*

Sang-Seung Yi

Department of Economics, Dartmouth College, Hanover, New Hampshire 03755

Received April 12, 1995

This paper argues that the sign of external effects of coalition formation provides a useful organizing principle in examining economic coalitions. In many interesting economic games, coalition formation creates either *negative* externalities or *positive* externalities for nonmembers. Examples of negative externalities are research coalitions and customs unions. Examples of positive externalities include output cartels and public goods coalitions. I characterize and compare stable coalition structures under the following three rules of coalition formation: the Open Membership game of Yi and Shin (1995), the Coalition Unanimity game of Bloch (1996), and the Equilibrium Binding Agreements of Ray and Vohra (1994). *Journal of Economic Literature* Classification Numbers: C72, C71.

1. INTRODUCTION

In recent years, coalition formation has gained increasing prominence across a broad spectrum of economic disciplines, from industrial organization to international trade. For example, research coalitions have become an increasingly important business strategy among oligopolistic firms. The IBM–Apple–Motorola ‘‘PowerPC’’ alliance in the computer industry is a well-publicized example. In international trade, there has been a recent trend toward the formation of regional trading blocs such as the European Union (EU) and the North American Free Trade Agreements (NAFTA) zone. An important feature of these economic coalitions is that they create *externalities* for nonmembers. For example, an important motivation for oligopolistic firms to form research alliances with competitors is to exploit

*This paper merges two working papers, ‘‘Stable Coalition Structures with Negative External Effects’’ and ‘‘Stable Coalition Structures with Positive External Effects.’’ I thank Francis Bloch, Jim Dana, Kenneth Kang, Lindsey Klecan, Eric Maskin, Paul Milgrom, Debraj Ray, Roberto Serrano, Tomas Sjostrom, Rajiv Vohra, Michael D. Whinston, the seminar audiences at Brown, Harvard, SNU, the 1993 Midwest theory conference, the 1995 SITE conference on endogenous coalitions, and the seventh World Congress of the Econometric Society for comments. I am especially grateful to two anonymous referees for detailed suggestions which greatly improved the exposition of this paper.

0899-8256/97 $25.00

complementarities of research assets of alliance partners. If members of a research coalition realize efficiency gains by pooling their complementary research assets, nonmember firms may suffer a competitive disadvantage against member firms. In the case of regional customs unions, the abolition of tariffs on trade among member countries and the readjustment of external tariffs may worsen nonmember countries' terms of trade with member countries.

The recent surge in the formation of economic coalitions with externalities has produced a new strand of literature on the noncooperative theory of coalition formation, which includes Bloch (1995, 1996), Ray and Vohra (1994, 1995), Yi (1996a, 1996b), and Yi and Shin (1995). These models allow for the formation of multiple coalitions and examine the equilibrium number and size of coalitions.[1] They also share the common framework of a two-stage structure. In the first stage, players (for example, oligopolistic firms) form coalitions. In the second stage, players engage in a noncooperative game (for example, a Cournot oligopoly game), given the coalition structure determined in the first stage. Under the simplifying assumption that the second-stage equilibrium is unique for any coalition structure, the second stage game is typically reduced to a *partition function*, which assigns a value to each coalition in a coalition structure as a function of the entire coalition structure, not just the coalition in question. Thus, an important novelty of these models is that they can capture the important possibilities of *externalities* across coalitions. In the traditional characteristic function approach, as in Aumann and Dreze (1974) and Shenoy (1979), these externalities across coalitions are assumed not to be present.

This paper makes two contributions to the field of endogenous coalition formation with externalities. First, I argue that the *sign* of externalities of coalition formation provides a useful organizing principle in examining economic games of coalition formation. I show that, in many interesting economic games, coalition formation creates either *negative* externalities or *positive* externalities on outside coalitions. Examples of positive externalities include output cartels in oligopoly and coalitions formed to provide public goods. Examples of negative externalities are research coalitions with complementary research assets and customs unions in international trade. I also show that the partition function derived from these economic games satisfy other interesting properties.

[1] See also Aumann and Myerson (1988), Chwe (1994), Economides (1986), Greenberg and Weber (1993), Hart and Kurz (1983), and Kamien and Zang (1990). Another recent development in the noncooperative theory of coalition formulation has centered around implementation of cooperative solution concepts, such as the core, Shapley value, and bargaining set. See the survey by Greenberg (1995) and the references therein.

Another contribution of the current paper is the exploration of the stability properties of the rules of coalition formation proposed in the recent models mentioned above. Although these models share the common objective of analyzing equilibrium coalition structures, each adopts a different notion of the stability of a coalition structure. Bloch (1995, 1996) examines an infinite-horizon ‘‘Coalition Unanimity’’ game in which a coalition forms if and only if all potential members agree to form the coalition. Ray and Vohra (1994) study the ‘‘Equilibrium Binding Agreements’’ rule under which coalitions are allowed to break up into smaller subcoalitions only. Yi and Shin (1995) investigate the ‘‘Open Membership’’ game in which nonmembers can join an existing coalition without the permission of the existing members. (Hence, a key difference between these rules of coalition formation lies in what can happen to the membership of a coalition once it is formed: Can an existing coalition break apart, admit new members, or merge with other coalitions?)

Different rules of coalition formation lead to different predictions about stable coalition structures. Due to the absence of a unified framework with which to examine these different approaches, one is left wondering about the underlying causes of these different predictions. The current paper attempts to fill in this gap by examining endogenous coalition formation among symmetric players under some weak conditions on the partition function (which are satisfied for the economic games mentioned above). Particular attention is paid to the analysis of the stability of the grand coalition under different membership rules.

The current paper is organized as follows. Section 2 outlines the two-stage approach to coalition formation among symmetric players. Section 3 briefly introduces the models of Bloch, Ray and Vohra, and Yi and Shin. Section 4 examines equilibrium coalition structures with negative externalities. The main result shows that, under a set of reasonable conditions on the partition function, the grand coalition is an equilibrium outcome under the Open Membership rule, but typically not under the Coalition Unanimity rule nor the Equilibrium Binding Agreements rule. I also identify conditions under which the Coalition Unanimity rule supports a more ‘‘concentrated’’ coalition structure (roughly speaking, a coalition structure with bigger coalitions) than does the Equilibrium Binding Agreements. Section 5 analyzes the opposite case of positive externalities. I show that, due to *free-rider* problems, the grand coalition is rarely an equilibrium outcome under the Open Membership rule. The Coalition Unanimity rule and the Equilibrium Binding Agreements rule do better than the Open Membership rule, but the grand coalition is typically not a stable outcome under these two rules either. Section 6 concludes.

2. COALITION FORMATION AMONG SYMMETRIC PLAYERS WITH EQUAL DIVISION OF COALITION PAYOFF

I analyze a two-stage game of coalition formation, which is the framework shared by Bloch (1996), Ray and Vohra (1994) and Yi and Shin (1995). I adopt the notation of Yi and Shin (1995). In the first stage, players form coalitions. In the second stage, players engage in a noncooperative game given the coalition structure. There are N players, labeled $P_1, P_2, \ldots, P_N$. I start with some definitions and assumptions.

DEFINITION 2.1. A coalition structure $C = \{B_1, B_2, \ldots, B_m\}$ is a partition of the player set $P = \{P_1, P_2, \ldots, P_N\}$. $B_i \cap B_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{N} B_i = P$.

Throughout the paper, I assume that all players are ex ante identical. More formally, let X^i be player i's strategy set in the second-stage game and let $\pi^i: \prod_{i=1}^{N} X^i \to R$ be player i's payoff.

ASSUMPTION 2.1. (1) $X^i = X^j$ for all $i, j = 1, \ldots, N$.

(2) $\pi^i(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N) = \pi^j(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N)$, for all $i, j = 1, \ldots, N$, and $\pi^k(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_N) = \pi^k(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_N)$, $k \neq i, j$, where the second strategy profile is obtained from the first by switching x_i and x_j.

Under Assumption 2.1, each player has the same strategy set in the second-stage game. Furthermore, the *identities* of the players do not matter. Obviously, the assumption of symmetric players is restrictive. Nonetheless, we shall soon see that significant complexities arise in the analysis of stable coalition structures even among symmetric players. In order to further simplify the analysis, I assume that the second-stage game has a unique Nash equilibrium outcome for any coalition structure. Under this assumption, the second-stage game can be reduced to the payoff functions $\pi^i: \mathbf{C} \to R$, where $\mathbf{C}$ is the set of all feasible coalition structures. (For simplicity, I am using the same notation π^i for player i's payoff.)

The symmetry and the uniqueness assumptions imply that, in a given coalition structure, a coalition's payoff (i.e., the sum of payoffs to its members) depends only on the number and the size of coalitions. However, the payoffs do not depend on which player belongs to which coalition. More formally, suppose that players P_i and P_j belong to coalition B_i, player P_k to B_k, and player P_l to B_l, respectively in a coalition structure C, where $i \neq j \neq k \neq l$. Let P_j and P_k switch their coalitions and call the new coalition structure C'. We have $\pi^i(C) = \pi^i(C')$: Player i's second-stage equilibrium payoff stays unchanged. (Of course, the payoffs of players j and k will, in general, be affected.) Similarly, let P_k and P_l

switch their coalitions, and call the new coalition structure C''. Again, $\pi^i(C) = \pi^i(C'')$. Thus, with a slight abuse of notation, I identify a coalition by its size. Specifically, I will write $C = \{n_1, n_2, \ldots, n_m\}$, where n_i is the size of the ith coalition B_i in $C = \{B_1, B_2, \ldots, B_m\}$.

Throughout this paper, I assume *equal sharing* of the coalition payoff among coalition members: Each player in a given coalition receives the same payoff as the other members. That is, I rule out any *side payments* with respect to membership decisions. I rely upon the assumption of ex ante identical players in order to justify the equal division of the coalition payoff. (Recent work by Ray and Vohra (1995) provides a justification for this assumption of equal division of coalition payoff. In an infinite-horizon model of coalition formation among symmetric players with *endogenous* sharing rules, they show that the *equal* sharing of coalition payoff emerges as the equilibrium sharing rule in any equilibrium without delay.)

Under the equal sharing assumption, we can denote by $\pi(n_i; C)$ the *per-member* payoff of a member of the size-n_i coalition in the coalition structure $C = \{n_1, n_2, \ldots, n_m\}$. (Thus, $\pi(n_i; C)$ is a *per-member partition function*. If the payoff of a coalition does not depend on what the rest of the players do, then $\pi(n_i; C) = \pi(n_i)$, and $\prod(n_i) \equiv n_i \pi(n_i)$ is the familiar *characteristic function*, under the added assumption that the identities of coalition members are payoff-irrelevant.) For example, $\pi(3;\{3,2\})$ is the payoff of a member of the size-3 coalition in a coalition structure $\{3,2\}$.

In order to compare the equilibrium coalition structures under different rules of coalition formation, I use the notion of *concentration*, which Yi and Shin (1995) introduced.

DEFINITION 2.2. $C = \{n_1, n_2, \ldots, n_m\}$ is a *concentration* of $C' = \{n'_1, n'_2, \ldots, n'_{m'}\}$, $m' \geq m$, if and only if there exists a sequence of coalition structures $C^1 = \{n^1_1, n^1_2, \ldots, n^1_{m(1)}\}$, $C^2 = \{n^2_1, n^2_2, \ldots, n^2_{m(2)}\}, \ldots, C^R = \{n^R_1, n^R_2, \ldots, n^R_{m(R)}\}$ such that

(1) $C = C^1$ and $C' = C^R$; and

(2) $C^{r-1} = C^r \setminus \{n^r_{i(r)}, n^r_{j(r)}\} \cup \{n^r_{i(r)} + 1, n^r_{j(r)} - 1\}$, $n^r_{i(r)} \geq n^r_{j(r)}$, for some $i(r), j(r) = 1, \ldots, m(r)$ and for all $r = 2, \ldots, R$.

$C = \{n_1, n_2, \ldots, n_m\}$ is a concentration of $C' = \{n'_1, n'_2, \ldots, n'_{m'}\}$ if one can obtain C from C' by a finite sequence of moving one member at a time from a coalition in C' to another coalition of *equal or larger size*. (Notice that $m' - m \geq 0$ coalitions are dissolved in the process.) Concentration, like the usual notion of refinement/coarsening of coalition structures, is a partial ordering. The next result shows that if a coalition structure C is coarser than another coalition structure C', then C is more concentrated than C'.

LEMMA 2.1. *If* $C = \{n_1, n_2, \dots, n_m\}$ *is a coarsening of* $C' = \{n'_1, n'_2, \dots, n'_{m'}\}$, *then* C *is a concentration of* C'.

Proof. Since C is coarser than C', C can be obtained from C' by merging coalitions in C'. Without loss of generality, suppose that $n_1 = n'_1 + \cdots + n'_i$, $n_2 = n'_{i+1} + \cdots + n'_j, \dots,$ and $n_m = n'_k + \cdots + n'_{m'}$. Consider the merger of $n'_1, n'_2, \dots, n'_i$ into n_1. (The other cases are analogous.) Without loss of generality, suppose that $n'_1 \geq n'_2 \geq \cdots \geq n'_i$. Decompose this merger into $n'_2 + \cdots + n'_i$ steps. First, move a member of the size-n'_2 coalition to the size-n'_1 coalition. Second, move a member of the size-$(n'_2 - 1)$ coalition to the size-$(n'_1 + 1)$ coalition. Repeat these steps n'_2 times. Next, move a member of the size-n'_3 coalition to the size-$(n'_1 + n'_2)$ coalition. Repeat these steps n'_3 times, and so on. In each of these steps, the new coalition structure is created by moving a member of a coalition to an equal-sized or larger coalition in the old coalition structure. Q.E.D.

Notice that the reverse of Lemma 2.1 is not true: There exist some coalition structures which cannot be ranked under refinement but which can be ranked under concentration. For example, $\{5, 1\}$ is more concentrated than $\{3, 3\}$, which in turn is more concentrated than $\{2, 2, 2\}$. These three coalition structures cannot be ordered under refinement.

3. RULES OF COALITION FORMATION

3.1. *Open Membership Game*

Yi and Shin (1995) examine a simultaneous-move "Open Membership" game in which membership in a coalition is open to all players who are willing to abide by the rules of the coalition. This game is designed to model an institutional environment in which players are allowed to form coalitions freely, as long as no player is excluded from joining a coalition.

In this game, each player announces an "address" simultaneously. The players that announce the same address belong to the same coalition. Formally, each player's action space is $A^i = \{a_1, a_2, \dots, a_N\}$. For each N-tuple of announcements $\alpha = \{\alpha^1, \alpha^2, \dots, \alpha^N\} \in A \equiv A^1 \times A^2 \times \cdots \times A^N$, the resulting coalition structure is $C = \{B_1, B_2, \dots, B_m\}$, where P_i and $P_j \in B_k$ if and only if $\alpha^i = \alpha^j$: They choose the same address. P_i's payoff is $\pi(n_k; C)$, where n_k is the size of the coalition B_k to which P_i belongs.

3.2. *Infinite-Horizon Coalition Unanimity Game*

Bloch (1996) analyzes what can be called an infinite-horizon sequential-move Coalition Unanimity game in which a coalition forms if and only if all potential members agree to form the coalition. First, P_1 makes a proposal for a coalition, e.g., $\{P_1, P_3, P_4, P_7\}$. Then, the player on P_1's list (not including P_1) with the smallest index—here, it is P_3—accepts or rejects the proposal. If P_3 accepts, then it is P_4's turn to accept or reject the proposal, and the process goes on until we reach the last player on P_1's list. If any of the potential members rejects P_1's proposal, then the current proposal is thrown out (there is no coalition formation among the players who agree to the original proposal), and the player who first rejects the proposal starts over by proposing another coalition. If, instead, all potential members accept P_1's proposal, then they form a coalition. The remaining players continue the coalition formation game, starting with the player with the smallest index making a proposal to the rest of the players. Notice that once a coalition forms, it cannot break apart, admit new members, or merge with other coalitions, regardless of how the rest of the players form coalitions.

Bloch's (1996) main result shows that the infinite-horizon Coalition Unanimity game yields the same stationary subgame perfect equilibrium coalition structure as the following "Size Announcement" game: P_1 first announces the size of his coalition s_1, and the first s_1 players form a size-s_1 coalition, and then P_{s_1+1} proposes s_2, and the next s_2 players form a size-s_2 coalition, and so on until P_N is reached. (See also Ray and Vohra (1995).) Intuitively, this equivalence theorem is a result of the symmetry assumption. In an equilibrium with no delay, P_1 makes a proposal which is going to be accepted immediately. Since the *identities* of the members do not matter, P_1 (and all other subsequent proposers) may as well pick the size of his coalition s_1 under the assumption that the next $s_1 - 1$ players will be his coalition partners. It is straightforward to see that this "Size Announcement" game has a (generically) unique subgame perfect equilibrium coalition structure.

3.3. *Equilibrium Binding Agreements*

Ray and Vohra (1994) conduct an elaborate analysis of equilibrium binding agreements and the stable coalition structures that form under such agreements, under the assumption that *coalitions can only break up into smaller subcoalitions*. Their equilibrium concept, which is defined recursively, is quite involved. But the key idea can be summarized as follows.

First, the degenerate coalition structure $\{1, 1, \dots, 1\}$ is *defined* to be stable. A nondegenerate coalition structure $C = \{n_1, n_2, \dots, n_m\}$ is stable if and only if there do not exist (1) a subcoalition $\hat{n}_i$ of a coalition n_i in C and (2) a more refined coalition structure C' (a stable outcome itself under the Equilibrium Binding Agreements) which can be "induced" by a deviation by these $\hat{n}_i$ "leading perpetrators," such that these $\hat{n}_i$ leading perpetrators are better off under C' than under C. The term "induced" needs to be more carefully defined. The deviation by $\hat{n}_i$ leading perpetrators may result in the breakup of the other coalitions (including the subcoalition consisting of $n_i - \hat{n}_i$ remaining members of the formerly size-n_i coalition) and/or the further breakup of the size-$\hat{n}_i$ subcoalition consisting of the leading perpetrators. (Remember that Ray and Vohra's rule permits the breakup of coalitions only.) This further refinement in coalition structure will occur unless the coalition structure created by the breakup of the size-n_i coalition into size-$\hat{n}_i$ and size-$(n_i - \hat{n}_i)$ subcoalitions is stable. The leading perpetrators look ahead at the end outcome of their deviation and decide to carry out the deviation if they are better off in the final coalition structure induced by their deviation than in the status quo coalition structure.

More formally, a nondegenerate coalition structure $C = \{n_1, n_2, \dots, n_m\}$ is stable under the Equilibrium Binding Agreements rule if and only if there do not exist $C^1 = \{n_1^1, n_2^1, \dots, n_{m(1)}^1\}$, $C^2 = \{n_1^2, n_2^2, \dots, n_{m(2)}^2\}, \dots,$ $C^R = \{n_1^R, n_2^R, \dots, n_{m(R)}^R\}$ such that

(1) $C^1 = C$ and $C^{r+1} = C^r \setminus \{n_{i(r)}^r\} \cup \{\hat{n}_{i(r)}^r, n_{i(r)}^r - \hat{n}_{i(r)}^r\}$, for some $i(r) = 1, \dots, m(r)$ and for all $r = 1, \dots, R-1$;

(2) C^R is stable but $C^2, C^3, \dots, C^{R-1}$ are not; and

(3) $\hat{n}_{i(1)}^1$ leading perpetrators are better off under the final coalition structure C^R than under the original coalition structure $C = C^1$.

4. STABLE COALITION STRUCTURES WITH NEGATIVE EXTERNALITIES

This section examines stable coalition structures for the case of negative external effects, i.e., the case in which the formation (or merger) of coalitions reduces the payoffs of players who belong to other coalitions. I show that some interesting economic coalitions, such as research coalitions with complementary research assets in oligopoly and customs unions in international trade, create negative externalities for nonmember players. In the case of research coalitions, a member firm of a research coalition gains access to the total pool of complementary research assets of all member firms. Hence, the formation of a research coalition confers on member firms a competitive edge against nonmember firms, thereby reducing the profits of nonmember firms. In the case of customs unions, the member countries of a customs union acquire a greater monopoly power in setting the terms of trade against nonmember countries. As a result, the formation of a customs union reduces the welfare of nonmember countries. I then show that the per-member partition function derived from these economic games of coalition formation satisfies other interesting conditions. Under these conditions on the per-member partition function, I characterize equilibrium coalition structures under the three rules of coalition formation discussed above. The main result in this section is that the Open Membership rule supports the grand coalition as the stable coalition structure but the Coalition Unanimity rule or the Equilibrium Binding Agreements rule typically does not.

4.1. *Conditions on the Per-Member Partition Function*: *Negative Externalities*

(N.1) $\pi(n_i; C) > \pi(n_i; C')$, where $\{n_i\} \subset C, C'$ and $C' \setminus \{n_i\}$ can be derived from $C \setminus \{n_i\}$ by merging coalitions in $C \setminus \{n_i\}$.

If coalitions merge to form a larger coalition, *outside* coalitions not involved in the merger are worse off.

Condition (N.1) is the defining feature of coalition formation with negative external effects across coalitions. The next two conditions are about the *internal* effects of changes in the coalition structure (i.e., the effects on players involved in the changes in the coalition structure).

(N.2) $\pi(n_j; C) < \pi(k; C')$, where (1) $\{n_1, n_2, \dots, n_j\} \subseteq C$; (2) $k = \sum_{i=1}^{j} n_i$ and $C' = C \setminus \{n_1, n_2, \dots, n_j\} \cup \{k\}$; and (3) $n_i \geq n_j$ for $i = 1, 2, \dots, j-1$.

A member of a coalition becomes better off if his coalition merges with larger or equal-sized coalitions.

(N.3) $\pi(n_j; C) < \pi(n_i + 1; C')$, where $C' = C \setminus \{n_i, n_j\} \cup \{n_i + 1, n_j - 1\}$, $n_i \geq n_j \geq 2$.

A member of a coalition becomes better off if he leaves his coalition to join another coalition of equal or larger size.

Notice that (N.2) and (N.3) are distinct from each other, because (N.2) concerns the *merger* of coalitions whereas (N.3) concerns an *individual* change in coalition affiliation.[2] It is worthwhile to emphasize that (N.2) does not imply that the merger of coalitions necessarily benefits the members of the *larger* coalitions involved. For example, when two coalitions combine, members of the larger coalition may earn lower payoffs. Similarly, (N.3) does not imply that, when a member of a coalition joins a larger one, the *existing* members of the larger coalition necessarily become better off. (By symmetry, the merger of two *equal-sized* coalitions benefits

[2] For example, consider $C = \{3, 2\}$. Under (N.2), $\pi(2; \{3, 2\}) < \pi(5; \{5\})$: A member of the size-2 coalition becomes better off if his coalition merges with the larger size-3 coalition. Under (N.3), $\pi(2; \{3, 2\}) < \pi(4; \{4, 1\})$: A member of the size-2 coalition becomes better off by leaving his coalition to join the larger size-3 coalition.

all members. Similarly, the existing members of a coalition become better off by admitting a new member from another coalition of equal size.) This fact will become important when comparing the stable coalition structures under different rules of coalition formation.

As mentioned above, condition (N.2) is silent on how a merger of coalitions affects the members of larger coalitions. The following definition concerns the effect of a merger with a one-player coalition on the members of a larger coalition.

DEFINITION 4.1. k_0 is the largest integer which satisfies $\pi(k; C) \geq \pi(k-1; C')$, for *all* coalition structures C and C', $C' = C \setminus \{k\} \cup \{k-1, 1\}$, and for *all* $k, 2 \leq k \leq k_0$.

The integer k_0 is the largest integer such that the existing members of a size-$(k-1)$ coalition, $k - 1 < k_0$, are made better off by merging with a singleton coalition, *holding the rest of the coalition structure fixed*. Two remarks are in order about k_0. First, k_0 is defined to be the largest integer which satisfies $\pi(k; C) \geq \pi(k-1; C')$ for all coalition structures $C \setminus \{k\}$ formed by the other players. For example, consider $N = 5$ and suppose that $\pi(5; \{5\}) < \pi(4; \{4, 1\})$, $\pi(4; \{4, 1\}) < \pi(3; \{3, 1, 1\})$, $\pi(3; \{3, 2\}) < \pi(2; \{2, 2, 1\})$, $\pi(3; \{3, 1, 1\}) > \pi(2; \{2, 1, 1, 1\})$, $\pi(2; \{2, 2, 1\}) > \pi(1; \{2, 1, 1, 1\})$, and $\pi(2; \{2, 1, 1, 1\}) > \pi(1; \{1, 1, 1, 1, 1\})$. In this example, $k_0 = 2$, because $\pi(3; \{3, 2\}) < \pi(2; \{2, 2, 1\})$. However, if $\pi(3; \{3, 2\}) > \pi(2; \{2, 2, 1\})$ with the other inequalities unchanged, then $k_0 = 3$. Second, notice that $k_0 \geq 2$ under (N.2). We will see that k_0 proves useful in characterizing equilibrium coalition structures with negative externalities.

4.2. *Economic Models of Coalition Formation with Negative Externalities*

Assumptions (N.1)–(N.3) are satisfied in many interesting economic games of coalition formation. This subsection illustrates this point by showing that these conditions are satisfied by simple models of (1) research coalitions with complementary research assets in oligopoly and (2) customs unions in international trade.

4.2.1. *Research Coalitions with Complementary Research Assets*

Consider a Cournot oligopoly with inverse demand $P(X) = A - X$, where X is the industry output. There are N ex ante symmetric firms, each of which has one unit of unique research asset or “knowledge.” If a set of firms form a research coalition, they pool their research assets and develop a new technology. They then compete with new technologies in the downstream product market in order to maximize their own profits. Suppose that the cost function under a new technology developed with ω units of research assets is given by $c(x, \omega) = \mu(\omega)x$, where x is output. Assume that $\mu'(\omega) < 0$: The more research assets firms use in developing the new process, the better the new process is. Hence, this model of research and development cooperation captures the efficiency gains from pooling research knowledge, an important motivation for firms to form research joint ventures.[3]

Now suppose that the coalition structure is $C = \{n_1, n_2, \ldots, n_m\}$. Then in the second-stage product–market competition, there are n_1 firms with constant marginal cost $\mu(n_1)$, n_2 firms with constant marginal cost $\mu(n_2), \ldots, n_m$ firms with constant marginal cost $\mu(n_m)$. In the unique Nash equilibrium of the product market, a firm with constant marginal cost $\mu(n_i)$, that is, a member of the size-n_i coalition, earns

$$\pi(n_i; C) = \left(A - (N+1)\mu(n_i) + \textstyle\sum_{j=1}^{m} n_j \mu(n_j)\right)^2 \Big/ (N+1)^2. \quad (4.1)$$

It is straightforward to show that the per-member partition function given in Eq. (4.1) satisfies the conditions in the previous subsection:

LEMMA 4.1. *Research coalitions in the Cournot oligopoly with the inverse demand function* $P(X) = A - X$ *and the cost function* $c(x, \omega) = \mu(\omega)x$, $\mu'(\omega) < 0$, *satisfy* (N.1)–(N.3).

Proof. See Appendix A.

The intuition for Lemma 4.1 is as follows. When coalitions (say the size-n_i and size-n_j coalitions, $n_i \geq n_j$) merge, their members combine their research assets and develop a technology with lower marginal costs. As a result, they steal business from other coalitions, reducing other coalitions' profits [(N.1)]. To see why the merger helps the members of the smaller size-n_j coalition, decompose the change in marginal costs into two steps. First, the marginal costs of members of the size-n_j coalition fall to the level of the members of the size-n_i coalition. Second, the marginal costs of the members of the merged coalition fall to the new, lower level. Both steps increase the profits of the members of the (formerly) size-n_j coalition [(N.2)]. (On the other hand, since the first step reduces and the second step increases the profits of the members of the (formerly) size-n_i coalition, they may earn higher or lower profits as a result of the merger.)

Finally, suppose that a member of the size-n_j coalition leaves his coalition to join the size-n_i coalition, $n_i \geq n_j$. We can decompose the change in the cost structure of the industry into three steps. First, the

[3] This model of research coalitions is an extension of Bloch (1995), who examines a linear $\mu(\omega)$ function: $\mu(\omega) = \overline{\omega} - \omega$, where $\overline{\omega}$ is a positive constant. The results of this section can be generalized to arbitrary downward-sloping demand functions and increasing cost functions. For details, see Yi (1996b).

deviator's marginal cost falls to the level of the members of the size-n_i coalition. Second, the marginal costs of the deviator and the existing n_i members of the (formerly) size-n_i coalition fall to the new, lower level. Third, the marginal costs of the remaining $(n_j - 1)$ members of the (formerly) size-n_j coalition rise to the new, higher level. All three steps increase the profit of the deviator [(N.3)]. (On the other hand, since the first step reduces and the second and third steps increase the profits of the existing members of the (formerly) size-n_i coalition, they may earn higher or lower profits as a result of admitting a new member.)

4.2.2. *Customs Unions in International Trade*

There are N ex ante symmetric countries. Each country produces a homogeneous good at a constant marginal cost c in terms of the numeraire good. The representative consumer in country i has a utility function of the form

$$u^i(Q_i; M_i) = aQ_i - \tfrac{1}{2}Q_i^2 + M_i, \tag{4.2}$$

where Q_i is country i's consumption of the nonnumeraire good and M_i is country i's consumption of the numeraire good. Let τ_{ij} be country i's (nonnegative) specific tariff on imports from country j. Then country j's effective marginal cost of exporting to country i is

$$c_{ij} = c + \tau_{ij}. \tag{4.3}$$

Countries compete by choosing their sales simultaneously in each country. Assume that the profits of the domestic firm and the tariff revenues are rebated back to the consumers. Then country i's welfare (denoted by W^i) consists of four components: the domestic consumer surplus (denoted by CS^i), the domestic firm's profit in home market (denoted by π^{ii}), the domestic firm's export profits (denoted by π^{ji}, $j \neq i$), and the tariff revenue (denoted by TR^i):

$$W^i = CS^i + \pi^{ii} + \sum_{j \neq i} \pi^{ji} + TR^i. \tag{4.4}$$

A customs union is defined as a group of countries with internal free trade and an external common tariff for joint welfare maximization. Suppose that the customs union structure is $C = \{n_1, n_2, \ldots, n_m\}$ and let

$W(n_i; C)$ be the equilibrium welfare of a member of the size-n_i customs union. In Appendix A, I show that

$$W(n_i; C) = \frac{1}{2} - q_o(n_i) + \sum_{\substack{j=1 \\ j \neq i}}^{m} n_j \big[q_o(n_j) \big]^2, \tag{4.5}$$

where

$$q_o(n_j) = \frac{1}{(N+1) + (n_j+1)(2n_j+1)} \tag{4.6}$$

is a nonmember country's equilibrium exports to a member country of the size-n_j customs union, $j = 1, \ldots, m$. The per-member partition function given in Eq. (4.5) satisfies (N.1)–(N.3).

LEMMA 4.2. *Customs unions with the utility function* $u(Q; M) = aQ - (1/2)Q^2 + M$ *and the cost function* $c(q) = cq$ *satisfy* (N.1)–(N.3).

Proof. See Appendix A.

The intuition why (N.1)–(N.3) hold is as follows. Suppose that two customs unions, of size-n_i and size-n_j, $n_i \geq n_j$, merge. Members of the merged customs union abolish tariffs among themselves and impose joint-welfare-maximizing tariffs on outsiders. As a result, terms of trade for outsiders deteriorate, reducing their welfare [(N.1)]. The members of the smaller customs union benefit from this merger, because they obtain a tariff-free access to n_i markets in return for granting members of the size-n_i customs union a tariff-free access to n_j markets, $n_i \geq n_j$ [(N.2)]. Now, suppose that a member of the size-n_j customs union leaves its customs union to join the size-n_i customs union, $n_i \geq n_j$. Essentially, this deviator gives up a tariff-free access to $n_j - 1$ countries in return for obtaining a tariff-free access to n_i countries. This deviator's welfare improves, since an increase in the number of markets with tariff-free access increases a country's welfare [(N.3)].[4]

[4]As in research coalitions with complementary assets, the existing members of the large customs union need not benefit by admitting a member of a small customs union. The reason is that they gain tariff-free access to a single country in return for granting this new member tariff-free access to all existing member countries. Similarly, the members of a large customs union need not gain from the merger with a small customs union.

4.3. *Equilibrium Coalition Structures with Negative Externalities*

4.3.1. *Open Membership Game*

It is easy to see that, under (N.2) and (N.3), the grand coalition is the unique (pure-strategy) Nash equilibrium outcome of the simultaneous-move Open Membership game.

PROPOSITION 4.1. *Assume* (N.2) *and* (N.3). *In the simultaneous-move Open Membership game,* $\{N\}$ *is the unique* (*pure-strategy*) *Nash equilibrium coalition structure*.

Proof. Take a coalition structure $C = \{n_1, n_2, \dots, n_m\}$, $m \geq 2$ and $n_1 \geq n_2 \geq \cdots \geq n_m$. C is not a Nash equilibrium outcome, because a member of the size-n_i coalition, $i \geq 2$, can earn a higher payoff by changing his address to the one announced by the members of the size-n_1 coalition. The grand coalition $\{N\}$ is a Nash equilibrium outcome since no player benefits by changing his address to form a one-player coalition. Q.E.D.

4.3.2. *Infinite-Horizon Coalition Unanimity Game*

Suppose that $s_1, s_2, \dots, s_m$ are the announcements of coalition sizes in the (generically) unique subgame perfect equilibrium of the Size Announcement game, and thus of the Coalition Unanimity game. s_i is the size of the ith coalition to form in the equilibrium path, $i = 1, 2, \dots, m$. I show that the last coalition to form is uniquely the smallest, and that the second-to-last coalition to form is uniquely the second smallest. Thus, a symmetric coalition structure is not an equilibrium outcome. Furthermore, the second-to-last coalition has at least k_0 members so that the number of equilibrium coalitions does not exceed $I(N/k_0)$, where $I(r)$ is the closest integer greater than or equal to r. The proposition also identifies a *necessary* condition for the grand coalition to be the equilibrium outcome.

PROPOSITION 4.2. (1) *Under* (N.2), $s_m < s_j$, $j = 1, \dots, m-1$.

(2) *Under* (N.1), $s_{m-1} \geq k_0$.

(3) *Under* (N.1) *and* (N.2), $s_{m-1} < s_j$, $j = 1, \dots, m-2$.

(4) *Under* (N.1) *and* (N.2), $m \leq I(N/k_0)$.

(5) *Under* (N.1), $\{N\}$ *is not* the subgame perfect equilibrium coalition structure of the infinite-horizon Coalition Unanimity game if there exists $\hat{k}$ such that $\pi(N; \{N\}) < \pi(\hat{k}; \{\hat{k}, N - \hat{k}\})$.

Proof. (1) Suppose not. Then, there exists j, $1 \leq j \leq m-1$, such that $s_m \geq s_j$. Suppose that $s_k = \min\{s_1, s_2, \dots, s_{m-1}\}$. The player announcing s_k

can increase his payoff by instead declaring a grand coalition among the remaining players. That is, let this player announce $s_k + s_{k+1} + \cdots + s_m$ instead of s_k. Since $s_k \leq s_{k+1}, s_{k+2}, \ldots, s_m$, this player earns a higher payoff by (N.2).

(2) Suppose that $s_{m-1} < k_0$. Then $\pi(s_{m-1}; \{s_1, \ldots, s_{m-1}, s_m\}) \leq \pi(s_{m-1}; \{s_1, \ldots, s_{m-1}, 1, s_m - 1\}) < \pi(s_{m-1} + 1; \{s_1, \ldots, s_{m-1} + 1, s_m - 1\})$. The first inequality follows from (N.1) and the second from $s_{m-1} + 1 \leq k_0$. (The nongeneric case where the members of the size-s_{m-1} coalition are indifferent to the merger with a singleton coalition is ignored.) Thus, the second-to-last announcer can earn a higher payoff by declaring $s_{m-1} + 1$ instead of s_{m-1}. (If the next announcer does not propose $s_m - 1$, the deviator becomes even better off by (N.1).)

(3) Suppose not. Then, there exists $j, 1 \leq j \leq m-2$, such that $s_{m-1} \geq s_j$. Suppose that $s_k = \min\{s_1, s_2, \ldots, s_{m-2}\}$. The player announcing s_k can increase his payoff by declaring $s_k + s_{k+1} + \cdots + s_{m-1}$ instead of s_k. To see why, first suppose that the size-s_m coalition does not break up. (That is, following the above deviation, the next announcer chooses s_m.) Then, as in part (1), the deviator is better off by (N.2). Second, if the size-s_m coalition breaks apart following the above deviation, the deviator's payoff increases even more by (N.1).

(4) Suppose that $m \geq I(N/k_0) + 1$. Then $N > \sum_{i=1}^{m-1} s_i \geq (m-1) \cdot k_0 \geq I(N/k_0)k_0 \geq N$, which is a contradiction.

(5) Under (N.1) and the condition stated in the proposition, announcing $\hat{k}$ dominates announcing N for P_1. (If the next player does not announce $N - \hat{k}$ following the announcement of $\hat{k}$ by P_1, then, by (N.1), P_1 earns an even higher payoff than $\pi(\hat{k}; \{\hat{k}, N - \hat{k}\})$.) Q.E.D.

When $k_0 \geq N/2$, the task of identifying the equilibrium coalition structure of the Coalition Unanimity game becomes much simpler, since Proposition 4.2, (4) shows that the number of equilibrium coalitions is at most two. Instead of looking at all feasible coalition structures, one only needs to compare $Q(N/2) + 1$ coalition structures which contain at most two coalitions, where $Q(r)$ is the integer part of r. For P_1, announcing k_0 dominates announcing k', $1 \leq k' < k_0$, because $\pi(k_0; \{k_0, N - k_0\}) \geq \pi(k'; \{k', 1, \ldots, 1, N - k_0\}) > \pi(k'; \{k', N - k'\})$ by the definition of k_0 and (N.1). Thus, the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game is $\{k^u, N - k^u\}$, where $k^u \in \arg\max_{k_0 \leq k \leq N} \pi(k; \{k, N - k\})$: P_1 chooses the "best" coalition structure from $\{k, N - k\}$, $k = k_0, k_0 + 1, \ldots, N$.

Proposition 4.2 shows that the last coalition is the unique smallest coalition and the second-to-last coalition is the unique second-smallest coalition. What about the third-smallest coalition? Is it strictly smaller than the fourth-smallest coalition and is it the third-to-last coalition to form? As is clear from the above proof, under (N.1) and (N.2), it is possible that $s_{m-2} \geq s_j$, for some $j = 1, \dots, m-3$. (For example, suppose that $s_{m-2} \geq s_{m-3}$. Can the announcer of s_{m-3} always be better off by instead choosing $s_{m-2} + s_{m-3}$? The answer is ambiguous because, following this deviation, the next announcer might pick $s_{m-1} + s_m$. Under (N.1), the merger of these last two coalitions reduces the payoff of the deviator.) Similarly, the third-smallest coalition need not be the third-to-last coalition to form.

4.3.3. *Equilibrium Binding Agreements*

The following result shows that, under (N.1), any coalition structure for which the size of the largest coalition is less than or equal to k_0 is a stable coalition structure under the Equilibrium Binding Agreements. It also identifies a *necessary* condition for the grand coalition to be stable under the Equilibrium Binding Agreements.

PROPOSITION 4.3. (1) *Under* (N.1), $C = \{n_1, n_2, \dots, n_m\}$, $n_i \leq k_0$ *for* $i = 1, \dots, m$, *is a stable coalition structure under the Equilibrium Binding Agreements rule*.

(2) *Under* (N.1), $\{N\}$ *is not stable under the Equilibrium Binding Agreements rule if* $k_0 \geq N/2$ *and if there exists* $N/2 \leq \hat{k} \leq k_0$ *such that* $\pi(N; \{N\}) < \pi(\hat{k}; \{\hat{k}, N - \hat{k}\})$.

Proof. (1) Consider $C = \{n_1, n_2, \dots, n_m\}$, $n_1 \geq n_2 \geq \cdots \geq n_m$. The proof proceeds by induction on n_1, the size of the largest coalition (and the number of size-k_0 coalitions). If $n_1 = 1$, then $C = \{1, 1, \dots, 1\}$, which is stable by definition. Now suppose that the claim holds for $n_1 = 1, 2, \dots, k_0 - 1$ and consider $n_1 = k_0$. By combining (N.1), the definition of k_0, and the assumption that $k_0 \geq n_1 \geq n_i$, $i = 2, \dots, m$, we have $\pi(n_i; C) \geq \pi(k; C') \geq \pi(k; C'') \geq \cdots \geq \pi(k; \hat{C})$, where $C' = C \setminus \{n_i\} \cup \{k, 1, 1, \dots, 1\}$, $C'' = C \setminus \{n_i\} \cup \{k, 2, 1, \dots, 1\}, \dots, \hat{C} = C \setminus \{n_i\} \cup \{k, n_i - k\}$, for all $k = 1, 2, \dots, n_i - 1$, and for all $i = 1, 2, \dots, m$. (The chain of coalition structures $C', C'', \dots, \hat{C}$ covers all possible coalition structures that can be obtained by breaking up the size-$(n_i - k)$ subcoalition.) Thus, if outside coalitions do not break apart, members of the size-n_i coalition cannot make themselves better off by breaking up to form smaller coalitions. If all outside coalitions have less than k_0 members, then, by the induction hypothesis, no outside coalitions break up in response to the breakup of the size-n_i coalition. (That is, the coalition structure created by the breakup of the size-n_i coalition is stable.) Now suppose that the number of outside coalitions with k_0 members is s. Further suppose that no size-k_0 coalition breaks up into smaller coalitions in response to the breakup of the size-n_i coalition for $s = 0, 1, \dots, m-2$. (We have just seen that this claim is true for $s = 0$.) For $s = m-1$, no size-k_0 coalition breaks up into smaller coalitions in response to the breakup of the size-n_i coalition by (N.1), the definition of k_0, and the induction hypothesis.

(2) By construction, $k_0 \geq \hat{k}, N - \hat{k}$. Hence, $\{\hat{k}, N - \hat{k}\}$ is stable under the Equilibrium Binding Agreements. $\{N\}$ is not stable since $\hat{k}$ players can profitably leave the grand coalition to form a size-$\hat{k}$ coalition which results in the coalition structure $\{\hat{k}, N - \hat{k}\}$. Q.E.D.

4.3.4. *Comparison of Stable Coalition Structures with Negative Externalities*

As shown above, the grand coalition is the unique pure-strategy Nash equilibrium outcome of the Open Membership game. However, the grand coalition is often *not* an equilibrium outcome of the Coalition Unanimity game or under the Equilibrium Binding Agreements. (That is, the necessary conditions in Propositions 4.2 and 4.3 for the grand coalition to be stable under these two rules are often violated.) Since the grand coalition is more concentrated (indeed, coarser) than any other coalition structure, we obtain the following observation.

Remark 4.1. Assume (N.1)–(N.3). The unique pure-strategy Nash equilibrium coalition structure of the Open Membership game is (weakly) more concentrated than the subgame perfect equilibrium coalition structure of the Coalition Unanimity game or any stable coalition structure under the Equilibrium Binding Agreements. If $k_0 \geq N/2$ and if there exists $N/2 \leq \hat{k} \leq k_0$ such that $\pi(N; \{N\}) < \pi(\hat{k}; \{\hat{k}, N - \hat{k}\})$, then the relationship is strict.

In general, a stable coalition structure under the Equilibrium Binding Agreements can be more concentrated than the subgame perfect equilibrium coalition structure of the Coalition Unanimity game. Consider the following example.

EXAMPLE 4.1. $N = 5$. The possible coalition structures are

$$\underset{50}{\{5\}},\ \underset{51,35}{\{4,1\}},\ \underset{52,39}{\{3,2\}},\ \underset{53,36,36}{\{3,1,1\}},\ \underset{48,48,37}{\{2,2,1\}},\ \underset{58,38,38,38}{\{2,1,1,1\}},\quad \text{and}$$

$$\underset{42,42,42,42,42}{\{1,1,1,1,1\}}.$$

The small numbers below a coalition are the per-member profits of that coalition in that coalition structure. In this example, $k_0 = 2$. In Appendix A, I show that $\{3, 2\}$ is the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game and that $\{4, 1\}$ is stable under the Equilibrium Binding Agreements. $\{4, 1\}$ is more concentrated than $\{3, 2\}$. (Note that $\{4, 1\}$ and $\{3, 2\}$ cannot be ranked under the usual binary relation of refinement, because they have the same number of coalitions.)

The next result identifies a condition under which the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game is at least as concentrated as any stable coalition structure under the Equilibrium Binding Agreements.

PROPOSITION 4.4. *Assume* (N.1) *and* (N.2) *and suppose that* $k_0 \geq N/2$. *Consider* $C = \{n_1, n_2, \ldots, n_m\}$, $n_1 > k_0$. *If there exists* $\hat{k}$, $N/2 \leq \hat{k} \leq k_0$, *such that* $\pi(n_1; C) < \pi(\hat{k}; C')$, $C' = C \setminus \{n_1\} \cup \{\hat{k}, n_1 - \hat{k}\}$, *then the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game is* (*weakly*) *more concentrated than any stable coalition structure under the Equilibrium Binding Agreements*.

Proof. Proposition 4.3 shows that $C = \{n_1, n_2, \ldots, n_m\}$, $n_1 \leq k_0$, is stable under the Equilibrium Binding Agreements. Under the condition stated in the proposition, I show that $C = \{n_1, n_2, \ldots, n_m\}$, $n_1 > k_0$ is *not* stable under the Equilibrium Binding Agreements. To see why, let $\hat{k}$ members leave the size-n_1 coalition to form a size-$\hat{k}$ coalition. Since $\hat{k} \geq N/2$, we have $n_1 - \hat{k}, n_2, \ldots, n_m \leq N/2 \leq k_0$: All other coalitions in the resulting coalition structure have less than or equal to k_0 members. By Proposition 4.3, C' is stable and the deviation is profitable. Hence, C is not stable and $\{k_0, N - k_0\}$ is the most concentrated coalition structure under the Equilibrium Binding Agreements.

Proposition 4.2 and the paragraph following it show that the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game is $\{k^u, N - k^u\}$, where $k^u \in \arg\max_{k_0 \leq k \leq N} \pi(k; \{k, N - k\})$: P_1 chooses the "best" coalition structure from $\{k, N - k\}$, $k = k_0, k_0 + 1, \ldots, N$. Since $k^u \geq k_0$, $\{k^u, N - k^u\}$ is more concentrated than $\{k_0, N - k_0\}$. Q.E.D.

Heuristically speaking, under the conditions of Proposition 4.4, $\{k^u, N - k^u\}$ is stable under the Coalition Unanimity game because, if $\hat{k}$ members of the size-k^u coalition leave it to form a size-$\hat{k}$ coalition, the remaining members of the (previously) size-k^u coalition can "credibly threaten" to form the size-$(N - \hat{k})$ coalition with the size-$(N - k^u)$ coalition. By (N.1), this "retaliation" reduces the payoff of the deviators. Indeed, in Proposition 4.4, this retaliation reduces the payoff of the deviators sufficiently enough to make the deviation not profitable. In contrast, $\{k^u, N - k^u\}$ is not stable under the Equilibrium Binding Agreements rule because $\hat{k}$ members of the size-k^u coalition can earn a higher payoff by breaking off to form a size-$\hat{k}$ coalition, without worrying about the response of the other players. Under the Equilibrium Binding Agreements rule, the remaining $k^u - \hat{k}$ members can only break up into smaller subcoalitions but cannot merge with the other coalition in response to the deviation of $\hat{k}$ leading perpetrators. However, if the remaining members (or members of other coalitions) do break up into smaller subcoalitions, the leading perpetrators become even better off by (N.1).

5. STABLE COALITION STRUCTURES WITH POSITIVE EXTERNALITIES

This section examines stable coalition structures for the case of positive external effects, i.e., the case in which the formation (or merger) of coalitions increases the payoffs of players who belong to other coalitions. Well-known economic coalitions, such as output cartels in oligopoly and coalitions formed to provide public goods, create positive externalities on nonmember players. In the case of output cartels, members of a cartel reduce their aggregate output in order to raise price. Nonmember firms earn higher profits by free-riding on the price increase induced by the output reduction by member firms of a cartel. Similarly, members of a public goods coalition increase their total contributions to the provision of the public good. Nonmember players benefit from the increased supply of the public good without increasing their own contributions. The per-member partition function derived from these classical economic coalitions satisfies other interesting conditions.

Under these conditions on the per-member partition function, I characterize equilibrium coalition structures under the three rules of coalition formation. Unlike the case of negative externalities, the Open Membership game typically does not support the grand coalition as an equilibrium outcome. Indeed, the most concentrated equilibrium coalition structure under the Open Membership game is less concentrated than the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game or the most concentrated stable coalition structure under the Equilibrium Binding Agreements.

5.1. *Conditions on the Per-Member Partition Function: Positive Externalities*

(P.1) $\pi(n_i; C) < \pi(n_i; C')$, where $\{n_i\} \subset C$, C' and $C' \setminus \{n_i\}$ can be derived from $C \setminus \{n_i\}$ by merging coalitions in $C \setminus \{n_i\}$.

If coalitions merge to form a larger coalition, *outside* coalitions not affected by the change are better off.

Condition (P.1) is the cornerstone condition of coalition formation with positive externalities and is the opposite of (N.1). The next condition ranks per-member payoffs of coalitions in a given coalition structure.

(P.2) $\pi(n_i; C) < \pi(n_j; C)$ if and only if $n_i > n_j$.

In any coalition structure, small coalitions have higher *per-member* payoffs than big coalitions.

Now suppose that a member of a coalition leaves his coalition to join a larger or equal-sized one. The next two conditions concern the effect on the remaining members of the (now) smaller coalition and the deviator, respectively.

(P.3) $\pi(n_j; C) < \pi(n_j - 1; C')$, where $C' = C \setminus \{n_i, n_j\} \cup \{n_i + 1, n_j - 1\}$, $n_i \geq n_j \geq 2$.

If a member of the size-n_j coalition leaves his coalition to join a larger or equal-sized coalition, then the remaining members of the (formerly) size-n_j coalition become better off.

(P.4) $\pi(n_j; C) > \pi(n_i + 1; C')$, where $C' = C \setminus \{n_i, n_j\} \cup \{n_i + 1, n_j - 1\}$, $n_i \geq n_j \geq 2$.

If a member of the size-n_j coalition leaves his coalition to join a larger or equal-sized coalition, then the deviator becomes worse off.

Notice that (P.4) is the opposite of (N.3). The following definition is useful in characterizing equilibrium coalition structures with positive external effects.

DEFINITION 5.1. $C = \{n_1, n_2, \ldots, n_m\}$ *is stand-alone stable* if and only if $\pi(n_i; C) \geq \pi(1; C_i)$, $C_i = C \setminus \{n_i\} \cup \{n_i - 1, 1\}$ for all $i = 1, \ldots, m$.

A coalition structure C is stand-alone stable if and only if no player finds it profitable to leave his coalition to form a singleton coalition, *holding* the rest of coalition structure constant (including his former coalition). Notice that, by definition, the degenerate coalition structure $\{1, 1, \ldots, 1\}$ is stand-alone stable.

5.2. *Economic Models of Coalition Formation with Positive Externalities*

This subsection shows that the above four conditions are satisfied by two interesting economic coalitions, output cartels in oligopoly and public goods coalitions.

5.2.1. *Output Cartels in a Linear Cournot Oligopoly*

Consider a Cournot oligopoly with inverse demand $P(X) = A - X$, where X is the industry output. Firm i's cost function is given by cx_i, where x_i is firm i's output and c is the common constant marginal cost with $A > c$. Suppose that the cartel structure is $C = \{n_1, n_2, \dots, n_m\}$ and consider the size-n_i cartel. Without loss of generality, suppose that the first n_i firms belong to this cartel. The members of this cartel choose their output to maximize their joint profit $\sum_{j=1}^{n_i}[A - c - X]x_j = [A - c - X] \sum_{j=1}^{n_i} x_j$. Under constant marginal (and average) cost, a big cartel does not enjoy any *strategic* advantage over a small cartel, since one plant is as good as several plants. As a result, in a given coalition structure, regardless of size, all cartels make the same profit in the unique Cournot equilibrium. Furthermore, only the number of cartels, not their sizes, determines profits. As a result, the per-member partition function for output cartels in the linear Cournot oligopoly is

$$\pi(n_i; C) = \frac{(A-c)^2}{n_i(m+1)^2}, \qquad i = 1, \dots, m. \tag{5.1}$$

It is easy to see that Eq. (5.1) satisfies the four conditions on the partition function with positive externalities:

LEMMA 5.1. *Output cartels in a Cournot oligopoly with the inverse demand function* $P(X) = A - X$ *and the cost function* $c(x) = cx$ *satisfy* (P.1)–(P.4).

I omit the obvious proof of Lemma 5.1. Instead, I discuss the economic idea behind this result. First, consider the merger of cartels. The members of the merging cartels reduce their output in order to internalize the positive externalities which output reduction creates on each other. The other cartels benefit from the merger by free-riding on the merging cartels' output reduction [(P.1)]. Next, recall that cartels earn the same total profit regardless of size in a given cartel structure. Hence, a small cartel earns a higher per-member profit than does a big cartel [(P.2)]. Finally, suppose that a player belonging to a nondegenerate cartel (that is, a cartel with two or more players) leaves his cartel to join a larger or equal-sized cartel. Since this deviation leaves the number of cartels unchanged, the remaining members of the deviator's former cartel each earns a higher profit [(P.3)] and the deviator earns a lower profit [(P.4)].

5.2.2. *Public Goods Coalitions*

Consider the following model of public goods coalitions. Each player is endowed with 1 unit of a private good. At cost $c(x_i)$, agent P_i can provide x_i units of the public good. Let $X = \sum_{i=1}^{N} x_i$ be the total amount of the public good. Each player enjoys the same benefit from consuming the public good, $g(X)$. Assume that $g'(X) > 0$, $g''(X) \le 0$, $c'(x_i) > 0$, $c''(x_i) > 0$, and $2[c''(x_i)]^2 > c'(x_i)c'''(x_i)$. Player P_i's net utility is given by

$g(X) - c(x_i)$.[5] (To be precise, P_i's net utility is $1 + g(X) - c(x_i)$. To save on notation, I subtract 1 from each player's utility.)

Suppose that the coalition structure is $C = \{n_1, n_2, \ldots, n_m\}$ and consider the size-n_i coalition. Without loss of generality, suppose that the first n_i players belong to this coalition. The members of the size-n_i coalition choose their provision of public goods to maximize their joint utility $\sum_{j=1}^{n_i}[g(X) - c(x_j)] = n_i g(X) - \sum_{j=1}^{n_i} c(x_j)$. The first-order condition for an optimal level of public goods provided by a member of the size-n_i coalition is

$$n_i g'(X) - c'(x_j) = 0, \quad \text{for } j = 1, \ldots, n_i. \tag{5.2}$$

Given the strict convexity of the cost function, the optimal solution of the size-n_i coalition is symmetric. Let $x(n_i; C)$ be the *per-member* provision of the public good by the size-n_i coalition and let $X(n_i; C) = n_i x(n_i; C)$ be the total public good provided by the size-n_i coalition. Finally, let $X(C) = \sum_{i=1}^{m} X(n_i; C)$ be the aggregate amount of the public good produced under the coalition structure $C = \{n_1, n_2, \ldots, n_m\}$. Then Eq. (5.2) becomes

$$n_i g'\big(X(C)\big) - c'\big(x(n_i; C)\big) = 0, \quad \text{for } i = 1, \ldots, m. \tag{5.3}$$

In Appendix B, I show that this model of public goods coalitions satisfies the four conditions on the partition function with positive externalities.

LEMMA 5.2. *Public goods coalitions with utility function $g(X)$ and cost function $c(x_i)$ satisfy* (P.1)–(P.3) *if* $g'(X) > 0$, $g''(X) \leq 0$, $c'(x_i) > 0$, $c''(x_i) > 0$, *and* $2[c''(x_i)]^2 > c'(x_i)c'''(x_i)$. *They also satisfy* (P.4) *for* $g(X) = X$ *and* $c(x) = cx^2$, $c > 0$.

Proof. See Appendix B.

The idea behind Lemma 5.2 is simple. Equation (5.3) shows that a member of the size-n_i coalition ($n_i \geq 2$) produces more public good than

[5] This specification is a slight variation on the standard model of public goods coalition (Ray and Vohra (1994)) in which each coalition decides how much to contribute to the provision of the public good which is produced according to an *economy-wide* production function. In the standard model, if a size-k coalition contributes x and others contribute z, total production of the public good is equal to $c^{-1}(x + z)$. In equilibrium, only the largest coalitions make positive contributions and the other coalitions make no contributions. Furthermore, if there is more than one largest coalition, the second-stage equilibrium outcome is not unique: While the total amount of the public good is fixed, the distribution of the contributions among the largest coalitions is indeterminate. (The total amount of public good in the coalition structure is implicitly defined by the first-order condition $kg'(X) - c'(X) = 0$, where k is the size of the largest coalition.) I adopt the current variation in order to avoid this multiplicity of the second-stage equilibria. (The current formulation has a unique second-stage equilibrium outcome for all coalition structures.) This change does not affect the analysis.

the amount which maximizes his *individual* utility given other agents' production of the public good. This result follows from the fact that, in determining the optimal amount of the public good to produce, a member of the size-n_i coalition takes into account the positive externality on members of his coalition. An inspection of Eq. (5.3) further reveals that, in a given coalition structure, a member of a large coalition produces more public good than a member of a small coalition does. Hence, a member of a large coalition enjoys lower net utility than a member of a small coalition does [(P.2)]. If coalitions merge, the merging coalitions increase their total production of the public good, thus benefiting members of other coalitions [(P.1)]. Similarly, if a member of a coalition leaves his coalition to join a larger or equal-sized one, the aggregate amount of the public good increases but the remaining members of the deviator's former coalition reduce their production of the public good. As a result, the remaining members of the deviator's former coalition become better off [(P.3)]. Finally, the deviator becomes worse off [(P.4)] (for example, for $g(X) = X$ and $c(x) = cx^2$, $c > 0$) because he bears (with the existing members of his new coalition) the burden of increasing the total amount of the public good. These results reflect the fundamental *free-riding* problems associated with the formation of public goods coalitions.

5.3. *Equilibrium Coalition Structures with Positive Externalities*

5.3.1. *Open Membership Game*

It is easy to see that the stand-alone stability is a *necessary* condition for a coalition structure to be a Nash equilibrium outcome of the Open Membership game. (Suppose that $C = \{n_1, n_2, \ldots, n_m\}$ is not stand-alone stable: For some $i = 1, \ldots, m$, we have $\pi(n_i; C) < \pi(1; C_i)$, $C_i = C \setminus \{n_i\} \cup \{n_i - 1, 1\}$. C cannot be supported as a pure strategy Nash equilibrium outcome, because a member of the size-n_i coalition can increase his payoff by instead forming a singleton coalition by announcing an address not chosen by other players.) Condition (P.4) further narrows down the set of Nash equilibrium coalition structures. Consider $C = \{n_1, n_2, \ldots, n_m\}$, $n_1 \geq n_2 \geq \cdots \geq n_m$, with $n_1 \geq n_m + 2$. Under (P.4), a member of the size-n_1 coalition becomes better off by leaving his coalition to join one of the smaller coalitions. Hence, such a coalition structure cannot be a Nash equilibrium outcome.

As a result, under (P.4), the only coalition structures which can be Nash equilibrium outcomes are $C = \{n_1, n_2, \ldots, n_m\}$, $n_1 \geq n_2 \geq \cdots \geq n_m$, with $n_1 \leq n_m + 1$. There are N such coalition structures: Ignoring integer constraints, these are $\{N\}$, $\{N/2, N/2\}$, $\{N/3, N/3, N/3\}$, $\{N/4, N/4, N/4, N/4\}, \ldots, \{2, 2, \ldots, 2\}$, $\{2, 2, \ldots, 2, 1, 1\}, \ldots, \{2, 1, 1, \ldots, 1\}$, and

$\{1, 1, \dots, 1\}$. (More precisely, consider $C = \{n_1, n_2, \dots, n_m\}$, $n_1 \geq n_2 \geq \cdots \geq n_m$, with $n_1 \leq n_m + 1$. Let $k = I(N/m)$ and $q = mk - N (\geq 0)$. Then $C = \{n_1, n_2, \dots, n_m\} = \{k, \dots, k, k-1, \dots, k-1\}$, where there are $m - q$ entries of k, and q entries of $(k-1)$. C is a "symmetric" coalition structure given the integer constraint.) Notice that $\{N\}$ is more concentrated than $\{N/2, N/2\}$, which in turn is more concentrated than $\{N/3, N/3, N/3\}$, and so on.

Appendix C shows that, under (P.4), the most concentrated stand-alone stable coalition structure among these N coalition structures is a Nash equilibrium outcome. Since $\{1,1, \dots, 1\}$ is stand-alone stable, under (P.4) there exists a Nash equilibrium coalition structure of the Open Membership game. The following proposition records these results.

PROPOSITION 5.1. *Consider* $C = \{n_1, n_2, \dots, n_m\}$, $n_1 \geq n_2 \geq \cdots \geq n_m$.

(1) *If C is not stand-alone stable, it cannot be a Nash equilibrium coalition structure of the Open Membership game.*

Under $(P.4)$,

(2) *If $n_1 > n_m + 1$, C is not a Nash equilibrium coalition structure of the Open Membership game.*

(3) *Suppose that $n_1 \leq n_m + 1$. Further suppose that C is stand-alone stable but that $C' = \{n'_1, n'_2, \dots, n'_m\}$, $n'_m + 1 \geq n'_1 \geq n'_2 \geq \cdots \geq n'_m$, is not stand-alone stable for $1 \leq m' < m$. Then C is a Nash equilibrium coalition structure of the Open Membership game. Furthermore, C is the most concentrated Nash equilibrium coalition structure of the Open Membership game.*[6]

(4) *Since $\{1, 1, \dots, 1\}$ is stand-alone stable, there exists a Nash equilibrium coalition structure of the Open Membership game.*

5.3.2. *Infinite-Horizon Coalition Unanimity Game*

Unlike the Open Membership game, it is hard to obtain a sharp characterization of the subgame perfect equilibrium coalition structure of the Coalition Unanimity game with positive externalities. But under (P.1)–(P.3), if an additional condition is satisfied, the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game consists of either the grand coalition or two coalitions.

PROPOSITION 5.2. *Suppose that $\{k, N-k\}$, $k \geq N/2$, is stand-alone stable and that $\pi(N-k; \{k, N-k\}) \geq \pi(1; \{N-2, 1, 1\})$. Under* (P.1)–(P.3), *the subgame perfect equilibrium coalition structure of the Coalition Unanimity game is $\{k^u, N - k^u\}$, where $k^u \geq k$.*

[6] For an odd N, if $C = \{2, 2, \dots, 2, 1\}$ is the most concentrated stand-alone stable "symmetric" coalition structure, assume that $\pi(1; C) > \pi(3; C')$, where $C' = C \setminus \{2, 1\} \cup \{3\}$.

Proof. The proof of Proposition 5.2 consists of three steps.

(Step 1) *If P_1 announces $N - k$, then P_{N-k+1} announces k.* To see why, notice that $\pi(k; \{k, N-k\}) \geq \pi(1; \{1, k-1, N-k\}) > \pi(s; \{s, k-s, N-k\}) > \pi(s; \{s, k-s-1, 1, N-k\}) > \cdots > \pi(s; \{s, 1, 1, \dots, 1, N-k\})$ for all $s, 2 \leq s \leq k-1$. The first inequality follows from the stand-alone stability of $\{k, N-k\}$, the second from (P.2) and (P.3), and the rest from a step-by-step application of (P.1). Hence, given P_1's announcement of $N-k$, P_{N-k+1}'s best strategy is to form a grand coalition among the remaining players.

(Step 2) *For P_1, announcing $N-k$ dominates announcing $N-s$, $1 \leq s \leq k-1$.* By announcing $N-k$, P_1 can secure the payoff $\pi(N-k; \{k, N-k\})$. If P_1 announces $N-s, 1 \leq s \leq k-1$, the *best* payoff P_1 can hope to obtain is $\pi(N-s; \{s, N-s\})$ under (P.1). (If the next player does not announce a grand coalition among the remaining players, P_1's payoff becomes smaller than $\pi(N-s; \{s, N-s\})$.) But under (P.2) and (P.3), $\pi(N-k; \{k, N-k\}) > \pi(N-s; \{s, N-s\})$ for $1 \leq s \leq k-1$.

(Step 3) *If P_1 announces $N-r$ in equilibrium, $r > k$, then P_{N-r+1} must announce r.* Suppose not: P_{N-r+1} announces $t, 1 \leq t < r$. Then, under (P.1), the best payoff P_1 can get is $\pi(N-r; \{N-r, t, r-t\})$. Under (P.1)–(P.3), this payoff is smaller than $\pi(1; \{N-2, 1, 1\})$, which in turn is smaller than $\pi(N-k; \{k, N-k\})$ by assumption. Q.E.D.

Unfortunately, the conditions in Proposition 5.2 are quite restrictive. For example, in the case of output cartels in the linear Cournot oligopoly model [Eq. (5.1)], the degenerate cartel structure $\{1, 1, \dots, 1\}$ is the *unique* stand-alone stable coalition structure for $N \geq 3$. Hence, Proposition 5.2 does not apply to output cartels in the linear Cournot oligopoly model.

5.3.3. *Equilibrium Binding Agreements*

It is easy to see that, under (P.1)–(P.3), a stand-alone stable coalition structure is stable under the Equilibrium Binding Agreements rule.

PROPOSITION 5.3. *Under (P.1)–(P.3), a stand-alone stable coalition structure is a stable outcome under the Equilibrium Binding Agreements.*

Proof. Suppose that $C = \{n_1, n_2, \dots, n_m\}$ is stand-alone stable. Let $D_i = C \setminus \{n_i\}$, $i = 1, \dots, m$. Then, $\pi(n_i; \{n_i\} \cup D_i) \geq \pi(1; \{n_i - 1, 1\} \cup D_i) > \pi(k; \{n_i - k, k\} \cup D_i)$, $k = 2, \dots, n_i - 1$, where the first inequality follows from the stand-alone stability of C and the second inequality from (P.2) and (P.3). Hence, the breakoff by k members of the size-n_i coalition is not profitable if the other players do not break up their coalitions in response to the breakoff by the k deviators. However, if the other players do break up their coalitions, then, by (P.1), the deviators end up even worse off. Q.E.D.

5.3.4. *Comparison of Stable Coalition Structures with Positive Externalities*

When coalition formation creates positive externalities, due to free-riding problems, the Open Membership game rarely supports the grand coalition as a Nash equilibrium outcome. Indeed, the equilibrium coalition structure with positive externalities in the Open Membership game is often very fragmented. For example, consider the output cartels in the linear Cournot oligopoly model. As we have seen in Section 5.3.2, the degenerate cartel structure $\{1, 1, \dots, 1\}$ is the unique Nash equilibrium outcome of the Open Membership game for $N \geq 3$.

The Coalition Unanimity game and the Equilibrium Binding Agreements are better than the Open Membership game in overcoming the free-riding problems which arise when coalition formation creates positive externalities on nonmembers. Propositions in the previous subsections show that, for the case of positive externalities, the Coalition Unanimity rule and the Equilibrium Binding Agreements rule support a more concentrated coalition structure as a stable outcome than the Open Membership rule does. This result is exactly the opposite of what happens in the case of negative externalities. However, the grand coalition is typically not a stable outcome under the Coalition Unanimity rule nor Equilibrium Binding Agreements rule.

So far, I have not been able to produce a general result which compares the equilibrium coalition structures for the case of positive externalities under the Coalition Unanimity rule and under the Equilibrium Binding Agreements rule. The main difficulty lies in the precise characterization of stable coalition structures under these two rules. (Propositions 5.2 and 5.3 provide only partial characterizations of equilibrium coalition structures in these two games.) In the remainder of this section, I discuss these difficulties through a simple example of the output cartels in the linear Cournot oligopoly model. This example also illustrates the differences in the endogenous stability properties of the Coalition Unanimity rule and the Equilibrium Binding Agreements rule.

In the case of output cartels in the linear Cournot oligopoly model, Bloch (1996) shows that the unique subgame perfect equilibrium coalition structure of the Coalition Unanimity game is $C^u = \{k^u, 1, 1, \dots, 1\}$, where $k^u = I((2N + 3 - (4N + 5)^{1/2})/2)$ is the size of the "minimum" profitable cartel identified by Salant *et al.* (1983): $\pi(k^u; \{k^u, 1, 1, \dots, 1\}) \geq \pi(1; \{1, 1, \dots, 1\})$, but $\pi(k; \{k, 1, 1, \dots, 1\}) < \pi(1; \{1, 1, \dots, 1\})$ for all k, $2 \leq k < k^u$. In equilibrium, the first $N - k^u$ players announce 1 and the next player announces k^u. Since $k^u < N$ for $N \geq 6$, the grand coalition is not the equilibrium outcome of the Coalition Unanimity game for $N \geq 6$.

Ray and Vohra (1994) show that the stability of the grand coalition under the Equilibrium Binding Agreements exhibits a "cycling" pattern: The grand coalition is stable for $N = 2$, not stable for $3 \leq N \leq 8$, and stable again for $N = 9$.

This example may suggest that one might be able to obtain a ranking of the stable coalition structures with positive externalities under the Coalition Unanimity rule and the Equilibrium Binding Agreements rule, as in the case of negative externalities. Unfortunately, the answer is negative in the case of cartel formation with a linear demand function. For $N = 6$, $\{5, 1\}$ is the unique subgame-perfect equilibrium coalition structure of the Coalition Unanimity game. $\{5, 1\}$ is more concentrated than $\{3, 2, 1\}$, which is the most concentrated stable coalition structure under the Equilibrium Binding Agreements rule. But for $N = 9$, $k^u = 8$ and, hence, $\{8, 1\}$ is the unique equilibrium coalition structure of the Coalition Unanimity game. But $\{9\}$ is stable under the Equilibrium Binding Agreements rule, as shown by Ray and Vohra (1994).

6. CONCLUDING REMARKS

In this paper, I have examined equilibrium coalition structures when coalition formation creates externalities on nonmembers. I have captured these externalities across coalitions through the partition function which assigns a value to a coalition as a function of the entire coalition structure. There are two main contributions of this paper. First, I have shown that many economic models of coalition formation create either positive externalities (output cartels or public goods coalitions) or negative externalities (research joint ventures or customs unions) on nonmembers. The per-member partition function derived from these economic games of coalition formation satisfies further interesting properties. These properties of the per-member partition function serve as important input in studying the endogenous stability property of different rules of coalition formation.

The second contribution of this paper is the characterization of stable coalition structures under these conditions on the partition function. I have paid particular attention to the study of the stability of the grand coalition. The main finding that emerges from this inquiry is that the Open Membership rule, which stipulates that a coalition admit new members on a nondiscriminatory basis, supports the grand coalition as an equilibrium outcome for the case of negative externalities. But coalition formation rules which allow for exclusivity in membership, such as the Coalition

Unanimity rule or the Equilibrium Binding Agreements rule, typically do not support the grand coalition as a stable outcome.

In contrast, for the case of positive externalities, the grand coalition is usually not an equilibrium outcome under all three rules of coalition formation examined in this paper. This results from the pervasive free-riding problems which arise when coalition formation generates positive externalities on nonmembers.

I conclude this paper with some remarks on future research. First, Propositions 5.2 and 5.3 provide only partial characterization of stable coalition structures with positive externalities under the Coalition Unanimity rule and under the Equilibrium Binding Agreements rule. A more complete characterization and comparison of stable coalition structures with positive externalities under these two rules await further research.

Second, I have assumed ex ante symmetric players in this paper. This symmetry assumption is common to the recent literature on coalition formation with externalities, such as Bloch (1995, 1996) and, to some extent, Ray and Vohra (1994). When players are not symmetric, it is no longer possible to identify a coalition by its size, which is a major simplifying assumption of this paper. As hard as the analysis may be, heterogeneity of players raises the interesting and important issue of the composition of coalitions: Do coalitions in a stable coalition structure (assuming that one exists) consist of similar players or dissimilar players or both? One way to begin the analysis in this direction might be to assume just two types of players and see if equilibrium coalitions consist of the same types or of different types.

APPENDIX A

Proof of Lemma 4.1. It is easy to see that (N.1) holds, because $\mu'(\omega) < 0$. More precisely, suppose that the size-n_i and the size-n_j coalitions merge. (The merger of more than two coalitions is analogous.) Since $n_i \mu(n_i) + n_j \mu(n_j) > (n_i + n_j)\mu(n_i + n_j)$, a member of the size-$n_k$ coalition, $k \neq i \neq j$, earns a lower profit as a result of the merger of the other two coalitions. To see why (N.2) holds, consider the merger of the size-n_i and the size-n_j coalitions and suppose that $n_i \geq n_j$. (The merger of more than two coalitions is analogous.) This merger is profitable to a member of the size-n_j coalition if and only if $(n_i + n_j)\mu(n_i + n_j) - (N + 1)\mu(n_i + n_j) > n_i \mu(n_i) + n_j \mu(n_j) - (N + 1)\mu(n_j)$, or $(N + 1 - n_j)[\mu(n_j) - \mu(n_i + n_j)] > n_i[\mu(n_i) - \mu(n_i + n_j)]$, which holds because $\mu(n_j) \geq \mu(n_i) > \mu(n_i + n_j)$ and $N \geq n_i + n_j$. Finally, to see why (N.3) holds, suppose that a member of the size-n_j coalition leaves it to join a size-n_i coalition, $n_i \geq n_j$. The deviator is better off if and only if $(n_i + 1)\mu$

$(n_i + 1) + (n_j - 1)\mu(n_j - 1) - (N + 1)\mu(n_i + 1) > n_i \mu(n_i) + n_j \mu(n_j) - (N + 1)\mu(n_j)$, or $(N + 1 - nj)[\mu(n_j) - \mu(n_i + 1)] > n_i[\mu(n_i) - \mu(n_i + 1)] - (n_j - 1)[\mu(n_j - 1) - \mu(n_i + 1)]$, which holds because $\mu(n_j - 1) > \mu(n_j) \geq \mu(n_i) > \mu(n_i + 1)$ and $N \geq n_i + n_j$. Q.E.D.

Derivation of Eq. (4.5) *and Proof of Lemma* 4.2. From (4.2), country i's inverse demand function for the non-numeraire good is given by

$$P_i = a - Q_i. \tag{A.1}$$

Given the specific tariff τ_{ij}, country j chooses its exports to country i in order to maximize its export profit:

$$\operatorname*{Max}_{q_{ij}} \pi^{ij} = \left[P_i - c - \tau_{ij}\right] q_{ij}. \tag{A.2}$$

Country j's first-order condition in country i is given by

$$\frac{\partial \pi^{ij}}{\partial q_{ij}} = \left[P_i - c - \tau_{ij}\right] - q_{ij} = 0. \tag{A.3}$$

Solving (A.3) simultaneously for $j = 1, \ldots, N$ yields country j's Cournot equilibrium output in country i:

$$q_{ij} = \frac{(a - c) - (N + 1)\tau_{ij} + \sum_{k \neq i} \tau_{ik}}{N + 1}. \tag{A.4}$$

Substituting (A.3) and (A.4) into (A.2) yields country j's Cournot equilibrium profit in country i:

$$\pi^{ij} = q_{ij}^2. \tag{A.5}$$

Country i's tariff revenue is given by

$$TR_i = \sum_{j \neq i} \tau_{ij} q_{ij}. \tag{A.6}$$

Without loss of generality, suppose that countries $1, 2, \ldots, n_i$ belong to the size-n_i customs union and consider country 1. It solves

$$\operatorname*{Max}_{\{\tau_{1j}\}_{j=n_i+1}^{N}} \sum_{k=1}^{n_i} W^k, \tag{A.7}$$

where $\tau_{1j} = 0$ for $j = 1, \dots, n_i$. Using (4.2)–(4.4) and (A.1)–(A.6), it is straightforward to show that the above maximization problem has a unique solution:

$$\tau(n_i; C) = \frac{1}{(N+1) + (n_i+1)(2n_i+1)}, \qquad i = 1, \dots, m. \quad \text{(A.8)}$$

Substituting (A.8) into (4.4) yields

$$W(n_i; C) = \frac{1}{2} - q_o(n_i) + \sum_{\substack{j=1,\\ j \neq i}}^{m} n_j \big[q_o(n_j)\big]^2, \quad \text{(A.9)}$$

where

$$q_o(n_j) = \frac{1}{(N+1) + (n_j+1)(2n_j+1)} \quad \text{(A.10)}$$

is a nonmember country's export volume to a member country of the size-n_j customs union. By (A.5), $[q_o(n_j)]^2$ is a nonmember country's export profit to a member country of the size-n_j customs union. Notice that $q_o(n_j)$ is a *decreasing* function of n_j. Hence, if customs unions merge, a *nonmember* country's export profits to the merging countries decrease [(N.1)]. The merger of customs unions benefits the members of the smallest customs union involved in the merger. For example, suppose that the size-n_i and the size-n_j customs unions merge, $n_i \geq n_j$. (The merger of more than two customs unions is analogous.) This merger benefits the (former) members of the size-n_j customs union [(N.2)] if and only if $q_o(n_j) - q_o(n_i + n_j) > n_i[q_o(n_i)]^2$, which in turn holds if and only if

$$\frac{n_i\big[2n_i + 2n_j + 1 + 2(n_j+1)\big]}{\big[(N+1) + (n_j+1)(2n_j+1)\big]\big[(N+1) + (n_i+n_j+1)(2n_i+2n_j+1)\big]} > \frac{n_i}{\big[(N+1) + (n_i+1)(2n_i+1)\big]^2}. \quad \text{(A.11)}$$

A tedious derivation shows that (A.11) holds. Finally, a member of the size-n_j customs union becomes better off by leaving its union to join the size-n_i union, $n_i \geq n_j$ [(N.3)] if and only if $q_o(n_j) - q_o(n_i + 1) > n_i[q_o(n_i)]^2 - (n_j - 1)[q_o(n_j - 1)]^2$. Since $q_o(n_j)$ is a decreasing function of

n_j, this last inequality holds if $q_o(n_j) - q_o(n_i + 1) > (n_i - n_j + 1)[q_o(n_i)]^2$, which in turn holds if and only if

$$\frac{\left[n_i - n_j + 1\right]\left[2(n_j + 1) + 2(n_i + 1) + 1\right]}{\left[(N+1) + (n_j + 1)(2n_j + 1)\right]\left[(N+1) + (n_i + 2)(2n_i + 3)\right]} > \frac{\left[n_i - n_j + 1\right]}{\left[(N+1) + (n_i + 1)(2n_i + 1)\right]^2}. \tag{A.12}$$

A straightforward derivation shows that (A.12) holds. Q.E.D.

EXAMPLE 4.1.

$$\underset{50}{\{5\}},\quad \underset{51,35}{\{4,1\}},\quad \underset{52,39}{\{3,2\}},\quad \underset{53,36,36}{\{3,1,1\}},\quad \underset{48,48,37}{\{2,2,1\}},\quad \underset{58,38,38,38}{\{2,1,1,1\}},\quad \text{and}\quad \underset{42,42,42,42,42}{\{1,1,1,1,1\}}.$$

In the unique subgame perfect equilibrium path of the Coalition Unanimity game, P_1 announces 3 followed by P_4's announcement of 2, and P_1 earns $\pi(3;\{3,2\})$. To see why, consider the other four alternatives. If P_1 announces 5, then $\{5\}$ forms and P_1 earns $\pi(5;\{5\})$. If P_1 announces 4, then $\{4,1\}$ forms and P_1 earns $\pi(4;\{4,1\})$. If P_1 announces 2, then P_3 chooses 2 and P_1 earns $\pi(2;\{2,2,1\})$. Finally, if P_1 announces 1, then P_2 chooses 3 and P_1 earns $\pi(1;\{3,1,1\})$. Announcing 3 is the equilibrium strategy for P_1, because $\pi(3;\{3,2\}) > \pi(5;\{5\})$, $\pi(4;\{4,1\})$, $\pi(2;\{2,2,1\})$ and $\pi(1;\{3,1,1\})$.

Under the Equilibrium Binding Agreements, $\{4,1\}$, $\{3,2\}$, $\{2,2,1\}$, $\{2,1,1,1\}$ and $\{1,1,1,1,1\}$ are stable. First, $\{2,1,1,1\}$ and $\{2,2,1\}$ are stable by Proposition 4.3. $\{3,1,1\}$ is *not* stable, since two members of the size-3 coalition can profitably deviate to $\{2,1,1,1\}$, which is stable. $\{3,2\}$ is stable, because members of the size-3 coalition do not find breakup (either to $\{2,2,1\}$ or to $\{2,1,1,1\}$) profitable and members of the size-2 coalition do not find breakup to $\{3,1,1\}$ profitable. (Since $\{3,1,1\}$ is not stable, the breakup of the size-2 coalition in $\{3,2\}$ leads to $\{2,1,1,1\}$. A member of the size-2 coalition in $\{3,2\}$ earns 39 before the deviation and 38 after the deviation.)

$\{4,1\}$ is stable, because the departure of two members from the size-4 coalition leads to $\{2,2,1\}$, which is not profitable for the deviators. The departure of three members from the size-4 coalition leads to $\{2,1,1,1\}$, because $\{3,1,1\}$ is not stable. But one of the leading perpetrators, the player who is left as a singleton coalition after the other two perpetrators deviate again to form the size-2 coalition in $\{2,1,1,1\}$, earns a lower payoff. The departure of one member from the size-4 coalition similarly leads to $\{2, 1, 1, 1\}$. Hence, the leading perpetrator is worse off in this deviation.

Finally, $\{5\}$ is not stable, because three members can profitably deviate to form the size-3 coalition in $\{3, 2\}$, which is stable.

APPENDIX B

Proof of Lemma 5.2. The proof of Lemma 5.2 consists of eight steps, which are presented in the following lemmas.

LEMMA B.1. *For* $C = \{n_1, n_2, \ldots, n_m\}$, $x(n_i; C) > x(n_j; C)$ *if and only if* $n_i > n_j$.

Proof. In a "cross-section" differentiation, one holds the coalition structure (and hence the aggregate amount of public good $X(C)$) fixed and examines how the increase in n_i changes the per-member amount of the size-n_i coalition. A "cross-section" differentiation of Eq. (5.3) with respect to n_i yields

$$\frac{d^{cs}x(n_i)}{dn_i} = \frac{g'(X(C))}{c''(x(n_i))} > 0. \tag{B.1}$$

Q.E.D.

Lemma B.1 leads to (P.2):

LEMMA B.2. $\pi(n_i; C) < \pi(n_j; C)$ *if and only if* $n_i > n_j$.

Proof. Suppose that $n_i > n_j$. $\pi(n_i; C) = g(X(C)) - c(x(n_i; C)) < g(X(C)) - c(x(n_j; C)) = \pi(n_j; C)$, where the inequality follows from $x(n_i; C) > x(n_j; C)$ and $c'(x) > 0$. Q.E.D.

Suppose that a member of the size-n_j coalition leaves his coalition to join the size-n_i coalition, $n_i \geq n_j$. What is the effect of this change in the coalition structure on the total amount of the public good produced? Since we are comparing two coalition structures, we are comparing two equilibria in the (second-stage) public good provision game. Comparison of two equilibria is difficult except for special functions with closed-form solutions. The following differential technique overcomes this difficulty by finding sufficient conditions on the utility and cost functions under which we can unambiguously sign the effect of change in the coalition structure on the equilibrium amount of the public good.

Start with a total differentiation of (5.3) which yields

$$g'(X(C))\,dn_i + n_i g''(X(C))\,dX(C) - c''(x(n_i;C))\,dx(n_i;C) = 0. \tag{B.2}$$

A total differentiation of $X(n_i;C) = n_i x(n_i;C)$ yields

$$dX(n_i;C) = x(n_i;C)\,dn_i + n_i\,dx(n_i;C). \tag{B.3}$$

Substituting (B.3) into (B.2) and rearranging,

$$dX(n_i;C) = -\lambda(n_i)\,dX(C) + \delta(n_i)\,dn_i, \tag{B.4}$$

where

$$\begin{aligned} \lambda(n_i) &\equiv -\frac{n_i^2 g''(X(C))}{c''(x(n_i;C))} \geq 0 \qquad \text{and} \\ \delta(n_i) &\equiv \frac{n_i g'(X(C))}{c''(x(n_i;C))} + x(n_i;C) > 0 \end{aligned} \tag{B.5}$$

Summing (B.4) over $i = 1, \dots, m$ and rearranging yield the effect of infinitesimal changes in the coalition structure on the aggregate amount of public good produced,

$$dX(C) = \frac{\sum_{i=1}^m \delta(n_i)\,dn_i}{1 + \Lambda}, \tag{B.6}$$

where $\Lambda \equiv \sum_{i=1}^m \lambda(n_i) \geq 0$.

Now consider an infinitesimal change in the coalition structure in which dn members leave the size-n_j coalition to join the size-n_i coalition. In a vector notation, it can be written as $d\mathbf{n} \equiv (0, \dots, 0, dn, 0, \dots, 0, -dn, 0, \dots, 0)$, where d_n appears in the ith entry and $-dn$ appears in the jth entry. The effect of this change on the equilibrium aggregate public good is

$$\frac{dX(C)}{d\mathbf{n}} = \frac{\delta(n_i) - \delta(n_j)}{1 + \Lambda}. \tag{B.7}$$

Suppose that one member of the size-n_j coalition leaves his coalition to join the size-n_i coalition. By integrating (B.7) from 0 to 1, we can obtain the effect of this change in the coalition structure on the total public good produced. If $\delta(n_i) > \delta(n_j)$ for $n_i > n_j$, then the total public good produced increases when dn members leave the size-n_j coalition to join the size-n_i coalition. I show that $\delta(n_i) > \delta(n_j)$ for $n_i > n_j$ under weak conditions on the cost function identified in the text as

LEMMA B.3. *$\delta(n_i) > d(n_j)$ if $2[c''(x)]^2 > c'(x)c'''(x)$.*

Proof. Holding the coalition structure (and hence $X(C)$) constant, a cross-section differentiation of $\delta(n_i)$ yields

$$\frac{d^{cs}\delta(n_i)}{dn_i} = \frac{1}{[c''(x)]^2}\left\{g'(X)c''(x) + \left([c''(x)]^2 - n_i g'(X)c'''(x)\right)\frac{d^{cs}x(n_i)}{dn_i}\right\}$$

$$\left(\frac{d^{cs}x(n_i)}{dn_i} = \frac{g'(X(C))}{c''(x(n_i))} \quad \text{by (B.1)}\right)$$

$$= \frac{g'(X)}{[c''(x)]^3}\left\{2[c''(x)]^2 - n_i g'(X)c'''(x)\right\}$$

$$\left(n_i g'(X(C)) = c'(x(n_i; C)) \quad \text{by Eq. (5.3)}\right)$$

$$= \frac{g'(X)}{[c''(x)]^3}\left\{2[c''(x)]^2 - c'(x)c'''(x)\right\} > 0$$

$$\text{if and only if } 2[c''(x)]^2 > c'(x)c'''(x).$$

Q.E.D.

The condition $2[c''(x)]^2 > c'(x)c'''(x)$ is quite weak. For example, this condition is satisfied by the constant-elasticity cost functions $c(x) = cx^{\gamma}$, $c > 0$ and $\gamma > 1$, because $2[c''(x)]^2 - c'(x)c'''(x) = \gamma(\gamma - 1)x^{2\gamma-2} > 0$ for $\gamma > 1$.

It follows from Lemma B.3 that the total amount of public good increases when one member of the size-n_j coalition leaves his coalition to join the size-n_i coalition, $n_i \geq n_j$.

LEMMA B.4. *$X(C) < X(C')$, where $C' = C \setminus \{n_i, n_j\} \cup \{n_i - 1, n_j + 1\}$ and $n_i \geq n_j$.*

Eq. (B.7) and Lemma B.3 show that the size-n_i and size-n_j coalitions jointly increase their production of the public good when one member of the size-n_j coalition leaves his coalition to join the size-n_i coalition,

$n_i \geq n_j$. In response, other coalitions reduce their production of the public good. From (B.4) and (B.7),

$$\frac{dX(n_k;C)}{d\mathbf{n}} = -\lambda(n_k)\frac{dX(C)}{d\mathbf{n}} \leq 0. \tag{B.8}$$

The following lemma records this result.

LEMMA B.5. *$x(n_k;C) \leq x(n_k;C')$, where $C' = C \setminus \{n_i, n_j\} \cup \{n_i - 1, n_j + 1\}$, $n_i \geq n_j$ and $k \neq i \neq j$.*

Lemma B.4 and B.5 imply that (P.1) holds in this model of public goods coalitions: If the coalition structure becomes coarser, indeed, more concentrated (which can be decomposed into finite steps of moving one member at a time from a coalition to a larger or equal-sized one), then members of the coalitions not affected by the change become strictly better off.

LEMMA B.6. *$\pi(n_k;C) < \pi(n_k;C')$, where $\{n_k\} \subset C, C'$ and $C' \setminus \{n_k\}$ is more concentrated than $C \setminus \{n_k\}$.*

Proof. Consider $C' = C \setminus \{n_i, n_j\} \cup \{n_i - 1, n_j + 1\}$ and $n_i \geq n_j$. (The general case can be decomposed into a sequence of moving one member at a time from a coalition to a larger or equal sized one.) $\pi(n_k;C') = g(X(C')) - c(x(n_k;C')) \geq g(X(C') + x(n_k;C) - x(n_k;C')) - c(x(n_k;C)) > g(X(C)) - c(x(n_k;C)) = \pi(n_k;C)$. The first inequality holds, because (1) $g(X) - c(x)$ is strictly concave with respect to x; and (2) $x(n_k;C) \geq x(n_k;C') \geq x'(n_k;C')$, where $x'(n_k;C')$ is the *individual* best response amount of the public good for a member of the size-n_k coalition in the coalition structure C'. The second inequality holds, because $X(C') > X(C)$, $x(n_k;C') \leq x(n_k;C)$, and $g'(X) > 0$. Q.E.D.

From Eqs. (B.3), (B.4), and (B.7),

$$\frac{dx(n_j;C)}{d\mathbf{n}} = -\frac{n_j g''(X(C))}{c''\big(x(n_j;C)\big)}\frac{dX(C)}{d\mathbf{n}} - \frac{g'(X(C))}{c''\big(x(n_j;C)\big)} < 0. \tag{B.9}$$

Hence, the *remaining* members of the formerly size-n_j coalition reduce their production of the public good. As a result, (P.3) holds: When one member of the size-n_j coalition leaves it to join the size-n_i coalition, $n_i \geq n_j \geq 2$, the *remaining* members of the (formerly) size-n_j coalition become better off.

LEMMA B.7. *$\pi(n_j; C) < \pi(n_j - 1; C')$ for $C' = C \setminus \{n_i, n_j\} \cup \{n_i + 1, n_j - 1\}$, $n_i \geq n_j \geq 2$.*

Proof. $\pi(n_j - 1; C') = g(X(C')) - c(x(n_j - 1; C')) \geq g(X(C') + x(n_j; C) - x(n_j - 1; C')) - c(x(n_j; C)) > g(X(C)) - c(x(n_j; C)) = \pi(n_j; C)$. The first inequality follows from (1) the strict concavity of $g(X) - c(x)$ with respect to x and (2) $x(n_j; C) > x(n_j - 1; C') \geq x'(n_j - 1; C')$, where $x'(n_j - 1; C')$ is the individual best response of a member of the size-$(n_j - 1)$ coalition. The second inequality holds, because $X(C') > X(C)$, $x(n_j - 1; C') < x(n_j; C)$, and $g'(X) > 0$. Q.E.D.

Finally, the following result shows that (P.4) holds for $g(X) = X$ and $c(x) = cx^2$, $c > 0$.

LEMMA B.8. *Suppose that $g(X) = X$ and $c(x) = cx^2$, $c > 0$. We have $\pi(n_j; C) > \pi(n_i + 1; C')$ for $C' = C \setminus \{n_i, n_j\} \cup \{n_i + 1, n_j - 1\}$, $n_i \geq n_j \geq 2$.*

Proof. The first-order condition for the size-n_i coalition Eq. (5.3) becomes $n_i - 2cx(n_i; C) = 0$. Hence, $x(n_i; C) = n_i/2c$, $X(n_i; C) = n_i^2/2c$, $X(C) = (1/2c)\sum_{j=1}^{m} n_i^2$, and $\pi(n_i; C) = (1/4c)\{2\sum_{j=1}^{m} n_j^2 - n_i^2\}$. A simple derivation shows that $\pi(n_i + 1; C') < \pi(n_j; C)$ if and only if $n_i + n_j > 3$. Hence, (P.4) holds. Q.E.D.

Proof of Proposition 5.1, (3). Since $n_1 \leq n_m + 1$, we can write $C = \{k, \ldots, k, k - 1, \ldots, k - 1\}$, where $k = I(N/m)$. Since C is stand-alone stable, no player gains by forming a singleton coalition by changing his address to one not chosen by the other players. There are three types of deviations we need to consider. First, a member of the size-k coalition can join another size-k coalition. If $k \geq 2$, then by (P.4), the deviation is not profitable. If $k = 1$, then $C = \{1, 1, \ldots, 1\}$. By assumption, $C' = \{2, 1, \ldots, 1\}$ is not stand-alone stable: $\pi(2; C^1) < \pi(1; C)$. Second, a member of the size-$(k - 1)$ coalition joins another size-$(k - 1)$ coalition, $k \geq 2$. If $k \geq 3$, by (P.4), the deviation is not profitable. If $k = 2$, then $C = \{2, \ldots, 2, 1, \ldots, 1\}$ and the new coalition structure is $C' = C \setminus \{1, 1\} \cup \{2\}$. By assumption, C' is not stand-alone stable: $\pi(2; C') < \pi(1; C)$. Third, a member of the size-$(k - 1)$ coalition joins the size-k coalition, $k \geq 2$. If $k \geq 3$, by (P.4) the deviation is not profitable. If $k = 2$, then $C = \{2, \ldots, 2, 1, \ldots, 1\}$ and the new coalition structure is $C'' = C \setminus \{2, 1\} \cup \{3\}$. There are two cases. If C'' contains singleton coalitions, $\pi(3; C'') < \pi(2; C') < \pi(1; C)$ by (P.4) and by the assumption that C' is not stand-alone stable. If C'' does not contain singleton coalitions, then $C = \{2, \ldots, 2, 1\}$. By assumption, we have $\pi(3; C'') < \pi(1; C)$. Q.E.D.

REFERENCES

Aumann, R. J., and Dreze, J. (1974). ''Cooperative Games with Coalition Structures,'' *Int. J. Game Theory* **3**, 217–237.

Aumann, R. J., and Myerson, R. B. (1988). ''Endogenous Formation of Links Between Players and of Coalitions: An Application of the Shapley Value,'' in *The Shapley Value: Essays in Lloyd S. Shapley* (A. E. Roth, Ed.), Cambridge: Cambridge Univ. Press.

Bloch, F. (1995). ''Endogenous Structures of Association in Oligopolies,'' *Rand J. Econ.* **26**, 537–556.

Bloch, F. (1996). ''Sequential Formation of Coalitions with Fixed Payoff Division and Externalities,'' *Games Econ. Behav.* **14**, 90–123.

Chwe, M. S. (1994). ''Farsighted Coalitional Stability,'' *J. Econ. Theory* **63**, 299–325.

Economides, N. S. (1986). ''Non-Cooperative Equilibrium Coalition Structures,'' Columbia University Discussion Paper Series, No. 273.

Greenberg, J. (1995). ''Coalition Structures,'' in *Handbook of Game Theory with Economic Applications* Vol. 2 (R. J. Aumann and S. Hart, Eds.), Amsterdam: North-Holland.

Greenberg, J., and Weber, S. (1993). ''Stable Coalition Structures with a Unidimensional Set of Alternatives,'' *J. Econ. Theory* **60**, 62–82.

Hart, S., and Kurz, M. (1983). ''Endogenous Formation of Coalitions,'' *Econometrica* **51**, 1047–1064.

Kamien, M. I., and Zang, I. (1990). ''The Limits of Monopolization through Acquisition,'' *Quart. J. Econ.* **105**, 465–499.

Ray, D., and Vohra, R. (1994). ''Equilibrium Binding Agreements,'' unpublished manuscript, Brown University.

Ray, D., and Vohra, R. (1995). ''Binding Agreements and Coalitional Bargaining,'' unpublished manuscript, Brown University.

Salant, S. W., Switzer, S., and Reynolds, R. J. (1983). ''Losses from Horizontal Mergers: The Effects of an Exogenous Change in Industry Structure on Cournot–Nash Equilibrium,'' *Quart. J. Econ.* **93**, 185–199.

Shenoy, P. (1979). ''On Coalition Formation: A Game Theoretical Approach,'' *Int. J. Game Theory* **8**, 133–164.

Yi, S.-S. (1996a). ''Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism Is Good,'' *J. Internat. Econ.* **41**, 151–175.

Yi, S.-S. (1996b). ''Endogenous Formation of Research Coalitions with Complementary Assets,'' Dartmouth College Department of Economics WP No. 95-11, revised.

Yi, S.-S., and Shin, H. (1995). ''Endogenous Formation of Coalitions in Oligopoly,'' Dartmouth College Department of Economics WP No. 95-2.