Voluntariness and the Coase Theorem*

Tore Ellingsen[†] and Elena Paltseva[‡]

16 February 2011

Abstract

The paper investigates the possibility, first suggested by Avinash Dixit and Mancur Olson (2000), that the Coase Theorem may be inconsistent with natural notions of voluntary participation. Retaining Dixit and Olson's definition of voluntariness, we consider bargaining over the provision of club goods. Although property rights are well defined and there are no transaction costs, we provide examples in which the unique equilibrium outcome is inefficient under a variety of bargaining protocols. The reason is that some agents may profit from not participating at the club good provision stage, but instead negotiate access ex post.

---

*Financial support from the Torsten and Ragnar Söderberg Foundation (Ellingsen) and the Wallander and Hedelius Foundation (Paltseva) is gratefully acknowledged. Thanks to Avinash Dixit, Drew Fudenberg, Bård Harstad, Oliver Hart, Erik Lindqvist, Volker Nocke, Larry Samuelson, and Robert Östling for helpful discussions. We are responsible for remaining errors.

[†]Address: Department of Economics, Stockholm School of Economics, Box 6501, S—11383 Stockholm, Sweden. E-mail: gte@hhs.se

[‡]Address: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, 1353 København K, Denmark. E-mail: elena.paltseva@econ.ku.dk

1 Introduction

A modern statement of the Coase Theorem is that voluntary bargaining suffices to attain efficient outcomes if (i) property rights are well defined and perfectly enforced, (ii) parties have complete information, and (iii) negotiation is costless.[1] The theorem, which is based on arguments by Coase (1960), is rarely stated formally and has never been proven in full generality, yet is usually held to be true.[2] However, Dixit and Olson (2000) argue with the help of nonexcludable public goods examples that if the voluntariness requirement is taken seriously, the efficient outcome is only one of many possibilities, and not a particularly likely one. The basic idea is that a free rider will always obtain higher utility than do public good providers.

Defenders of the Coase Theorem may object that the nonexcludable public goods example is inadmissible, because it violates the condition that property rights are well defined. In particular, owners do not have exclusion rights. Another objection is that Dixit and Olson's game has efficient equilibria, and it is somewhat unclear why players are unable to coordinate on these outcomes.

In this paper we take Dixit and Olson's approach a small but crucial step further by offering a related set of excludable public goods (i.e., club goods) examples in which *all* equilibrium outcomes are inefficient under the Coasean assumptions. By focusing on club goods, we respect the property rights prerequisite. Since our game does not have efficient equilibria, we eliminate the coordination problem. Thus, under Dixit and Olson's definition of voluntariness, the Coase Theorem is false.

To fix ideas, consider the following example. Two families live on an island without

[1]The original formulations of the Coase Theorem were less restrictive. The complete information prerequisite was introduced after Myerson and Satterthwaite (1983) showed that incomplete information can lead to unavoidable efficiency losses.

[2]Coase's claim was baptised as theorem by George Stigler. For examples of its widespread acceptance among leading microeconomists, see for example Mas-Colell, Whinston and Green (1995, page 357) and Hart (1995, page 32).

electricity.[3] Both households would like to have an electricity cable to the island. Suppose cables come in one size but in two qualities. Each family values the high quality cable at 24 and the low quality cable at 20. The price of the low quality cable is 25 and the price of the high quality cable is 32.

The jointly optimal outcome for the two families is to buy the high quality cable, because the total net benefit of $2 \cdot 24 - 32$ is greater than the total net benefit of $2 \cdot 20 - 25$. Supposing that the families have equal bargaining power, each family gets a net surplus of 8 under the efficient agreement. But the efficient agreement may not come about. Suppose one of the families, say family A, irrevocably refuses to contribute anything to the purchase of the cable, relying instead on family B to purchase the cable by themselves – both families foreseeing that nothing prevents them from engaging in a new round of negotiations regarding access once the cable is in place. Sticking to the previous bargaining assumption, so that ex post negotiations also lead to an equal split of the available net surplus of 20 (the value of access to family A), it is straightforward to compute the outcome under this scenario. If family B buys the high quality cable, family B's surplus becomes $24-32+24/2 = 4$. If instead family B buys the low quality cable, its surplus is $20 - 25 + 20/2 = 5$. Hence, it prefers to invest in a low quality cable, thus failing to maximize joint surplus. It now only remains to check that family A is wise to refuse an initial agreement. By refusing, family A's net payoff becomes $20/2 = 10$, which is greater than the 8 that it would get under the efficient initial agreement.

Effectively, the situation boils down to a game of Chicken. Each family has the choice of participating at the investment stage, strategy P, and not participating, strategy N. The associated payoffs are depicted in Figure 1.

[3] In the Stockholm archipelago, there are many islands like this - with two summer houses and no electricity. One of the authors used to rent one such house. Recently, the owner of the other house wanted to have electricity, but the author's landlord did not want to chip in. The final outcome is eagerly awaited.

	P	N
P	$8, 8$	$5, 10$
N	$10, 5$	$0, 0$

Figure 1: The participation decision

The game has two pure strategy equilibria, (P,N) and (N,P), and one mixed strategy equilibrium, but only the non-equilibrium outcome (P,P) is efficient.

While the example is stylized, we think that the situation is both generic and economically important. Much private infrastructure, such as telecommunication networks or ATM networks, are essentially club goods, and there are many examples in which access is negotiated after the infrastructure is in place, sometimes by firms that previously refrained from participating at the investment stage.

Does our analysis disprove the Coase Theorem? We think one might as well ask: Is the Dixit-Olson voluntariness definition the most reasonable? The essential properties of their definition are: (a) It is possible to commit not to hear or respond to others' contract offers for as long as one may wish. (b) Such a non-participation commitment precludes making an own pre-emptive contract offer. We think that Property (a) ought to be uncontroversial. Property (b) is more debatable. In Section 4, we briefly discuss the consequences of dropping it.

Before articulating the model more generally, let us relate our analysis first to the classical informal arguments in favor of the Coase Theorem and then to the more abstract and formal work on coalition formation games.

Coase's (1960) celebrated analysis was based on a set of numerical examples concerning externalities between two agents. All of them considered negative externalities, so free riding was never an issue. The subsequent literature soon came to treat positive externalities analogously with negative externalities; see Cheung (1973) and Johnson

(1973). A typical illustration was the relationship between a bee keeper and the owner of an apple orchard. The presence of bees is beneficial for the apple farmer, and the presence of apple trees is beneficial for the bee keeper. According to Coase's logic, the two parties ought to be able to negotiate their way to an efficient size of both operations. These examples differ fundamentally from ours, because neither the apple farmer nor the bee keeper has any incentive to abstain from negotiations. In our club good examples, abstention is caused by the substitutability between investments. In the example of the bee keeper and the apple farmer, the two parties' investments are not substitutes, but complements. Thus, the previous informal arguments are not illogical, but incomplete: They fail to address a relevant class of situations, that of positive externalities from substitutable investments.

A large literature on coalition formation games studies the issue of whether multiplayer bargaining under complete information always admits efficient outcomes. As it turns out, the answer depends on the formulation of the bargaining game as well as the choice of solution concept. Authors taking a cooperative game approach noted that the Core may fail to exist, and some took this as proof that the Coase Theorem fails (Aivazian and Callen, 1981). The claim is invalid for the simple reason that a non-prediction is different from a failure prediction; for a version of this counter-argument, see Coase (1981). Using other solution concepts, a rich subsequent literature has constructed coalition formation games that sometimes fail to have efficient equilibria; see for example Chatterjee et al (1993), Seidmann and Winter (1998), and Gomes and Jehiel (2005). Recent work by Bloch and Gomes (2006) and Hyndman and Ray (2007) explains how many of these inefficiency results are caused by restrictive assumptions concerning agents' ability to write binding contracts, ability to renegotiate such contracts, or by restrictive concepts of equilibrium – notably Markov perfect equilibria.

Bloch and Gomes (2006) is perhaps most closely related to the present paper, addressing the consequences of exit possibilities when there are externalities between coalitions. A crucial difference is that Bloch and Gomes, like other work in this vein, assume that there is always at least one contract proposal before any player is allowed to exit the negotiation. This assumption effectively violates Dixit and Olson's voluntariness requirement (Property (a)) and immediately implies that all bilateral negotiations yield efficient outcomes. On the other hand, Bloch and Gomes assume that a player who exits cannot contract with remaining players. This feature is in line with Dixit and Olson (Property (b)), and we believe that this aspect of voluntariness is responsible both for the inefficiency that we report and for the inefficiency that arises in Bloch and Gomes' multi-player games. However, the two approaches are sufficiently different that we cannot offer a clean translation between them.

Finally, our analysis could inform the literature on incomplete contracts and property rights, initiated by Grossman and Hart (1986). There, inefficient outcomes are due to agents' inability to write complete contracts. We show that, in the case of substitute investments, similar contractual incompleteness and inefficiency may arise even if the agents are able to costlessly write complete contracts. Temporarily refraining from contracting may simply be a better strategy for an individual agent.

The paper is organized as follows. In Section 2, we consider a simple parametric n-player simultaneous move club–formation game. We show that in equilibrium the club will typically not comprise all n players, because some players are better off negotiating access to the club good ex post. As a result, there is underprovision of the club good. In Section 3, we consider a sequential game in which each player has exactly one opportunity to join the club. (Appendix 2 demonstrates that the results are robust to the addition of more bargaining rounds.) Section 4 briefly discusses whether there might be other reasonable ways of specifying the contracting game such that efficiency is always assured. That is, is it likely that a Coase Theorem can be proved?

2 The simultaneous membership model

A group of identical players are to decide on production of and access to a club good. The game has three stages. At stage 1, all players independently and simultaneously decide whether or not to join a club. At stage 2, the club members jointly decide on their contributions to the club good, and the club good is produced.[4] The contribution of each club member is fully observable and club members can write binding and enforceable agreements among themselves. Outsiders are unable to contract with club members at this stage. At stage 3, the club negotiates with outsiders the terms under which the latter are allowed access to the club good. Let N denote the set of all players and M the set of members in the club, with n and m denoting the number of elements in the respective sets. We assume that n is finite and greater than 1. The key to all our inefficiency results is that some player has an incentive to stay outside the club in any equilibrium, and that the club's investment incentive is therefore too small. We focus on pure strategy equilibria. While the model also has equilibria in which players use mixed strategies when making their entry decision at stage 2, these equilibria are less efficient than the pure strategy equilibria.

Each player is endowed with a fixed budget, b. The budget can be spent on private goods or club goods. There is only one type of each good. Private goods have a fixed quality and can be bought in any quantity at a constant price. The club good can only be produced in a fixed quantity, normalized to 1;[5] however the quality of the club good is variable and depends on the sum of the players' contributions. Let e_i denote the contribution of player i to the club good.

[4]Like Dixit and Olson we assume that a player who does not join the club is committed not to contribute to the club good investment. Indeed, attaining this commitment is the point of not joining.

[5]For example, the club good may require a specific and unique site. The restriction to one club good simplifies our analysis by allowing us to disregard investment by outsiders – investment that might in principle be used to strengthen their bargaining position in ex post access negotiations.

The quality of the club good is assumed to be fixed after stage 2. Thus, it depends only on total stage 2 contributions. Since our primary aim is to provide a conterexample to the Coase Theorem, and not to establish an alternative theorem, we adopt the simple production function

$$q\left(e_1,...,e_m\right)=\begin{cases} 2\left(\sum_{i\in M} e_i-\varepsilon\right)^{1/2} & \text{if } \sum_{i\in M} e_i-\varepsilon\geq 0; \\ 0 & \text{otherwise.}\end{cases}$$

Note that ε can be interpreted as the fixed cost associated with producing the lowest possible quality. For most of the paper we assume for convenience that $\varepsilon < 1$, which is small enough never to threaten production. The utility of any player j is assumed to be linear in the quantity of private goods and in the quality of the club good. In order to avoid corner solutions, we assume that the budget b is large enough to ensure that each player's marginal dollar is spent on private goods. Thus, player j's utility can be written

$$U(e_j,q)=q\left(e_1,...,e_m\right)-e_j.$$

Note that the perfect substitutability of the players' efforts implies that it is unimportant for overall efficiency which set of players actually produces the public good.

While the club good is non-rivalrous in consumption, providers have enforceable property rights to it. More precisely: each member of the provision club has access, outsiders can be denied access, and the club can sell access rights. (These properties define club goods.)

Suppose for a moment, counterfactually, that the club members were forced to deny access to any outsiders. The game would then only contain the first two stages – club formation and negotiation among club members. Suppose that at date 0 exactly m players were to enter the club. Since contributions are contractible among club members, contributions would maximize their joint utility. If positive, the contribution of member i must thus solve the problem

$$\max_{e_j} 2m \left(\sum_{i \in M} e_i - \varepsilon \right)^{1/2} - e_j,$$

which implies that

$$\sum e_i = m^2 + \varepsilon.$$

Incentives to enter the club are affected by the club's cost-sharing arrangement. As will become clear, asymmetric cost sharing makes it harder to attain full efficiency. Thus, for our purposes it suffices to consider symmetric cost sharing. Under symmetric sharing, the utility of each club member j is given by

$$U(e_j, q) = m - \varepsilon / m.$$

By staying outside the club, player j would instead get the reference utility of 0. Since $\varepsilon < 1 \leq m$, we see that j is better off in the club than outside it. As the argument holds for any $m \geq 1$, in a subgame-perfect equilibrium of this game everyone wants to join the club and the production of the club good is socially efficient.

However, the premise that the club will deny access to any outsiders is untenable. *Ex post* there is always an incentive for the club to sell access to any willing outsider. Therefore, we return to the original timing, allowing the club to bargain with the outsiders at stage 3. Suppose that ex post bargaining gives the outsider a share β of the gains from trade. That is, if at stage 2 the club produces quality Q, an outsider receives a net utility of βQ and the club receives a payment of $(1 - \beta)Q$. Assuming symmetric sharing of payoffs among club members, the game has the following equilibrium outcome.

Proposition 1 *The equilibrium size of the club is*

$$m^* = \min \left\{ n, \;\; \text{integer} \left(\left(\frac{n^2 (1-\beta)^2 + \beta^2 - \varepsilon}{\beta^2} \right)^{1/2} \right) + 1 \right\}.$$

Proof. First, derive the quality of club good produced at date 1, conditional on m players having decided to join the club at date 0. Taking into account the outcome of bargaining with $n-m$ outsiders at date 2, the club of m members maximizes

$$2m\left(\sum_{i\in M} e_i-\varepsilon\right)^{1/2}-\sum_{i\in M} e_i+(1-\beta)2(n-m)\left(\sum_{i\in M} e_i-\varepsilon\right)^{1/2} \tag{1}$$

subject to the non-negativity constraint. Since $\varepsilon<1$ the solution is interior and can be written

$$\sum e_i=(\beta m+(1-\beta)n)^2+\varepsilon. \tag{2}$$

Thus the quality produced by a club of size m is

$$q_m=2\left(\beta m+(1-\beta)n\right). \tag{3}$$

From equation (2) it follows that the payoff of staying outside of club of size m is equal to

$$U_m^{Out}=2\beta\left(\beta m+(1-\beta)n\right). \tag{4}$$

Given symmetric payoffs, the payoff of a member of a club of size m is

$$U_m^{In}=\frac{(\beta m+(1-\beta)n)^2-\varepsilon}{m}. \tag{5}$$

Now turn to the participation decision at date 0. A club of size m constitutes an equilibrium if (i) no outsider wants to join, i.e., if

$$U_m^{Out}\geq U_{m+1}^{In}, \tag{6}$$

and (ii) no insider wants to leave, i.e., if

$$U_m^{In}\geq U_{m-1}^{Out}. \tag{7}$$

Let us assume for expositional simplicity that when the outsider is indifferent, she joins, while if the insider is indifferent, she stays in the club. Substituting (4) and (5) into (6) and (7) and solving the resulting system yields the optimal club size

$$m^*=\min\left[n,\text{integer}\left(\frac{\left(n^2\left(1-\beta\right)^2+\beta^2-\varepsilon\right)^{1/2}}{\beta}\right)+1\right],$$

(see Appendix 1 for details). The quality of the club good produced in equilibrium q^* is given by formula (3) for $m = m^*$. ■

We are now ready to characterize some of the properties of the equilibrium.

Proposition 2 *The equilibrium size of the club m^* and the equilibrium quality provision q^* are weakly increasing in the size of the society n.*

Proof. See Appendix 1. ■

The intuition is quite simple. For a given size of the club and a given quality of the club good, a larger population means that the number of outsiders increases. This raises the payoff of insiders while keeping the outsiders' payoff unchanged. Therefore outsiders have additional incentive to join the club. At the same time, for a given club size the insiders' marginal return from increasing provision and selling access goes up, therefore amplifying the insiders' incentive to invest. If the additional player joins the club, the latter effect can only be strengthened, because the cost per member goes down.

Proposition 3 *The equilibrium size of the club m^* and the equilibrium level of club good provision q^* are weakly decreasing in the bargaining power of the outsiders β.*

Proof. See Appendix 1. ■

For a given size of the club and quality level of the club good, higher bargaining power of outsiders increases their payoff and lowers the payoff of the insiders. Therefore, the marginal club member may now be better off outside the club. For a given club size the marginal return to quality investment goes down due to the increased bargaining power of the outsiders. If any member leaves, the club invests even less, because the leaving player ends up contributing less toward financing the marginal quality unit.

Proposition 4 *In the simultaneous club formation game there exists a threshold $\widehat{\beta} < 1$ such that for any $\beta \in \left[\widehat{\beta}, 1\right]$ the unique subgame perfect equilibrium outcome is a club size $m^* < n$. In this equilibrium quality provision is inefficiently low.*

Proof. Suppose that at date 0 player n expects all other $n-1$ players to join the club. From formula (2) it follows that a club of the size $n-1$ produces the quality

$$\beta(n-1)+(1-\beta)n=n-\beta.$$

Thus if player n decides to stay outside, her payoff is

$$U_{n-1}^{Out}=2\beta(n-\beta).$$

If instead she decides to join the club of size $n-1$ (so that it becomes the club of size n), she gets the payoff

$$U_n^{In}=n-\frac{\varepsilon}{n}.$$

Thus she prefers to stay outside if and only if

$$2\beta(n-\beta)>n-\frac{\varepsilon}{n}. \tag{8}$$

As $\beta\in[0,1]$, the expression $2\beta(n-\beta)$ reaches its maximum at $\beta_{\max}=1$. For this $\beta_{\max}$ the inequality (8) becomes

$$2(n-1)>n-\frac{\varepsilon}{n},$$

or, equivalently

$$n>2-\frac{\varepsilon}{n}. \tag{9}$$

As long as $\varepsilon>0$, inequality (9) holds for any $n\geq 2$. By continuity, inequality (8) is satisfied for some segment $\left[\widehat{\beta},1\right]$. This, in turn, implies that for any $\beta\in\left[\widehat{\beta},1\right]$ the n-th player chooses not to join the club. ■

Again, the intuition is simple. As long as outsiders have enough bargaining power, it is more tempting to stay outside than to join the club at date 0. If n and/or ε is large the critical bargaining power (the threshold $\widehat{\beta}$) can in fact be quite low for free riding incentives to preclude the efficient outcome.

While the results were derived under the assumption that club members divide utility equally, inefficiency can only get worse if asymmetric splitting is assumed. The reason is that players who expect to carry more than the average burden have a stronger incentive to stay outside. Asymmetric splitting can thus only be helpful in combination with sequential moves at the club formation stage.

3 The single-round sequential membership model

To demonstrate that the inefficiency is not due to simultaneous moves in the coalition formation game, we now assume instead that the players enter the club formation negotiations in a sequential order. At each stage, if no club was created before, player i has a choice of organizing the club or staying outside. If the club is already created, the organizer of the club bargains over entry with each new potential member i. If there is agreement, player i joins the club; otherwise she stays outside. Finally, once all players have had their opportunity to join, the club produces the club good whereafter the club organizer bargains with the outsiders over access.

In the entry negotiations, the outside option of a potential entrant is the utility she anticipates having from subsequent access negotiations. In the access negotiation, on the other hand, the outside option is 0. In an alternating offers bargaining model, Binmore, Shaked and Sutton (1988) shows that, when a player has a relatively good outside option it enters as a constraint; the player gets exactly the outside option, but no more. We therefore assume that each new member at the coalition formation stage obtains the expected utility from access negotiations. It is as if the organizer makes a take-it-or-leave-it offer to the entrant. (This assumption also maximizes the incentive for founding the club, and thus gives efficiency its best shot.) Finally, in the access negotiation the entrant's outside option is not binding. Instead, as before, we assume that the entrant gets a fraction β of the surplus.

Although the analysis becomes somewhat technical, the inefficient outcome rests on two elementary arguments. First, early movers may choose not to set up a club, relying instead on later movers to do so. Second, for every player that the club's founder signs up, it becomes more expensive to sign up subsequent players, as the value of their outside option increases.

Lemma 1 *If player 1 organizes the club, she signs up player n in any subgame-perfect equilibrium.*

Proof. For each branch of the game, let Θ denote the "history" of that branch - that is, the collection of actions of players $1, ..., n$ along the branch. Let $P(i|\Theta)$ denote the payoff of player i conditional on Θ, and let $V(m)$ denote the aggregate payoff of a club of m members.

Consider a game tree node X where player 1 makes an offer to player n. Let Θ_{n-1} denote the "history" of this node - that is, the actions of players $i \in \{1, 2, 3, ..., n-1\}$. In other words, Θ_{n-1} summarizes the offers from player 1 to players $i \in \{2, 3, ..., n-1\}$ and whether they accepted or rejected these offers. Suppose that $m \leq n-2$ players accepted the offer along this game branch before it reached the node X, and let $M(\Theta_{n-1})$ denote the set of players who accepted the offer.

If player 1 wants to sign player n in, she must offer player n at least her outside option. Among the continuations of the game beyond node X at which player n joins the club, the only candidate for SPNE is the one bringing n the minimum "joining" payoff. A history Θ_{n-1} followed by this continuation is denoted Θ_n^{In}. All the continuations of the game beyond node X at which player n does not join, are payoff-equivalent for all participants, so we denote them all Θ_n^{Out}. The minimum payoff to sign in player n is thus determined by the equation

$$P(n|\Theta_n^{In}) = P(n|\Theta_n^{Out}). \tag{10}$$

Note that the payments to the club members $i \in M(\Theta_{n-1})$ are already set when the game reaches node X and do not depend on whether player n accept or rejects the offer. That is,

$$P(i|\Theta_n^{In}) = P(i|\Theta_n^{Out}). \tag{11}$$

Consider the difference between the maximum payoff of player 1 in case she signs in player n and her maximum payoff in case she keeps player n outside:

$$\begin{aligned} P(1|\Theta_n^{In}) - P(1|\Theta_n^{Out}) &= \left(V(m+1) - \sum_{i \in M(\Theta_{n-1})} P(i|\Theta_n^{In}) - P(n|\Theta_n^{In}) \right) \\ &\quad - \left(V(m) - \sum_{i \in M(\Theta_{n-1})} P(i|\Theta_n^{Out}) \right). \end{aligned} \tag{12}$$

Applying formulas (10), (11) to equality (12) transforms it into

$$P(1|\Theta_n^{In}) - P(1|\Theta_n^{Out}) = V(m+1) - V(m) - P(n|\Theta_n^{Out}). \tag{13}$$

The payoff to the coalition of m members is

$$V(m) = (\beta m + (1-\beta)n)^2 - \varepsilon, \tag{14}$$

and the payoff of being outside such a coalition is

$$P_m^{Out} = 2\beta\left(\beta m + (1-\beta)n\right). \tag{15}$$

Using (14) and (15) to substitute for the terms in equation (13) yields

$$\begin{aligned} P(1|\Theta_n^{In}) - P(1|\Theta_n^{Out}) &= \left(\beta^2(2m+1) + 2n\beta(1-\beta) - 2\beta\left(\beta m + (1-\beta)n\right)\right) \\ &= \beta^2 > 0. \end{aligned}$$

That is, no matter how large the current club is, player 1 is always better off by signing in player n, than she is by keeping n outside. ■

The intuition is that when the club's founder bargains with the last potential member, the founder is already committed and cannot escape. Since player n is willing to join if offered the utility emanating from staying outside, and since both parties can gain from a higher quality level than that which will result if n does stay outside, the founder will sign in player n.

However, a similar logic does not hold when player 1 negotiates with other players. The reason is that each entering club member is associated with higher quality, and thus improves outside options for subsequent potential entrants, making it more costly for the founder to sign them into the club.

Lemma 2 *If player* 1 *organizes the club, she does not sign up player* $n-1$ *in any subgame-perfect equilibrium.*

Proof. Consider a game tree node X' where player 1 makes an offer to player $n-1$. Similarly to above, denote by Θ_{n-2} the "history" of node X'. Denote the set of players who accepted the offer along this game branch before it reached node X' by $M'(\Theta_{n-2})$, with m' being its cardinality.

Again, to sign player $n-1$ in, player 1 should offer player $n-1$ at least her outside option. According to Lemma 1, in any SPNE player n is always signed in. If player $n-1$ stays outside, in a SPNE player 1 signs in player n for the minimum payoff, so that the final club size is $m'+1$. The respective game tree branch is denoted by Θ_{n-1}^{Out}. If player $n-1$ accepts the offer, the subgame-perfect continuation implies a club of $m'+2$ members. Again, among all the continuation games with both players $n-1$ and n signed player 1 prefers the one where both $n-1$ and n join for the minimum acceptable payoff. This game branch is denoted by Θ_{n-1}^{In}. So player $n-1$ joins if she is offered at least

$$P(n-1|\Theta_{n-1}^{In}) = P(n-1|\Theta_{n-1}^{Out}). \tag{16}$$

Again, like in (11) the payments to the club members joining the club prior to the node X' being reached, $i \in M'(\Theta_{n-2})$ do not depend on the continuation of the game, so

$$P(i|\Theta_{n-1}^{In}) = P(i|\Theta_{n-1}^{Out}). \tag{17}$$

Compare the payoff of player 1 along Θ^{In}_{n-1} and Θ^{Out}_{n-1} respectively:

$$\begin{aligned} P(1|\Theta^{In}_{n-1}) - P(1|\Theta^{Out}_{n-1}) &= \left(V(m'+2) - \sum_{i \in M'(\Theta_{n-2}) \cup \{n\}} P(i|\Theta^{In}_{n-1}) - P(n-1|\Theta^{In}_{n-1}) \right) \\ &\quad - \left(V(m'+1) - \sum_{i \in M'(\Theta_{n-2}) \cup \{n\}} P(i|\Theta^{Out}_{n-1}) \right), \end{aligned}$$

or equivalently,

$$\begin{aligned} P(1|\Theta^{In}_{n-1}) - P(1|\Theta^{Out}_{n-1}) &= \left(V(m'+2) - P(n|\Theta^{In}_{n-1}) - P(n-1|\Theta^{In}_{n-1}) \right) \\ &\quad - \left(V(m'+1) - P(n|\Theta^{Out}_{n-1}) \right), \end{aligned} \tag{18}$$

where the last equality follows from (17). Substituting (16) into (18) we get

$$\begin{aligned} P(1|\Theta^{In}_{n-1}) - P(1|\Theta^{Out}_{n-1}) &= V(m'+2) - P(n|\Theta^{In}_{n-1}) \\ &\quad - \left(V(m'+1) - P(n|\Theta^{Out}_{n-1}) \right) - P(n-1|\Theta^{Out}_{n-1}). \end{aligned}$$

There are $m'+2$ club members along the game tree branch Θ^{In}_{n-1}, so $P(n|\Theta^{In}_{n-1})$ can be found from (10) setting $m = m'+1$. Similarly, $m'+1$ players join the club along the path Θ^{Out}_{n-1} and (10) for $m = m'$ determines $P(n|\Theta^{Out}_{n-1})$. Finally, $P(n-1|\Theta^{Out}_{n-1})$ is the outsider's payoff when the club consists of $m'+1$ members. Using these observations as well as relations (14) and (15) we obtain

$$P(1|\Theta^{In}_{n-1}) - P(1|\Theta^{Out}_{n-1}) = -\beta^2 < 0. \tag{19}$$

■

We make two observations before proceeding to the next step of the argument. First, for a given parameters set, the payoff of player n is fully determined by the number of players who join the club before she gets to choose, and not affected by their identity. Second, Lemma 2 holds independently of the number of players who already agreed to join the club prior to the node X' being reached. These two observations together with backward induction allow us to extend the statement of Lemma 2 to all remaining players $i = n-2, n-3, ..., 2$.

Lemma 3 *If player* 1 *organizes the club, she does not sign up player* $i = 2, ..., n-2$ *in any subgame-perfect equilibrium.*

Proof. As above, consider a game tree node X'' where player 1 makes an offer to player $n-2$. Lemmas 1 and 2 describe the SPNE continuation of the subgame starting in node X''. As the player $n-1$ does not join the club in any SPNE, she does not impact on the allocation of the payoffs within the club. Thus, we can repeat the argument in the proof of Lemma 2, replacing $n-1$ by $n-2$. As a result, we get an analogy of property (19): Player 1 is better off by not signing in player $n-2$, because

$$P(1|\Theta_{n-2}^{In}) - P(1|\Theta_{n-2}^{Out}) = -\beta^2 < 0. \tag{20}$$

The result for $i = n-3, ..., 2$ is obtained in exactly the same way. ■

An immediate consequence of last three Lemmas is that whenever player 1 organizes the club, the resulting club only consists of two players: player 1 and player n.

Corollary 1 *As long as outsiders have some bargaining power, the grand coalition never forms.*

One might suspect that failure to form the grand coalition entails inefficiency. Is that correct, and if so how severe is the problem? To tackle the question, we proceed to derive the pure-strategy SPNE of the game. So far, we have merely supposed that player 1 chooses to organize the club, but would she do so in equilibrium? Clearly, in making the decision player 1 compares the payoff of being a club organizer (and signing player n only) to the payoff of staying outside and allowing subsequent players to form the club. Let us therefore first derive the subgame perfect equilibrium of the continuation game that follows if player 1 decides not to organize the club.

Again, once the player opts out of the club she does not influence the subsequent distribution of payoff in the club. Thus we can extend the logic of Lemmas 1, 2 and 3 to the case where player k gets to organize the club.

Lemma 4 *If players $i = 1, ..., k-1$ choose not to participate in the production of the public good, and player $k < n$ organizes the club, she only signs in player n in any SPNE.*

It remains to investigate whether anybody but player n will be in the club, and if so who. Consider first the node where player $n-1$ gets her chance to organize a club. If she chooses to organize the club, she becomes the residual claimant in a club of size 2, as players $i = 1, ..., n-2$ have already chosen to be outside, and player n is going to be signed in by Lemma (4). If instead player $n-1$ chooses not to organize the club, she receives the payoff of an outsider of a club of size $m = 1$, as player n obviously would organize the club in this case (the alternative is to have no good produced at all). So player $n-1$ will choose not to organize the club if and only if

$$V(2) - P_1^{Out} \leq P_1^{Out},$$

where P_1^{Out} is the minimum payment required to attract player n. Using (14) and (15) and simplifying, we find that player $n-1$ will pass up the opportunity to organize the club if and only if

$$n^2(1-\beta)^2 - \varepsilon \leq 0. \tag{21}$$

Now move one step backwards – to the node where player $n-2$ gets to organize the club. If she chooses to do so, she becomes the residual claimant of a club of size 2, as players $i = 1, ..., n-3$ already chose to be outside, and only player n will be subsequently signed (by Lemma 4). If she decides to opt out, there are two possible continuation games depending on a parameter values: In the case that (21) holds, the continuation game only has player n joining the club. Thus by opting out player $n-2$ receives a payoff of an outsider of a club of size 1. Note that in this case the choice of player $n-2$ is exactly identical to the choice of player $n-1$, and is also determined by (21). As a result, player $n-2$ chooses to stay outside too. By backward induction the same result holds for players $n-3, ..., 1$. So for these parameter values the SPNE of the game is highly inefficient; the club consist of player n only, and the quality of the club good is merely

$$\underline{q} = 2\left(\beta + (1-\beta)n\right),$$

as given by equation (3).

If instead the parameters induce player $n-1$ to organize the club (so that (21) is violated), by opting out player $n-2$ becomes an outsider of a club of size 2. Thus she opts out if and only if

$$V(2) - P_1^{Out} \leq P_2^{Out}. \tag{22}$$

Using (14) and (15) and simplifying, we conclude that player $n-2$ decides not to organize the club if and only if

$$n^2(1-\beta)^2 - 2\beta^2 - \varepsilon \leq 0. \tag{23}$$

Note that independently of the decision of player $n-2$ the club that follows her organization decision has 2 members: either player $n-2$ and player n, or player $n-1$ and player n. Thus the choice of player $n-3$ in the node where she gets to organize the club is also determined by condition (23). By backward induction, the players $n-4, ..., 1$ face the same choice as player $n-2$, and consequently, make the same decision. So if condition (21) fails but condition (23) holds, it is better to be an outsider of a club of size 2, than to be its organizer. Thus, in the SPNE the club comprises players $n-1$ and n and produces

$$\overline{q} = 2\left(2\beta + (1-\beta)n\right),$$

according to equation (3).

If both (21) and (23) fail, each player $i = n-1, n-2, ..., 1$ prefers being a residual claimant in a club of size 2 to being an outsider of a club of the same size. Thus, in this case the club consists of two members, 1 and n, and again produces the quality $\overline{q}$. This completes our characterization.

Proposition 5 *Let* $\beta \in (0,1)$*. (a) If the outsider's bargaining power is sufficiently high, i.e., if* $n^2(1-\beta)^2-\varepsilon \leq 0$, *the club consists of player* n *only and provides the quality* $\underline{q} = 2(\beta + (1-\beta)n)$*. (b) For intermediate values of* β*, satisfying* $n^2(1-\beta)^2 - \varepsilon > 0 \geq n^2(1-\beta)^2 - 2\beta^2 - \varepsilon$*, the club comprises players* $n-1$ *and* n *and provides the quality* $\overline{q} = 2(2\beta + (1-\beta)n)$*.(c) For lower values of* β*, the club comprises players* 1 *and* n *and provides the quality* $\overline{q}$*.*

The intuition for why the club becomes so small under sequential negotiations is that with each additional member signed up by the founder it becomes more expensive to sign up subsequent members. A striking fact is that the grand coalition fails to form even when the club's outsiders has very little bargaining power (β is close to zero). The result is not as paradoxical as it first appears, because as β goes to zero the equilibrium quality converges monotonically to the optimal level $2n$. That is, even if the equilibrium club size is small, the efficiency level can be large.

In Appendix 2, we demonstrate that the above results generalize naturally to the case of multiple rounds of negotiation.

4 Concluding remarks

Dixit and Olson (2000) asked whether voluntary participation undermines the Coase Theorem. We find that the answer is affirmative under their concept of voluntariness. Private property rights in combination with costless bargaining fail to preclude free-riding.

One objection to our conclusion is that a club might set up a constitution under which subsequent trade with outsiders must occur on terms that extract all the outsiders' surplus. We think that this objection is intriguing, but does not necessarily resolve the problem. If the club can make strategic commitments prior to access negotiations, then the outsiders also should be able to make such commitments. As Ellingsen and Miettinen (2008) show, such bilateral commitment games tend not to have efficient solutions.[6]

Another objection is that it is possible to imagine other definitions of voluntariness. Specifically, assume now that a party that exits from negotiations can leave behind a credible promise. To see why this matters, consider again the introductory example. Suppose that family A, as it exits, leaves behind the promise that it is willing to make a contribution of 1.5 towards any purchase of a high quality cable. If this contract is binding, family B's net surplus becomes $4 + 1.5 = 5.5$ if it buys the high quality cable, which is preferable to 5. Family A's surplus in turn increases from 10 to $24/2 - 1.5 = 10.5$.

Unfortunately, this observation does not establish that the opportunity to make pre-emptive promises ensures the existence of efficient equilibria. Rather it raises two questions: First, what happens if promises or other contract proposals always precede any exit decision? Second, is the latter timing assumption compatible with sensible notions of voluntariness?

Leaving the second question aside, we already have some partial answers to the first question. Jackson and Wilkie (2005) show that when two players can make promises before playing some normal form game, the resulting outcome is not generally efficient.[7]

[6]To be precise, in the case of zero transaction costs and perfect commitment, it follows from Ellingsen and Miettinen's (2008) analysis of bilateral bargaining that iterated elimination of weakly dominated strategies admits only the outcome in which both parties demand the whole surplus – and thus fail to agree.

[7]The analysis of Jackson and Wilkie (2005) is not quite conclusive, however, since they do not prove that the game of promises has an equilibrium - only that any equilibrium, if it exists, will be inefficient.

On the other hand, Ellingsen and Paltseva (2011) argue that optimal outcomes are attainable in Jackson and Wilkie's setting if the players can not only make unilateral promises, but can also propose and sign agreements before playing the normal form game. (To us, allowing agreements seems crucial for evaluating the Coase Theorem.) However, all these results concerning pre-play non-cooperative contracting apply only to normal form games. It is not known whether efficiency is also attainable when players – as in our club goods example – are engaged in an extensive form game with recurring negotiation opportunities along the path of play.

To summarize, Dixit and Olson have proposed a definition of voluntariness under which we have shown that the Coase Theorem fails. Whether there are other sensible definitions of voluntariness under which a Coase Theorem holds remains an open issue.

References

[1] Aivazian, V.A. and Callen, J.L. (1981): The Coase Theorem and the Empty Core, *Journal of Law and Economics* 24, 175-181.

[2] Binmore, K.G., Shaked, A., and Sutton, J. (1988). An Outside Option Experiment, *Quarterly Journal of Economics* 104, 753-770.

[3] Bloch, F. and Gomes, A. (2006): Contracting with Externalities and Outside Options, *Journal of Economic Theory* 127, 172-201.

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

[5] Cheung, S.N.S. (1973): The Fable of the Bees: An Economic Investigation, *Journal of Law and Economics* 16, 11-33.

[6] Coase, R.H. (1960): The Problem of Social Cost, *Journal of Law and Economics* 3, 1-44.

[7] Coase, R.H. (1981): The Coase Theorem and the Empty Core: A Comment, *Journal of Law and Economics* 24, 183-187.

[8] Dixit A.K. and Olson, M. (2000): Does Voluntary Participation Undermine the Coase Theorem? *Journal of Public Economics* 76, 309-335.

[9] Ellingsen, T. and Miettinen, T. (2008). Commitment and Conflict in Bilateral Bargaining, *American Economic Review* 98, 1629-1635.

[10] Ellingsen, T. and Paltseva, E. (2011). Non-Cooperative Contracting, Manuscript.

[11] Gomes, A. and Jehiel, P. (2005): Dynamic Processes of Social and Economic Interactions: On the Persistence of Inefficiencies, *Journal of Political Economy* 113, 626-667.

[12] Grossman, S.J. and Hart, O.D. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, *Journal of Political Economy* 94, 691-719.

[13] Hart, O. (1995): *Firms, Contracts, and Financial Structure,* Oxford: Clarendon Press.

[14] Hyndman, K. and Ray, D. (2007): Coalition Formation with Binding Agreements, *Review of Economic Studies* 74, 1125–1147.

[15] Jackson, M.O. and Wilkie, S. (2005): Endogenous Games and Mechanisms: Side Payments among Players, *Review of Economic Studies* 72, 543-566.

[16] Johnson, D.B. (1973): Meade, Bees, and Externalities, *Journal of Law and Economics* 16, 35-52.

[17] Mas-Colell, A., Whinston, M.D., and Green, J.R. (1995): *Microeconomic Theory,* Oxford: Oxford University Press.

[18] Myerson, R.B. and Satterthwaite, M.A. (1983): Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory* 28, 265-281.

[19] Seidmann, D. and Winter, E. (1998): Gradual Coalition Formation, *Review of Economic Studies* 65, 793-815.

A Appendix 1: Proofs.

A.1 Proof of Proposition 1.

Using (4) and (5) equation (6) takes the form

$$2\beta\left(\beta m+(1-\beta)n\right)>\frac{\left(\beta\left(m+1\right)+(1-\beta)n\right)^{2}-\varepsilon}{m+1}. \tag{24}$$

Solving for m yields the following restriction on the size of the club m:

$$m>\frac{\left(n^{2}\beta^{2}+\beta^{2}-2n^{2}\beta+n^{2}-\varepsilon\right)^{1/2}}{\beta} \tag{25}$$

Similarly, (7) can be rewritten as

$$\frac{\left(\beta m+(1-\beta)n\right)^{2}-\varepsilon}{m}\geq 2\beta\left(\beta\left(m-1\right)+(1-\beta)n\right), \tag{26}$$

which, in turn, is equivalent to

$$m\leq 1+\frac{\left(n^{2}\beta^{2}+\beta^{2}-2n^{2}\beta+n^{2}-\varepsilon\right)^{1/2}}{\beta}. \tag{27}$$

Recognizing that the equilibrium club size m^* is a natural number between 1 and n, we have

$$m^{*}=\min\left[n,\text{integer}\left(\frac{\left(n^{2}\left(1-\beta\right)^{2}+\beta^{2}-\varepsilon\right)^{1/2}}{\beta}\right)+1\right].$$

A.2 Proof of Proposition 2.

Denote

$$\widetilde{m}=\frac{\left(n^{2}\left(1-\beta\right)^{2}+\beta^{2}-\varepsilon\right)^{1/2}}{\beta^{2}},$$

so that

$$m^{*}=\min\left[n,\text{integer}\,\widetilde{m}+1\right].$$

Note that if β is sufficiently close to 0, then $\widetilde{m}\to\infty$ and $m^*=n$. Otherwise,

$$\frac{\partial\widetilde{m}}{\partial n}=\frac{n\left(\frac{1}{\beta}-1\right)^{2}}{\left(n^{2}\left(\frac{1}{\beta}-1\right)^{2}+1-\frac{\varepsilon}{\beta^{2}}\right)^{1/2}}>0,$$

and thus m^* also (weakly) increases in n, the actual increase occurring when $\widetilde{m} \in \mathbb{N}$. According to formula (3)

$$q^* = 2\left(\beta m^* + (1-\beta)n\right),$$

which immediately implies that q^* also increases in n.

A.3 Proof of Proposition 3.

If β is sufficiently close to 0, then $\widetilde{m} \to \infty$ and $m^* = n$, which does not change with β. Otherwise,

$$\frac{\partial \widetilde{m}}{\partial \beta} = -\frac{\left(n^2\left(1-\beta\right)-\varepsilon\right)}{\beta^2\left(n^2\left(1-\beta\right)^2+\beta^2-\varepsilon\right)^{1/2}}.$$

Note that whenever

$$n^2\left(1-\beta\right)-\varepsilon > 0,$$

$\frac{\partial \widetilde{m}}{\partial \beta} < 0$ and thus m^* also (weakly) increases in β. Now consider the case when

$$n^2\left(1-\beta\right)-\varepsilon \leq 0.$$

Then,

$$\widetilde{m} = \frac{\left(n^2\left(1-\beta\right)^2+\beta^2-\varepsilon\right)^{1/2}}{\beta} < \left(\frac{\beta-\varepsilon}{\beta}\right)^{1/2} < 1.$$

Thus for all such $\widetilde{m}$, $m^* = 1$ and does not change with β. We conclude that m^* always (weakly) increases in β.

Rewriting (3) we get

$$q^* = 2\left(\beta m^* + (1-\beta)n\right) = 2\left(n - \beta\left(n - m^*\right)\right). \tag{28}$$

As m^* (weakly) decreases in β, so does the term $-\beta\left(n-m^*\right)$ and thus the entire expression (28).

B Appendix 2: Multi-round membership negotiations. [Not for publication?]

An objection to the single-round membership negotiation model is that it precludes relative losers, notably the last player, from making efficiency-enhancing offers to the relative winners at the club formation stage. What would happen if we added bargaining rounds to the above game? Under some circumstances, efficiency may then be attainable. Suppose in particular that the decision to enter the club is irreversible, whereas the decision not to enter is fully reversible. Then, it can be shown that the unique subgame perfect equilibrium of the game entails full efficiency provided that there are sufficiently many bargaining rounds.

We think that this objection is inadmissible, as it contradicts the idea of voluntariness. By adding rounds to the bargaining game, while insisting that the decision not to enter is only binding over a single negotiation round, the modeller implicitly forces reluctant players to participate in a negotiation that they may prefer to leave. In our view, a more reasonable definition of voluntary bargaining is that the players have the ability to exit the negotiation. In other words, a player ought to have the ability to not make or take offers.

To formalize our argument, we assume that a modified version of the above sequential game is repeated a finite number of times K. Each player has K options to move and within each round the moves are always in the same sequence. As above, at each stage, if no club was created before, player i has a choice between organizing the club or staying outside. Since the game is repeated, we need to make explicit our commitment assumptions. We assume that the decision to organize or to join the club is binding, whereas the decision to stay outside may or may not be binding. If player i who chooses to stay outside at time t receives another offer later in the game, or chooses to organize a club (in case it is not created yet), she is free to do so. At the same time, players cannot be forced to sit at the bargaining table. That is, at her move, in addition to the options of creating the club and staying outside, an uncommitted player i can choose to exit the bargaining process until after date K, by which time the good has to be provided.[8] We refer to the three possible actions for an uncommitted player as "in", "wait", and "exit". Observe that we do not allow players to make conditional commitments to future choices. For example, it is impossible to commit to exit at a future date unless some prerequisite has been fulfilled by that date. Any promise or threat is credible only to the extent that it is actually desirable to carry it out when the time arrives.[9] Moreover, the option to leave the negotiation table may also be available to the club's founder and not only to presently uncommitted players. However, it is easily seen that the founder will not ever want to exit negotiations, either before or after date K. If the founder exits the negotiation before interacting with all players, she loses some potentially valuable members; if instead she exits after date K but before having negotiated with all outsiders, she loses revenues from selling access.

Except for the distinction between waiting and exiting, which could not be made in a single-round framework, everything is as before: The organizer of the club makes unconditional monetary offers to the potential members, and once the club is created (that is, either the final period of the game is over or every player committed to join the club or to exit), the organizer bargains with any remaining outsiders over the terms of access.

As it turns out, the case of $K = 2$ rounds contains virtually all the insights of the

[8]The assumption that there is a specific provision date greatly simplifies the analysis.

[9]In our introductory example, efficiency can be attained by letting family B commit to only take part in provision if A also does so. The problem is that carrying out such a commitment is typically not credible once it is clear that family A will not be providing. (This is not to say that conditional commitment is always impossible. In some cases one party has a strong reputation or can rely on other sources of authority. For a general analysis of which outcomes can be attained under this assumption, see Segal,1999.)

general case. Moreover, the final outcome in the $K = 2$ case is identical to the $K = 1$ case.

Lemma 5 *If player* 1 *chooses to organize the club at her first move, then players* $2, ..., n-1$ *choose to exit in the first round of the game and player* n *joins the club. Thus, the resulting club consist of players* 1 *and* n *only.*

Proof. Consider the choice of player n at her first opportunity to move, that is, at the end of the first round. First, suppose that by this time every player $2, ..., n-1$ already either joined the club or exited. Assume that the number of club members is given by m. Then if player n chooses to exit, nobody joins the club in the next period and thus player n becomes an outsider of a club of size m with the payoff P_m^{Out}. If player n waits, she gets signed in the second round of the game for the payoff of P_m^{Out} and the resulting club size is $m+1$. Thus, as above, player 1 signs in player n for the payoff of P_m^{Out} and the resulting club has the size $m+1$.

Suppose now that a set of players $W \subseteq \{2, ..., n-1\}$ waited in the first round. Let $j = \max W$. Again, assume that the current club has m members. If player n exits, by Lemma 1 player j gets signed up in the second round of the game for the payoff P_m^{Out}. Thus the resulting club consists of $m+1$ members and, being an outsider, player n receives P_{m+1}^{Out}. If player n waits instead, she will get signed up in the second round of the game, receiving $P_m^{Out} < P_{m+1}^{Out}$. (We know from the single-round game that neither player j not any other player who waited will be signed up in the last round.) That is, for player n in round 1 waiting is strictly dominated by exiting. Thus, to sign player n at the end of the first round costs P_{m+1}^{Out}. If player 1 chooses to do so, she continues by signing player j in the second round for the payoff P_{m+1}^{Out}. The resulting club consists of $m+2$ members. Therefore, the difference between the payoffs of player 1 when she signs player n in the first round and when she does not is given by

$$\left[V(m+2) - P_{m+1}^{Out} - P_{m+1}^{Out}\right] - \left[V(m+1) - P_m^{Out}\right] = -\beta^2 < 0. \qquad (29)$$

Thus, in this case player 1 does not sign player n in the first round, and player n chooses to exit. Player j gets signed in the second round and the resulting club again consists of $m+1$ members.

We proceed to the choice of player $n-1$ in the first round of the game. Suppose m players joined the club before player $n-1$ gets her first move. If she exits, we have just shown that in the continuation of the game player 1 signs in one more player for the payoff of P_m^{Out}. The resulting club will consist of $m+1$ members and player $n-1$ receives P_{m+1}^{Out}.

If player $n-1$ waits, then player n exits in the first round and player $n-1$ gets signed in the second round for the payoff $P_m^{Out} < P_{m+1}^{Out}$, so player $n-1$ never chooses this strategy. Thus, to sign in player $n-1$ in the first round player 1 has to pay P_{m+1}^{Out}. If she signs player $n-1$, the club expands to $m+1$. As we shown above, player 1 then proceeds by signing in one more player for the payoff of the outsider of the club of size $m+1$, P_{m+1}^{Out}. Therefore, the net benefit to player 1 from signing player $n-1$, is again given by expression (29). We see that it is too costly for the player 1 to sign player $n-1$ in the first round. Thus, player 1 does not sign in player $n-1$, and player $n-1$ chooses to exit.

By backward induction the same result holds for players $n-2, ..., 2$. To summarize, along the branch where player 1 organizes the club, players $2, ..., n-1$ exit at the first round, and player n joins the club. ■

The result continues to hold when the club is organized by player i, $1 < i < n$.

Lemma 6 *If player i, $1 < i < n$, chooses to organize the club at her first move, the resulting club comprises only player i and player n. Players $1, ..., i-1, i+1, ..., n-1$ exit in the first negotiation round.*

Proof. This result follows from two observations. First, it can be proven exactly along the lines of the Lemma 5, that players $i+1, ..., n-1$ choose to exit.

Suppose that each player $1, ..., i-1$ has chosen to exit too. Then we showed above that player n joins the club, so that the resulting club consists of player i and player n.

Suppose now that a set of players $W' \subseteq \{1, ..., i-1\}$ waited in the first round. Let $k = \max W'$. Then player n exits in the first round, and the second round of the game results in the club consisting of player i and player k. Therefore, player k receives the payoff of an outsider of the club of size 1. But by choosing to exit, she would get the payoff of an outsider of the club of size 2. Thus, in equilibrium player k never prefers waiting to exiting, which contradicts the definition of player k. ■

The situation may potentially be different when player n gets her first chance to organize the club. To analyze the resulting equilibria, we consider three cases depending on the parameter values (n, β, ε). These three cases are motivated by the one-round game of the previous section. Denote by P_2^{Org} the payoff of the organizer of club of size 2,

$$P_2^{Org} = V(2) - P_1^{Out},$$

where P_i^{Out} denotes the payoff of an outside of a club of size i, recall formula (15).

Case 1.

Assume that the parameters n, β and ε satisfy the inequality

$$n^2(1-\beta)^2 - 2\beta^2 - \varepsilon > 0. \tag{30}$$

This case corresponds to the equilibrium of the one-round game when the club consists of players 1 and n. As we have seen above, in this equilibrium the payoff to an outsider of club of size 2 is below the payoff to the organizer of the club of size 2,

$$P_2^{Org} > P_2^{Out}. \tag{31}$$

There are three possibilities: First, assume that every player $1, ..., n-1$ has chosen to exit. Then in equilibrium player n chooses to be the single provider of the public good, and players $1, ..., n-1$ get the payoff of an outsider of club of size 1.

Second, suppose that only one player $h \in \{1, ..., n-1\}$ waited. If player n chooses to organize the club, she signs in player h, creates the club of size 2 and gets P_2^{Org}. If she chooses not to organize, but does not exit, she gets signed in by player h in the second round for the payoff P_1^{Out} (as the alternative would be to be an outsider of club of size 1). If she exits, she also receives P_1^{Out}, which is less than P_2^{Out}. By assumption (31) player n chooses to organize the club in the very first round, if she gets this opportunity.

Third, assume that there are more than one player $1 \leq h_1, ..., h_j < n$ that have waited (and the remaining players exited). The only difference with the previous case appears when player n chooses to exit. In this case by the argument of the previous section the remaining players organize a club of size 2 in the second round and player n receives P_2^{Out}. But due to the assumption (31) this option is dominated by her preferred choice - to organize the club.

Now let us return to the options of the player $n-1$. If she does not organize the club, it follows from the above discussion that her payoff is at most that of an outsider of a club of size 2. If she organizes the club (of 2 members), she receives $P_2^{Org} > P_2^{Out}$ by assumption (31). Thus, player $n-1$ would choose to organize the club in the first round, if ever given this opportunity. By the same logic we see that for this set of parameters the equilibrium is characterized by player 1 organizing the club of size 2 at the very first round, players $2, ..., n-1$ immediately exiting and player n joining the club. That is, for this set of parameters the outcome of the two-rounds game exactly replicates the outcome of the one-round game.

Case 2.

Assume now instead that

$$n^2(1-\beta)^2 - \varepsilon > 0 \geq n^2(1-\beta)^2 - 2\beta^2 - \varepsilon. \tag{32}$$

In this case the payoff of an outsider of a club of size 2 is higher than the payoff of the organizer of the club of size 2, which is in turn higher than the payoff of an outsider of the club of size 1 :

$$P_2^{Out} > P_2^{Org} > P_1^{Out}. \tag{33}$$

In the equilibrium of the corresponding one-stage game, the club consists of players $n-1$ and n.

Again, we consider three possibilities: If every player $1,...,n-1$ has chosen to exit, then in equilibrium player n chooses to be the single provider of the public good. Then players $1,...,n-1$ get the payoff P_1^{Out}.

Suppose now that only one player $h \in \{1,...,n-1\}$ waited. If player n organizes the club, she signs in player h into the club of size 2 and gets the payoff P_2^{Org}. If n chooses to wait, she gets signed by h in the second round and receives the payoff of an outsider of the club of size 1. If n exits, she also receives P_1^{Out}. Also in this case, by assumption (33) player n chooses to organize the club in the very first round if she gets this opportunity.

The difference with the previous case arises when there are more than one players $1 \leq h_1 < ... < h_j < n$ that have waited (while the other players exited). If player n chooses to exit, players h_{j-1} and h_j organize the club of size 2 in the second round. Therefore player n receives P_2^{Out}, which, by assumption (33) exceeds her payoff under the other two options. Thus, player n exits in this case. Now let us move backwards. In round 2, player h_j joins the club and receives no more than the maximum of the payoff of the organizer of club of size 1 (if she is the only one to join) and that of an outsider of the club of size 1 (if she is the last one to join the club of size 2). Suppose that instead of waiting, player h_j exits in the first round before player n gets to move. As there was at least one more person h_1 who has chosen not to organize the club and not to exit, player h_j receives P_2^{Out}. By assumption (33) she thus exits in round 1, which contradicts the definition of h_j. More generally, in equilibrium no player $i = 1,...,n-1$ chooses to wait if any of her predecessors waited.

Now consider the options of the player $n-1$. If everyone exited before, she chooses to organize the club, as this brings her P_2^{Org}, while alternative options bring her P_1^{Out}. If instead there is a player $h_1 < n-1$ who already has chosen to wait, we are back to our discussion of the incentives of the player h_j for $h_j = n-1$. Thus in this case player $n-1$ exits and receives P_2^{Out}.

Let us proceed backwards. If player $n-2$ exits, she gets the payoff P_2^{Out}. If she organizes the club, she signs up player n and receives the payoff P_2^{Org}, which is below the payoff of the exit option. Finally, if she waited (which she would only do if all players $1, ..., n-3$ exited), she gets signed in the club by player n and receives the payoff P_1^{Out}. Thus, player $n-2$ chooses to exit. By backward induction, so do the players $n-3, ..., 1$. Thus, in the resulting equilibrium players $1, ..., n-2$ exit and player $n-1$ organizes the club of size 2 by signing player n already in the first round. Again, this equilibrium outcome exactly replicates that of the single-round game.

Case 3.

Finally, assume that

$$n^2(1-\beta)^2 - \varepsilon \leq 0. \tag{34}$$

In the single-round game, the club in this case consists of player n only, and the payoff of the outsider of club of size 1 exceeds the payoff of the organizer of the club of size 2,i.e.,

$$P_2^{Out} > P_1^{Out} > P_2^{Org}. \tag{35}$$

Thus, no player wants to be the club organizer, as this brings the lowest payoff.

If every player $1, ..., n-1$ has chosen to exit, then in equilibrium player n chooses to be the single provider of the public good. Then players $1, ..., n-1$ get the payoff P_1^{Out}.

If there is any player $h \in \{1, ..., n-1\}$ that has chosen to wait, player n chooses to exit. By assumption (35) she does not want to organize the club. If she stays, the SPNE of the one-stage game will force her to do so in the second round.

Thus, staying outside without exit entails a loss of the first mover advantage – if player $h < n$ has chosen to stay, every player $h+1, ..., n$ immediately exits. On the other hand, an early exit insures the player the payoff of an outsider of the club of size 1. In the respective equilibrium players $1, ..., n-1$ exit in the first round of the game and player n organizes the club of size 1.

So we demonstrated that in the SPNE of the 2-rounds game every player makes a binding decision already by the end of the round 1 and the respective equilibrium outcomes exactly replicate the ones of the one-round game. By backward induction, the same is true for a game with any finite number of rounds K.

Proposition 6 *In the multi-round sequential game with a finite number of rounds, the grand coalition never forms if $n > 2$. If $n = 2$, the grand coalition may or may not form, depending on parameters. In either case the equilibrium size and composition of the club, as well as the club good provision level, replicate the outcomes of the one-stage game.*

In some respects, Case 3 above is perhaps the most important. It identifies parameters such that the grand coalition fails to form even when $n = 2$. Since the scope for varying the extensive form is quite limited in bilateral bargaining games, we are confident that the two-player counterexamples to the Coase Theorem are highly robust. For the sake of concreteness, suppose that $n = 2$, $\beta = 1/2$, and $\varepsilon \in (1, 9/4)$, satisfying the inequality that defines Case 3. (Moreover, even though $\varepsilon > 1$, it is easy to check that the solution remains interior.) For all our three extensive forms, the only equilibrium outcome is for one of the players to provide the club good and invest in quality $q = 3$, instead of the optimum quality $q = 4$; both numbers follow from equation (3). Under the chosen parameters, we believe that this is the unique kind of equilibrium outcome of any bargaining game with voluntary participation: Both players cannot be better off as joint providers than as lone outsiders, and each player prefers being the lone provider to having no provision.