# Transnational public goods: strategies and institutions

## Daniel G. Arce M. [a,*], Todd Sandler [b]

[a] *Robert D. McCallum Professor of Economics and Business, Department of Economics, Rhodes College, 2000 North Parkway, Memphis, TN 38112-1690, USA*
[b] *Robert R. and Katheryn A. Dockson Chair of International Relations and Economics, School of International Relations, University of Southern California, Los Angeles, CA 90089-0043, USA*

## Abstract

Weaker-link and better-shot public goods are prevalent in examples of transnational collective action. Instances include dike building, atmospheric monitoring, cyberspace virus control, deforestation, disease control, and peacekeeping. This paper analyzes essential game-theoretic features of such public goods, which allow correlated strategies to provide Pareto-improving alternatives to the Nash equilibria. Correlation is justified as providing a formal structure for the veil of uncertainty and political leadership. Weaker-link and better-shot public goods differ in terms of the appropriate institutional design. We also consider the consequences of diminishing returns on game forms and institutional prescriptions. © 2001 Elsevier Science B.V. All rights reserved.

## 1. Introduction

There are many cases of transnational public goods and externalities, for which allocative decisions in one country have consequences in another that is not party to the decisions, and no market exists to compensate for the associated costs or

---

* Corresponding author. Tel.: +1-901-843-3121; fax: +1-901-843-3736.
*E-mail address:* darce@rhodes.edu (D.G. Arce M.).

benefits.[1] Examples include curbing global warming, reducing acid rain, controlling plagues and agricultural pests, protecting biodiversity, and curtailing financial market contagion. New technologies not only create novel public bads and externalities (e.g., plutonium, ozone-depleting chlorofluorocarbons), but also allow humankind to monitor the earth and its atmosphere. Such monitoring leads to the discovery of additional externalities.

When public goods and externalities are present, markets are anticipated to result in sub-optimal resource allocation (Olson, 1965; Olson and Zeckhauser, 1966; Cornes and Sandler, 1996). At times, the underlying game form for such public goods and externalities is that of the Prisoner's Dilemma with its Pareto-inferior Nash equilibrium; but, at many other times, the game form is that of Chicken or Coordination with multiple Nash equilibria (Sandler, 1997). The underlying game form is dependent on the manner in which individual contributions or actions determine the overall level of the public good or externality. The relationship between individual contributions and the resulting provision of the public good is known as the *technology of public supply aggregation* (Sandler and Sargent, 1995).

Because independent actions often result in non-optimal Nash equilibria for transnational public goods and externalities, there is a need for institutional arrangements. To foster cooperation, the international community has devised institutions.[2] In some areas (e.g., international shipping, telecommunications, commercial aviation), institutions have evolved in the form of conventions and allocative mechanisms, reflecting the common interest in establishing standards of behavior concerning jurisdictional rights, competitive practices, and accident avoidance (Zacher, 1996). For environmental and security concerns, treaties and alliances have been used with varying degrees of success. Recent environmental treaties (e.g., Helsinki Protocol on sulfur emissions in Europe, the Montreal Protocol on ozone-shield depletion, and the Sofia Protocol on nitrogen oxide emissions in Europe) tend to *codify* Nash behavior and, as such, do not represent much in the way of cooperative gains.[3] For example, the majority of Helsinki Protocol signers had either met or were well on their way to meeting the treaty's mandated 30% reduction of emissions from 1980 levels by the time of adoption (8 July 1985); and, on average, two-thirds of the mandated reduction had been

---

[1] On such global and transnational public goods, see Kaul et al. (1999), Sandler (1997, 1998) and Young (1989, 1998).

[2] In the political science literature, this is often referred to as *regime building* (e.g., Young, 1989, 1991, 1998; Zacher, 1996). In economics, these cooperative linkages are referred to as *institution building* (Arce M., 1997; Schotter, 1981), where the specification of the game results in an institution.

[3] On the Montreal Protocol, see Murdoch and Sandler (1997), and on the Helsinki and Sofia Protocols see Murdoch et al. (1997). A similar message comes from studies on treaty making by Barrett (1994, 1998), who established that treaties which achieve little in the way of a Pareto improvement are the most likely to be agreed upon.

achieved prior to the treaty being framed (Sandler, 1997, p. 127). Even allies in a long-standing alliance, such as NATO, have been shown to abide by Nash behavior (Sandler and Murdoch, 1990).

Treaties, alliances, and other supranational institutions that are intended to provide public goods or correct externalities, thus result in Nash equilibria that are sub-optimal. There is consequently a need to consider more effective institutions that can improve upon these equilibria.

Our primary purpose in this paper is to examine the allocative efficiency of *weaker-link* (*better-shot*) public goods, for which smaller (larger) contributions have the greater marginal impact on provision. We investigate an institutional design that limits transaction costs and requires little loss of autonomy on behalf of the participants, while improving their well-being. The institutional arrangements we consider are self-enforcing in the sense that every country gains from following rules based on *correlating strategies*, which are not the same as mixed strategies and need not be a Nash equilibrium (Arce M., 1997; Aumann, 1974, 1987; Skyrms, 1996). A supranational institution (e.g., an alliance or a treaty) that employs a correlated strategy sends costless signals, not unlike that of preplay communication, which allows the participants to condition their play so as to avoid bad outcomes and improve their expected payoffs over Nash equilibria. Indeed, correlation can improve expected payoffs beyond even the effect of public randomization over Nash equilibria.

Another purpose for this paper is to formalize the *veil of uncertainty* (Brennan and Buchanan, 1985) and *leadership* (Frohlich and Oppenheimer, 1978; Kindleberger, 1981; Young, 1989), which are two prominent concepts associated with transnational institution building. We show that *self-enforcing* institutional arrangements exist for weaker-link public goods that result in Pareto-improving outcomes that *do not* involve any of the Nash equilibria. In general, the appropriate institutional structure results in more cooperation for both the weaker-link and better-shot cases.

## 2. Preliminaries

We shall restrict our analysis to the case of the voluntary provision of a pure public good. Initially, all agents are assumed to be identical, and so game matrices are symmetric.[4] Symmetry is present in many scenarios, where countries collec-

---

[4] With heterogeneous agents, three issues beyond the scope of this paper would have to be addressed. First, there is the concern of strategic misrepresentation of preferences so as to manipulate the participation mechanism. Second, there would be a need for a Rawls' (1971) veil of ignorance, so that agents view themselves a priori as equal when deciding provision rules. Without such a veil, an endless cycling of voting would result from distributional concerns. Third, income effects associated with heterogeneous tastes would have a significant impact on the payoffs.

tively confront a public good concern and frame a treaty or erect a supranational structure. More homogeneous agents are anticipated to have an easier time in forming collectives in terms of framing and signing agreements (Olson, 1965). In the case of protocols (i.e., Helsinki and Sofia) signed under the Long-Range Transboundary Air Pollution Convention (LRTAP), the ratifiers possessed similar supply-side and demand-side characteristics, with respect to the pollutant being controlled (Sandler, 1997). Countries such as the United Kingdom, which differ greatly from other potential ratifiers in terms of spillins and self-pollution, did not sign. A similar observation applies to the initial ratifiers of the Montreal Protocol on ozone-layer depletion, which included the major producer and consumer countries of CFCs.

Public goods can take a variety of forms and can differ from one another in a manner in which individual efforts contribute to the good's overall level (i.e., the *aggregation technology*). In an important contribution, Hirshleifer (1983) extended the standard public good, as defined by the *sum* of contributions, to weakest-link and best-shot technologies.

For weakest-link public goods, the smallest individual effort determines the overall level of the public good. Hirshleifer used the example of a dike built on a circular island, for which the lowest portion of the dike determines the level of safety for all islanders. This aggregation technology results in a coordination game with Nash equilibria that involve matching contributions, because nothing is gained by exceeding the smallest contribution.

For best-shot, the overall level of the public good is determined by the greatest individual contribution; for example, where the greatest research effort achieves the breakthrough for everyone. Best-shot leads to just a single agent providing the public good, since others' efforts are redundant. The Nash equilibria are characterized by free-riding by all but one agent.

The extreme forms and voluntary contribution characteristics that weakest-link and best-shot public goods imply, however, are not commonly observed among the global and transnational collective action problems confronting humankind. This suggests that less strict forms—weaker-link and better-shot public goods— might be more germane (Cornes, 1993; Sandler, 1997).

For weaker-link public goods, the smallest contribution has the largest marginal influence on utility, followed by the second smallest contribution, and so on. The reverse holds for better-shot: the largest contribution has the greatest marginal impact on utility, followed by the second largest effort, and so on. Weaker-link public goods include pest eradication (e.g., insects, crop diseases); as individual effort beyond the least may add to well-being, but progressively, by less, as effort levels exceed the minimal one. For the circular island dike, differences in the island's topography may allow unilaterally larger efforts to have some positive marginal impact. Examples of other weaker-link goods include computer viral protection stemming from prophylactic measures applied by networked users; fiscal and monetary discipline deriving from restraint among monetary union

members; pollutant dispersion arising from containment efforts among countries; and network compatibility coming from common practices among linked agents. Another weaker-link public good is the enforcement of transnational treaties (e.g., efforts to stem the smuggling of endangered species or CFCs). Weaker-link public goods resemble strategic complements.

Better-shot public goods are also present in the global community. Consider the development of a medicine or a vaccine against a disease. For the Salk (in-activated) and Sabin (oral) polio vaccines, the second-best effort generated benefi-cial spillovers. In the same way, efforts below the largest, to monitor the atmosphere, may still provide some marginal gains, although at a smaller level than the best. Better-shot public goods also include instituting proactive policies against transnational terrorism, deploying neutralizing agents against pollutants, verifying treaties, and promoting nuclear weapon nonproliferation. Better-shot public goods are akin to strategic substitutes.

Equilibrium behavior and institutional implications, which distinguish the un-derlying games of weaker-link and better-shot public goods from games associated with previously studied aggregation technologies, can be captured by some specific numerical representations. Generalizing the analysis is problematic be-cause there are competing functional forms for weaker-link and better-shot (Mueller, 1989; Cornes, 1993), and correlated equilibrium is not uniquely defined for continuous strategy spaces. The cases studied illustrate how institutional mechanisms such as leadership, and the veil of uncertainty result in correlated strategies, and consequently have the potential to improve upon the codification of Nash behavior in weaker-link and better-shot environments.

## 2.1. Weaker-link

For weaker-link public goods, a unit contributed has a greater marginal benefit when matched by the other player.

**Definition.** The technology of public supply aggregation is *weaker-link* if:

$$0 \le q_j < q_i \Rightarrow \frac{\Delta U}{\Delta q_i} < \frac{\Delta U}{\Delta q_j} \quad \forall \Delta q_i = \Delta q_j, \tag{1}$$

where $q_i$ represents the contribution of individual '$i$' and $q_j$ is that of individual '$j$'.

Condition (1) indicates that, for identical increases in contributions to the public good, the increase to the smallest contribution has the greatest marginal impact on utility. In addition, unilateral action by one agent can yield a positive (marginal) increase in the public good, even if the other agent is not providing any public good. This cannot occur under weakest-link.

Two numerical examples of a weaker-link technology are shown in Tables 1 and 2. We adopt the following notation: $U(q_i, q_j)$ denotes the utility or benefits for a contribution combination of $q_i$ and $q_j$ prior to deduction of provision costs. Strategic combinations are, henceforth, indicated in parentheses—i.e. (1,0) represents player '$i$' contributing one unit ($q_i = 1$) and player '$j$' contributing zero units to the public good ($q_j = 0$). Each table represents a weaker-link technology, because, if $q_i > q_j$ then utility is increased at a greater rate by increasing $q_j$ (up until the value of $q_i$), as compared to an equal marginal increase in $q_i$.

For example, starting from $U(1,0)$, we see that in either table $U(1,1)$ exceeds $U(2,0)$ even though aggregate contributions are $q_i + q_j = 2$. The difference between the two tables has to do with the benefits of unilateral action and the rate at which marginal utility decreases. In Table 1, additional increments in public good provision increase total utility at a rate that is greater than that for Table 2. For example, when $(q_i, q_j) = (2,1)$, an additional contribution by '$j$' ($q_j = 2$) increases utility by 2 in Table 1, but the same increase in $q_j$ raises utility by only 1.5 in Table 2.

We rule out any situation where providing every additional unit is *always* a dominant strategy; otherwise, there is no collective action problem for any aggregation technology. In addition, we assume throughout that the marginal cost of public goods provision is constant and equal to 2, and deduct the total cost from the utilities in Tables 1 and 2 in order to construct the corresponding collective action game. The payoff for player '$i$', $V_i$, is thus:

$$V_i(q_1, q_2) = U(q_1, q_2) - 2q_i; \quad i = 1,2. \tag{2}$$

The coefficient on $q_i$ on the right-hand side of Eq. (2) is the marginal cost of provision.

The corresponding games are shown in Figs. 1 and 2, respectively, where player 1's contribution levels are indicated in the left-hand column, and player 2's contribution levels are denoted in the top row. Each player can contribute 0, 1, or 2 units. For each of the nine cells denoting the strategic combinations, the number on the left is $V_1$—the *net payoff* of player 1 after applying Eq. (2)—and the number on the right is the net payoff of player 2, $V_2$.

Table 1
The weaker-link (weak diminishing returns)

| $U(q_i, q_j)$ | $q_i$ | $q_j$ |
|---|---|---|
| 0 | 0 | 0 |
| 2 | 1 | 0 |
| 4 | 1 | 1 |
| 3 | 2 | 0 |
| 6 | 2 | 1 |
| 8 | 2 | 2 |

Table 2
The weaker-link (strong diminishing returns)

| $U(q_i,q_j)$ | $q_i$ | $q_j$ |
|---|---|---|
| 0 | 0 | 0 |
| 3 | 1 | 0 |
| 4 | 1 | 1 |
| 3.5 | 2 | 0 |
| 6.5 | 2 | 1 |
| 8 | 2 | 2 |

To use Eq. (2) in constructing a representative payoff in Fig. 1, for example, that associated with $(q_1,q_2) = (1,0)$, we consult Table 1 and find that $U(1,0) = 2$. For player 1, we subtract the cost of provision from $U(1,0)$, thus giving this player a net gain of zero. For player 2, who free-rides, nothing gets deducted from $U(1,0)$. Hence, in Fig. 1 the pair of payoffs corresponding to the row $q_1 = 1$ and column $q_2 = 0$ are $V_1 = 0$ and $V_2 = 2$, respectively. Similarly, for $(q_1,q_2) = (1,1)$, the net gains in Fig. 1 for this strategic combination are 2 for both players. Throughout this paper, we shall compute net payoffs in the strategic-form matrices by applying Eq. (2).

We have the following definitions.

**Definition.** A *strategic-form game* is a set of players, $I$, the strategy set for each player $i \in I$, $S_i$, and a payoff function for player '$i$', $V_i$. The generic (pure) strategy of player '$i$' is denoted as $s_i \in S_i$. A joint strategy is a vector, $s$, consisting of a strategy $s_i$ for each player $i \in I$. The joint strategy space is denoted as $S \equiv \times_{i \in I} S_i$. Finally, a joint strategy, $s \in S$, can be written as $(s_i, s_{-i}) \in S$ where $s_i$ is $i$'s strategy and $s_{-i}$ is the joint strategy for all players other than '$i$'. It follows that $S_{-i} \equiv \times_{j \neq i} S_j$ is the joint strategy space corresponding to the elements $s_{-i} \in S_{-i}$.

**Definition.** Strategy combination $(s_i, s_{-i})$ is a Nash equilibrium for a strategic form game if it satisfies $V_i(s_i, s_{-i}) \geq V_i(\delta_i, s_{-i}) \, \forall \delta_i \in S_i$, for all $i \in I$.

Voluntary contributions are assumed to be characterized by the Nash equilibria for the underlying game. The weaker-link game in Fig. 1 (derived from Table 1)

| $\downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0, 0 | 2, 0 | 3, -1 |
| 1 | 0, 2 | 2, 2 | 4, 2 |
| 2 | -1, 3 | 2, 4 | 4, 4 |

Fig. 1. Weaker-link game derived from Table 1.

| ↓$q_1$/$q_2$→ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0, 0 | 3, 1 | 3.5, -.5 |
| 1 | 1, 3 | 2, 2 | 4.5, 2.5 |
| 2 | -.5, 3.5 | 2.5, 4.5 | 4, 4 |

Fig. 2. Weaker-link game derived from Table 2 (2×2 version is Chicken).

has seven Nash equilibria, from which neither player would unilaterally change his or her strategy. Only the diagonal (matching) strategies, (0,0), (1,1) and (2,2) characterize weakest-link technology. This, in itself, demonstrates that the game is not a variant of the Prisoner's Dilemma. In particular, the Pareto-optimal outcome (2,2) is possible; yet free-riding may also occur, with Nash equilibria at (1,0) and (0,1). Finally, we have 'easy-riding' Nash equilibria (see Cornes and Sandler, 1984), at (2,1) and (1,2). Here, each agent is supplying a positive amount of the public good, but one is providing more than the other (the easy rider).

The game corresponding to Table 2 is shown in Fig. 2 and illustrates the difference between weaker-link and weakest-link. The 2 × 2 version of this game contained within the darkened borders is 'Chicken', with Nash equilibria at (1,0) and (0,1). In Chicken, one player must take on a costly Pareto-improving move (here, set $q_i = 1$). In the 3 × 3 version, the pure-strategy equilibria again correspond to the free-riding outcomes (1,0) and (0,1).

Increased diminishing returns in Table 2 cause agents to be much less forthcoming in providing the public good. For example, the (2,2) outcome is a strict Pareto improvement over the Nash equilibrium; hence, in this second characterization of the weaker-link, the public good is inefficiently underprovided. The difference between equilibria, for these two games, illustrates a trade-off between diminishing returns and the aggregation technology of public supply, not previously identified. In particular, sufficient diminishing returns cause weaker-link behavior to be completely delineated from that of weakest-link because matching contributions are no longer equilibrium outcomes.

Table 3
Better-shot

| $U(q_i, q_j)$ | $q_i$ | $q_j$ |
|---|---|---|
| 0 | 0 | 0 |
| 4 | 1 | 0 |
| 6 | 1 | 1 |
| 7 | 2 | 0 |
| 9 | 2 | 1 |
| 10 | 2 | 2 |

| ↓$q_1$/$q_2$→ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0, 0 | 4, 2 | 7, 3 |
| 1 | 2, 4 | 4, 4 | 7, 5 |
| 2 | 3, 7 | 5, 7 | 6, 6 |

Fig. 3. Better-shot game derived from Table 3.

## 2.2. Better-shot

In better-shot cases, the greatest marginal gain is derived from the player already making the highest overall level of contribution. A formal definition of better-shot technology is:

**Definition.** The technology of public supply aggregation is *better-shot* if:

$$q_j < q_i \Rightarrow \frac{\Delta U}{\Delta q_i} \geq \frac{\Delta U}{\Delta q_j} \quad \forall \Delta q_i = \Delta q_j. \tag{3}$$

For example, consider the entry in Table 3 where $(q_i, q_j) = (1,0)$. An increase in $q_i$ of 1 leads to the combination $(q_i, q_j) = (2,0)$ with the corresponding increase in utility equal to 3. By comparison, if $q_j$ is instead increased by 1, the resulting combination $(q_i, q_j) = (1,1)$ increases utility by 2. The associated game is shown in Fig. 3. It has four pure-strategy Nash equilibria: (2,0) and (0,2) are the usual best-shot Nash equilibria with free-riding, while (2,1) and (1,2) are better-shot Nash equilibria. These latter easy-riding outcomes Pareto-dominate the best-shot equilibria.

## 3. Correlation and institutions: two-player case

Because the agents are identical in terms of the utility received from the public good and the marginal cost of provision, each is subject to Rawls' (1971) notion of the veil of ignorance. A problem arises, however, because in the weaker-link in Fig. 2, the Nash equilibria involve asymmetric provision of the public good: either (1,0) or (0,1). In the same way, the Pareto-superior Nash outcomes for the better-shot in Fig. 3 involve the easy-riding contributions of (2,1) and (1,2). Such alternatives leave open the question of how the participants would decide on who will take on the larger costs of provision? In the better-shot case, the individuals must have a method for deciding who receives the lowest payoff in an easy-riding outcome. Skyrms (1996) denoted such problems as the *curse of symmetry*. As there is no a priori distinction among these outcomes, the individuals are *cursed* in the sense that they must somehow choose who makes the Pareto-improving move.

One way to address the curse of symmetry is through private randomization, which takes the form of a mixed-strategy Nash equilibrium. Consider, for example, the $2 \times 2$ bordered version of the weaker-link in Fig. 2, where each player selects the strategy of $q = 1$ with probability $\sigma$ and $q = 0$ with probability $1 - \sigma$. A symmetric Nash equilibrium in the mixture $\sigma = 0.5$ exists, where in equilibrium, each player is indifferent between the strategy $q = 0$ and $q = 1$. The main drawback of this solution is that it is clearly inefficient, because with probability $(1 - \sigma)^2$, the (0,0) outcome is reached and this limits expected payoffs.

An alternative possibility is that the players build uncertainty into the institutional arrangement itself. Young (1989) has identified Brennan and Buchanan's (1985) veil of uncertainty as a characteristic of international environmental treaties that overcomes the paralysis of collective action associated with the curse of symmetry, and further serves as a Pareto-improving compliance mechanism. In international regimes which exhibit the veil,

> an inability on the part of individual members to predict exactly how the operation of a regime will affect their own welfare enhances incentives to devise equitable and flexible arrangements, and to comply with their requirements once they have entered into force (Young and Demko, 1996, p. 237).

Moreover, in contrast to the veil of ignorance,

> Those facing a 'veil of uncertainty' know what roles they occupy in the relevant social systems. They simply find it hard to foresee the consequences of institutional arrangements because of the generality and longevity of the rights and rules included in various regimes (Young, 1989, p. 232).

Hence, uncertainty built into a regime need not be detrimental to the achievement of its goals and may even create a wider scope of agreement.

Both Schweizer (1990) and Wärneryd (1993) used incomplete information games in order to model the veils of ignorance and uncertainty, where incomplete information (payoff uncertainty) is a consequence of an exogenous randomization process.[5] In contrast, the true power of the veil of uncertainty is that players may purposely design ambiguity into an institution so as to take advantage of the benefits of uncertainty. The uncertainty will incorporate some characteristics of rationality, not available from pure natural chance. For example, in the $2 \times 2$ version of Fig. 2, we would like players to correlate their randomization in such a way that the (0,0) outcome *is never a possibility*, but (1,1) *is*. It follows that this type of randomization should be *public* through institutionally given signals rather than private, and should also be *correlated* to fully capture the potential for Pareto improvements.

---

[5] Schweizer used the Harsanyi transformation for games of incomplete information to capture own-type uncertainty and Wärneryd employed Selten's asymmetric contest for role uncertainty in asymmetric evolutionary games.

Diagram 1 illustrates an example of the veil of uncertainty as applied to the $2 \times 2$ weaker-link game in Fig. 2. The veil ($v$) is represented by a probability distribution, $\rho$, explicitly designed by the players in the construction of a transnational treaty. The distribution $\rho$ represents a randomization over the *joint* strategy space of the players. Specifically, $\rho(q_1, q_2)$ is the joint probability that player 1 takes action $q_1$ and player 2 takes action $q_2$. This distribution is common knowledge, implying that there is no disagreement about each player's ex ante beliefs that they will have to take action $q_i$ within the treaty.

In other words, there is no misunderstanding or disagreement about the various ways in which the treaty can unfold. The actual operationalization of the treaty is the original game itself, which is given in the nodes labeled as $\Gamma_i$. Each $\Gamma_i$ is the *identical* $2 \times 2$ Chicken game embedded within Fig. 2, the difference, being on the move, each player believes he should take under the treaty at the information set associated with that node. In this way players have imperfect information—because they do not know the action to be taken by the other player—but have complete information about his or her payoff because in the end the same Chicken game is played at each node. There is no uncertainty about type/payoffs in the sense of the veil of ignorance.

When player '$i$' believes that he should take action $q_i$ under the treaty, this player finds himself at information set $h_i(q_i)$. For example, if player 1 believes he
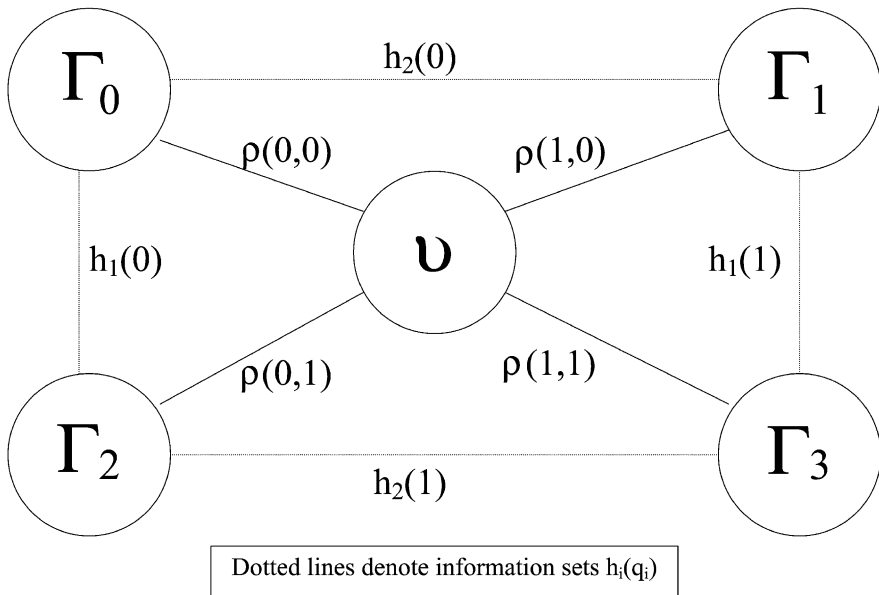


Diagram 1. The veil of uncertainty.

is able to free-ride, then player 1 is at information set $h_1(0)$, at which he cannot tell whether he is at node $\Gamma_0$ or $\Gamma_2$—he has imperfect information regarding player 2's action. Player 1 then forms his beliefs by assigning a conditional probability that he is either at $\Gamma_0$ or $\Gamma_2$ based on the treaty's stipulations, as represented by $\rho(q_1, q_2)$, and the application of Bayes' rule given the information that he is at $h_1(0)$.

We define rational institutional design under the veil of uncertainty as:

**Definition.** A collective action game, $\Gamma$, can be transformed to address the veil of uncertainty if there exists a distribution over joint strategies, $\rho$, such that each player takes action $q_i$ at information set $h_i(q_i)$ in a Bayes/Nash equilibrium of the transformed game.

For example, if the operationalization of the treaty calls for agent 1 to free-ride, this agent should find it in his interest to set $q_1 = 0$ at $h_1(0)$. Alternatively, if it calls for agent 1 to be a provider, then it should be a rational action (in the Nash sense) to set $q_1 = 1$ at $h_1(1)$. Clearly, any Nash equilibrium distribution satisfies this definition. For example, if $\rho(1,0) = 1$, then each player knows that he or she is at node $\Gamma_1$. As $(1,0)$ is a Nash equilibrium, there is no reason why player 1 has an incentive to deviate from $q_1 = 1$, nor does player 2 have an incentive to deviate from $q_2 = 0$. It follows that $\rho(0,1) = 1$ can also occur under the veil, and so can the mixed-strategy replication $\rho(0,0) = \rho(1,0) = \rho(0,1) = \rho(1,1) = 0.25$. The importance of the veil of uncertainty is that it acts as a correlation device to *expand* upon the set of Nash equilibria.

Indeed, Aumann (1974, p. 71) showed that in the $2 \times 2$ version of the weaker-link game in Fig. 2 (Chicken), the veil of uncertainty corresponding to $\rho(0,1) = \rho(1,0) = \rho(1,1) = 1/3$ is an equilibrium outcome that Pareto-dominates any pure or mixed-strategy equilibrium for the game. This follows because the expected payoff is 2 under the veil of uncertainty, operationalized by the institution's signals, while this expected payoff is only 1.5 under private randomization (mixed strategies). Hence, the institution-based veil is a preferred method for breaking the curse of symmetry for situations that resemble Chicken. In essence, the mechanism allows each player to profit from limiting his or her information, provided that the other player knows of this limitation and is directed to play in a helpful fashion. Moreover, what is novel about correlation in Chicken—in comparison, say, to publicly randomizing over the Nash equilibria of a Battle of the Sexes game—is that one can achieve an equilibrium outcome that is *outside* the convex hull of the Nash equilibrium payoffs.[6] In this way, we are foreshadowing

---

[6] For example, if $2.5 < V_i(1,1) < 3$ in the $2 \times 2$ version of Fig. 2, then correlating over $(1,0)$, $(0,1)$ and $(1,1)$ is a strict Pareto improvement over publicly randomizing over $(1,0)$ and $(0,1)$.

what is to come; in Result 1 below, we show that it is entirely possible to reach a Pareto improvement using correlation that does not involve Nash strategies *whatsoever*, but remains self-enforcing.

For those who might object to the explicit modeling of randomization in collective action, we offer the following analogies. First, when two automobiles meet at a perpendicular intersection, the law in many states requires that the right of way is yielded to the driver on the right. For the drivers themselves, a perpendicular encounter is truly a random event; hence, each is correlating on the assignment of action given this chance encounter. Many institutional forms correlate by agreeing to rules, not unlike yielding right-of-way on roads, as to who acts and how. The rules for avoiding accidents at sea, developed by the International Maritime Organization, are an example. In another international context, Young (1998, pp. 44–45) discussed how Farman et al.'s (1985) public announcement of the Antarctic ozone hole mobilized international collective action beyond prior conventions over the ozone problem. This literal sunspot served to correlate actions and led governments to consider strategies that they had not been willing to take individually.[7] Young attributed this exogenous shock to have been influential in leading to the 1987 Montreal Protocol, and lamented that no comparable exogenous shock has led to the mobilization of efforts over global climate change.

At times, correlation comes from an understanding negotiated within an international organization (e.g., NATO or the UN) and, at other times, a country (e.g., the United States) engineers it. For example, in the case of international peacekeeping efforts, correlation can be based on the conflict's location, so that crises in the Pacific Rim (e.g. East Timor) draw more peacekeepers from Oceanic countries (e.g. Australia), while crises in Europe draw peacekeepers more heavily from allies in NATO. At the time that such contingency arrangements are negotiated, signatories do not know the location of future crises. This correlation can result in a Pareto improvement for better-shot public goods such as peacekeeping, as established by Result 2 below.

Technically, the requirements for a Bayes/Nash equilibrium under the veil of uncertainty imply that the distribution $\rho(s_i, s_{-i})$ should be a *correlated equilibrium* (Aumann, 1974, 1987).

**Definition.** Consider the joint strategy set $S \equiv S_i \times S_{-i}$. A *correlated strategy*, $\rho(s) \equiv \rho(s_i, s_{-i})$, is a probability distribution over the joint strategy set:

$$\sum_{s \in S} \rho(s) = 1; \rho(s) \geq 0 \quad \forall s \in S. \tag{4}$$

---

[7] Interestingly enough, Peck and Shell (1991) showed the formal relationship between sunspot equilibria (due to market uncertainty) and the type of correlation illustrated in Diagram 1.

Moreover, $\rho$ generates a *correlated equilibrium* if it satisfies Eq. (5):

$$\sum_{s_{-i} \in S_{-i}} \rho(s_i, s_{-i}) \left[ V_i(s_i, s_{-i}) - V_i(\delta_i, s_{-i}) \right] \geq 0 \quad \forall s_i, \delta_i \in S_i, \forall i. \quad (5)$$

Eq. (4) states that we are creating a veil of uncertainty transformation of the type given in Diagram 1. At each information set, $h_i(s_i)$, the player knows which strategy, $s_i$, he believes he should choose under $\rho(s_i, s_{-i})$, but not the $s_{-i}$ strategies of the other players. Given $s_i$, he updates his beliefs about $s_{-i}$ through his knowledge of $\rho(s_i, s_{-i})$ and Bayes' rule. Eq. (5) is an incentive compatibility condition that states that it is a Bayes/Nash equilibrium for each player to take action $s_i$ at information set $h_i(s_i)$ rather than to deviate to some other action $\delta_i \neq s_i$ at $h_i(s_i)$.

Equating correlated equilibrium with the veil of uncertainty establishes the following incentive for the veil in collective action.

**Result 1.** For the $3 \times 3$ weaker-link game in Fig. 2, the veil of uncertainty represented by $\rho(1,2) = \rho(2,1) = \rho(2,2) = 1/3$ is a Pareto-improving correlated equilibrium. Note: All proofs are provided in Appendix A.

Recall that in Fig. 2, the free-riding outcomes (1,0) and (0,1) remain as Nash equilibria. We have now established that the easy-riding outcomes (2,1) and (1,2) —which were *not* Nash outcomes of the $3 \times 3$ game—can be included under a properly veiled institutional setting. Moreover, the cooperative (2,2) outcome can also occur in combination with the easy-riding outcomes! Such an institutional agreement Pareto-dominates its Nash counterpart and leads to an overall increase in the provision of the public good.

This result is substantially different from the use of correlation in Chicken. The veil of uncertainty can be used to achieve actions that would not otherwise be individually rational. The correlated equilibrium in Chicken relies on the strategies in the Nash equilibrium profiles (1,0) and (0,1). Result 1, by contrast, does not involve the combination of Nash strategies whatsoever. Instead, the uncertainty created by the veil allows players to coordinate their actions where they would otherwise have no incentive to do so. This coordination of collective action yields an expected payoff of 3.67 versus a maximum payoff of 3 in a free-riding Nash equilibrium.

As an application, consider the monetary and fiscal discipline that promotes the stability of a monetary union, such as the EU. In such union, monetary stability is a weaker-link public good, since strengthening the weakest monetary system in the union has the greatest marginal impact on the union's monetary stability. Result 1 indicates that a device which directs members' monetary restraint conditional on centrally given signals can increase stability. Yet some uncertainty must remain, so that it is in the interests of the participants to act insofar as they cannot be sure of the actions of others.

The veil of uncertainty also increases collective action under the better-shot scenario and further reduces the curse of symmetry.

**Result 2.** For the $3 \times 3$ better-shot game (Fig. 3), the veil of uncertainty represented by $\rho(1,2) = \rho(2,1) = \rho(2,2) = 1/3$ is a Pareto-improving correlated equilibrium.

In Fig. 3, the better-shot Nash equilibria—(2,1) and (1,2)—are Pareto improvements over the best-shot equilibria—(2,0) and (0,2). However, these better-shot equilibria suffer from the curse of symmetry, because one player behaves as an easy rider. Overcoming the curse through correlation increases the expected payoff of $q_i = 2$ from 5 to 5.5, because it is now just as probable that player '$j$' fully contributes ($q_j = 2$) as it is that he is easy-riding ($q_j = 1$). Unlike a best-shot situation, actions by more than one player can add to the participants' well-being in a better-shot scenario, and this increases the effectiveness of an institution that correlates action.

We have already offered the manner of assigning peacekeeping operations, in part, by location as an institutional example. Another example of better-shot concerns coordinating efforts to develop medicines, where an institution (e.g. the Centers for Disease Control) issues signals to scientific teams, where the signals do not completely eliminate uncertainty but do serve to correlate actions.

We conclude this section by relating correlated equilibria and the veil of uncertainty to the notion of *political entrepreneurship* (leadership). From Results 1 and 2, correlation has the potential to create a surplus over the payoffs that can be generated through the codification of Nash actions. Frohlich and Oppenheimer (1978) proposed that it is precisely the existence of this surplus which creates an incentive for political entrepreneurship. They defined a political entrepreneur as an individual who coordinates and combines the actions of others in order to supply collective goods and claim part of the surplus. Furthermore, the process of political entrepreneurship often involves the manipulation of information in order to coordinate expectations. This coincides nicely with the most common interpretation of correlation: that the distribution is created by an outside mediator (Aumann, 1974). The discussion of the importance of mediators is widespread within the literature on international relations, yet there are few formalizations of the concept. Correlation is a natural interpretation of international regime mediation.

## 4. Correlation and institutions: three-player case

The conventional wisdom about public goods is that underprovision increases with the number of players. Indeed, the underlying motivation for Hirshleifer (1983) is whether this intuition is robust for weakest-link and best-shot. Hirshleifer intuitively argued that underprovision mildly increases under weakest-link, but that the difference between the Pareto-efficient and voluntary solutions *decreases* under best-shot. We now investigate whether these characteristics persist under the

Table 4
Three-player weaker-link

| $U(q_i,q_j,q_k)$ | $q_i$ | $q_j$ | $q_k$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 3 | 1 | 0 | 0 |
| 4 | 1 | 1 | 0 |
| 7 | 1 | 1 | 1 |
| 3.5 | 2 | 0 | 0 |
| 6.5 | 2 | 1 | 0 |
| 9.5 | 2 | 1 | 1 |
| 8 | 2 | 2 | 0 |
| 12 | 2 | 2 | 1 |
| 13.5 | 2 | 2 | 2 |

weaker-link and better-shot. We also gauge the potential for correlation in light of these effects. We use three-player versions of Figs. 2 and 3.

The three-player weaker-link game is extended in Table 4, so that whenever the third player ($k$) is not contributing, utility is exactly as given in Table 2. The first column denotes the representative player's gross gain (prior to provision costs being deducted) from the combined strategy combinations in each row. The last three columns of Table 4 represent each of the three players' contributions. The corresponding game is given in Fig. 4. Here, player 1 picks the row, player 2 the column, and player 3 the matrix. For example, the first matrix is equivalent to the

| $q_3=0; \downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0, 0, 0 | 3, 1, 3 | 3.5, -.5, 3.5 |
| 1 | 1, 3, 3 | 2, 2, 4 | 4.5, 2.5, 6.5 |
| 2 | -.5, 3.5, 3.5 | 2.5, 4.5, 6.5 | 4, 4, 8 |

| $q_3=1; \downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 3, 3, 1 | 4, 2, 2 | 6.5, 2.5, 4.5 |
| 1 | 2, 4, 2 | 5, 5, 5 | 7.5, 5.5, 7.5 |
| 2 | 2.5, 6.5, 4.5 | 5.5, 7.5, 7.5 | 8, 8, 10 |

| $q_3=2; \downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 3.5, 3.5, -.5 | 6.5, 4.5, 2.5 | 8, 4, 4 |
| 1 | 4.5, 6.5, 2.5 | 7.5, 7.5, 5.5 | 10, 8, 8 |
| 2 | 4, 8, 4 | 8, 10, 8 | 9.5, 9.5, 9.5 |

Fig. 4. Three-player weaker-link game derived from Table 4.

two-player game because player 3's contribution is zero. The order of payoffs is $V_1$, $V_2$, $V_3$.

The free-riding and underprovision that occur in the two-player weaker-link extend immediately to this three-player version. The strategies $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ are Nash equilibria. In addition, easy-riding now emerges in the form of the Nash equilibria $(2,2,1)$, $(2,1,2)$ and $(1,2,2)$. The easy-riding equilibria are Pareto improvements over the free-riding outcomes. In contrast to Hirshleifer's weakest-link analysis, we have an example of *convergence* to cooperation through the addition of another player in the weaker-link.

**Result 3.** For the three-player weaker-link game in Fig. 4, the following veil of uncertainty breaks the curse of symmetry and maximizes collective action: $\rho(2,1,2) = \rho(2,2,1) = \rho(1,2,2) = 0.2$;  $\rho(2,2,2) = 0.4$.

This result provides further reason for establishing a veil of uncertainty through correlation. The outcome is not as extreme as it is in Result 1—we are not correlating over entirely non-Nash profiles—but this is primarily due to the easy-riding Nash outcomes in the three-player case that did not occur for two players. The easy-riding Nash equilibria are Pareto improvements over their free-riding counterparts. Room for further improvement is left only through the cooperative outcome, $(2,2,2)$, whose inclusion further illustrates the potential for the convergence of collective action under the weaker-link.

A three-player version of the better-shot is shown in Table 5, where the utilities associated with $q_k = 0$ are those given in the two-player version in Table 3. By its nature, unilateral and bilateral action are productive under better-shot, implying that the incentives for free- and easy-riding are increased. This intuition is confirmed in Fig. 5. The Nash equilibria of this better-shot game include the free-riding outcomes $(2,0,0)$, $(0,2,0)$ and $(0,0,2)$, which are normally associated with best-shot technology. As in the case of the two-player better-shot, equilibria exist that Pareto-dominate their best-shot counterparts. These are $(2,1,0)$, $(1,2,0)$, $(2,0,1)$, $(1,0,2)$, $(0,2,1)$, and $(0,1,2)$.

Table 5
Three-player better-shot

| $U(q_i,q_j,q_k)$ | $q_i$ | $q_j$ | $q_k$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 4 | 1 | 0 | 0 |
| 6 | 1 | 1 | 0 |
| 6.5 | 1 | 1 | 1 |
| 7 | 2 | 0 | 0 |
| 9 | 2 | 1 | 0 |
| 9.5 | 2 | 1 | 1 |
| 10 | 2 | 2 | 0 |
| 10.5 | 2 | 2 | 1 |
| 11 | 2 | 2 | 2 |

| $q_3=0$; $\downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0, 0, 0 | 4, 2, 4 | 7, 3, 7 |
| 1 | 2, 4, 4 | 4, 4, 6 | 7, 5, 9 |
| 2 | 3, 7, 7 | 5, 7, 9 | 6, 6, 10 |

| $q_3=1$; $\downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 4, 4, 2 | 6, 4, 4 | 9, 5, 7 |
| 1 | 4, 6, 4 | 4.5, 4.5, 4.5 | 7.5, 5.5, 75 |
| 2 | 5, 9, 7 | 5.5, 7.5, 7.5 | 6.5, 6.5, 8.5 |

| $q_3=2$; $\downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 7, 7, 3 | 9, 7, 5 | 10, 6, 6 |
| 1 | 7, 9, 5 | 7.5, 7.5, 5.5 | 8.5, 6.5, 6.5 |
| 2 | 6, 10, 6 | 6.5, 8.5, 6.5 | 7, 7, 7 |

Fig. 5. Three-player better-shot game derived from Table 5.

Furthermore, Hirshleifer's intuition about cooperation and the better-shot is verified and extended. The better-shot equilibria are *not* Pareto-dominated by the cooperative outcome. Convergence between the better-shot and cooperative outcomes occurs because the two cannot be Pareto-ranked.

The benefits of collective action are limited in better-shot, because the participation of the triad is unnecessary to achieve a cooperative payoff. Specifically, the cooperative outcome is no longer $(2,2,2)$ but rather $(2,2,0)$, $(2,0,2)$ and $(0,2,2)$.[8] Now it appears that *cooperation* itself is plagued by the curse of symmetry because two players have to agree to allow the third to free-ride. Yet the structure of the cooperative outcomes suggests a characteristic that has not been previously associated with the better-shot, that of *club formation*.

**Result 4.** The curse of symmetry associated with the $3 \times 3$ better-shot can be broken through club formation. For example, $\rho(2,1,0) = \rho(1,2,0) = \rho(2,2,0) = 1/3$ is a correlated equilibrium for Fig. 5. Permutations of this equilibrium also exist that involve the clubs $\{1,3\}$ and $\{2,3\}$.

Clubs here would allow the internalization of externalities (Cornes and Sandler, 1996). Result 4 looks at a club involving players 1 and 2, which correspond to the

---

[8] The cooperative outcome is generally taken as the utilitarian one who maximizes the sum of the agents' payoffs.

top matrix bordered in bold in Fig. 5. In this two-player game, $(2,2,0)$ is the cooperative outcome, but this is neither a Nash equilibrium nor a Pareto optimum. Atypically, this club equilibrium generates an optimal payoff, or spillover, for the nonmember (free-rider). Clubs are most often associated with partially rival goods with excludable benefits, but this is not the case with a better-shot public good. Instead, two players act in concert to produce a cooperative outcome among them which concurrently happens to maximize the payoff for the nonmember.

This institutional prescription for subgroups to form and then to correlate their activities within, but not between, subgroups is reflected in real-world examples. Our previous observation regarding the use of location of a crisis as a signal for potential participants' responsibility in peacekeeping can be extended to another level for three or more countries. Based on Result 4, it is advisable for two or more subgroups of potential peacekeepers to form depending on location. In NATO, for example, two such subgroups have formed to address contingencies in Europe—the Western European Union with members from both the EU and NATO, and a subgroup with just France and Germany. This is consistent with our result that such subgroups should characterize better-shot, but not weaker-link, public goods.

We are left with several insights regarding better-shot public goods. First, cooperation does not require collective action on the part of the grand coalition. Each permutation of the $(2,2,0)$ outcome generates a level of societal welfare that is greater than determined by $(2,2,2)$. As such, requiring the grand coalition to act is unnecessary. Second, the gap between cooperation and voluntary collective action can decrease as the number of players increase, and moreover, correlation can further decrease this gap. Third, correlation involves club-like behavior that produces its own positive externality (spillover) for the nonmember.

## 5. Differences in marginal cost

For public good problems, such as pollution or disease control, differences in the cost of provision or abatement can be a function of a country's level of development. We must, however, be cautious and acknowledge that the direction of the difference in marginal cost between developed and developing nations depends on the level of the activity and the underlying technology in the two regions. Consider, for example, the case of limiting sulfur emissions from power plants to reduce acid rain and particulates in the air. The marginal cost of abatement increases with its level and may be near zero for low abatement (International Institute for Applied Systems Analysis, 1993). Thus, even if the developed world's marginal cost is smaller than developing countries for the same level of abatement, the marginal cost in the former may well exceed that of the latter, if the former undertakes abatement to a greater extent. Hence, the marginal

| $\downarrow q_1/q_2 \rightarrow$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0, 0 | 3, 2 | 3.5, 1.5 |
| 1 | 1, 3 | 2, 3 | 4.5, 4.5 |
| 2 | -.5, 3.5 | 2.5, 5.5 | 4, 6 |

Fig. 6. Asymmetric weaker-link.

cost difference may be of either sign depending on abatement levels and/or the costs associated with the available technologies.

To indicate the influence of such asymmetry, we can introduce differences in marginal costs. For example, for the weaker-link game presented in Fig. 2, we leave the marginal cost of the row player unchanged at 2, and lower the marginal cost of the column player to 1. The altered strategic-form game is shown in Fig. 6, where the payoffs for player 2 in the two right-most columns are raised by the savings in marginal cost (i.e., one time the number of units provided by the agent).

There are two pure-strategy Nash equilibria at (0,1) and (1,2). The first is the original symmetric case, while the second is new. The latter easy-rider equilibrium represents a greater degree of provision, even in the absence of leadership. The earlier free-rider Nash equilibrium at (1,0) no longer applies, so that the curse of symmetry is eliminated by cost asymmetry.[9] Cost asymmetry can be dealt with through the correlated equilibrium $\rho(0,1) = \rho(2,1) = \rho(2,2) = 1/3$, which strictly improves the well-being of the low-cost provider and leaves the high-cost provider's expected welfare unchanged. This implies that the greater the difference in marginal costs, the smaller the need for the veil of uncertainty or leadership associated with correlating strategies.

## 6. Concluding remarks

This paper has focused on weaker-link and better-shot public goods. Such goods represent two prevalent classes of transnational public goods. Weaker-link voluntary contribution games are characterized by multiple Nash equilibria, none of which may be efficient. The design of an institution that fully bypasses inefficient Nash equilibria is important, because many existing transnational treaties and alliances have been characterized as merely codifying Nash behavior. We have shown that such institutions can be associated with a new type of game form, where a Pareto-improving correlated equilibrium exists which does not

---

[9] Although the cost asymmetries make a difference for weaker-link public goods, this is not the case for weakest-link public goods where equilibria will remain along the diagonal with matching behavior. With cost asymmetries, the high-cost agent will be the limiting agent for the match.

involve Nash strategies. For example, correlated strategies represent a formaliza-tion of both the veils of uncertainty and leadership, which are essential concepts of government intervention that, heretofore, lacked any formal structure. Further-more, abiding by the leader or institution's correlated instructions is self-enforcing, and as such, may limit transaction/enforcement costs.

In the case of better-shot public goods, the Nash equilibria also diffuse away from the hallmark free-riding outcomes to include easy-riding equilibria, where more than one agent participates, but by unequal amounts. Since efficient better-shot correlated equilibria do not require full participation, the veil of uncertainty's results in club-like behavior. This implies that institutional structures for weaker-link public goods cannot necessarily be applied to better-shot public goods. *Alternative aggregation technologies* affect payoff patterns and, in so doing, *influence the kinds of institutional structures that can foster efficiency.*

## Appendix A. Proofs of Results 1–4

**Proof of Result 1.** We begin with Eq. (5) under the condition that $q_1 = 1$. This corresponds to agent 1's actions at information set $h_1(1)$.

**Case 1.** $\delta_1 = 0$ at $h_1(1)$. Then Eq. (5) becomes $\rho(1,2) [4.5 - 3.5] \geq 0$, which strictly holds.

**Case 2.** $\delta_1 = 2$ at $h_1(1)$. Then Eq. (5) becomes $\rho(1,2) [4.5 - 4] \geq 0$, which strictly holds.

These two cases imply that $\rho(1,2)$ is unrestricted by any deviation $\delta_1 \neq 1$. By symmetry $\rho(2,1)$ is also unrestricted for all $\delta_2 \neq 1$. We now turn to the conditions under which $q_1 = 2$ at $h_1(2)$.

**Case 3.** $\delta_1 = 0$. Then Eq. (5) becomes $\rho(2,1) [2.5 - 3] + \rho(2,2) [4 - 3.5] \geq 0$. This implies $\rho(2,2) \geq \rho(2,1)$.

**Case 4.** $\delta_1 = 1$. Then Eq. (5) becomes $\rho(2,1) [2.5 - 2] + \rho(2,2) [4 - 4.5] \geq 0$. This implies $\rho(2,1) \geq \rho(2,2)$.

Together, Cases 3 and 4 imply that $\rho(2,2) = \rho(2,1)$. By symmetry, the require-ment that $q_2 = 2$ at $h_2(2)$ implies the restriction $\rho(2,2) = \rho(1,2)$. Hence, $\rho(1,2) = \rho(2,2) = \rho(2,1) = 1/3$ is a correlated equilibrium.    □

**Proof of Result 2.** If either player is at the information set $h_i(1)$, he knows that $q_j = 2$. Since $(q_i, q_j) = (1,2)$ is a Nash equilibrium $\forall i, j = 1,2$, $i \neq j$, we know that '$i$' will take action $q_i = 1$ at information set $h_i(1)$. It, therefore, remains to be

shown that $q_i = 2$ at $h_i(2)$. By symmetry, we can, without loss of generality, check that Eq. (5) holds for $q_1 = 2$ at $h_1(2)$.

For either $\delta_1 = 0$ or 1, Eq. (5) becomes: $(1/3)[5 - 4] + (1/3)[6 - 7] \geq 0$, which weakly holds.                                                                                 □

**Proof of Result 3.** At information set $h_i(1)$ player '$i$' knows that $q_{-i} = (2,2)$ and $q_i = 1$ is the best response to $q_{-i} = (2,2)$. It follows that '$i$' has no incentive to deviate to some $\delta_i \neq 1$ at $h_i(1)$. The remainder of the proof requires that we show $q_i = 2$ at $h_i(2)$. Again, without loss of generality we focus on the case of player 1.

**Case 1.** $\delta_1 = 0$ at $h_1(2)$. Eq. (5) becomes: $0.2[8 - 6.5] + 0.2[8 - 6.5] + 0.4[9.5 - 8] \geq 0$.

**Case 2.** $\delta_1 = 1$ at $h_1(2)$. Eq. (5) becomes: $0.2[8 - 7.5] + 0.2[8 - 7.5] + 0.4[9.5 - 10] \geq 0$.                                                                                 □

**Proof of Result 4.** In the club equilibrium, $\rho(2,1,0) = \rho(1,2,0) = \rho(2,2,0) = 1/3$, it is always a best reply for $q_3 = 0$ at information set $h_3(0)$. Hence, we need to check Eq. (5) for players 1 and 2 only. Furthermore, $q_i = 1$ is the best reply to $q_{-i} = (2,0)$, implying that $q_i = 1$ at $h_i(1) \forall i = 1,2$. This leaves us with the question of whether $q_i = 2$ at $h_i(2) \forall i = 1,2$.

**Case 1.** $\delta_1 = 0$ at $h_1(2)$: $\rho(2,1,0)[5 - 4] + \rho(2,2,0)[6 - 7] \geq 0$, which implies $\rho(2,1,0) \geq \rho(2,2,0)$.

**Case 2.** $\delta_1 = 1$ at $h_1(2)$: $\rho(2,1,0)[5 - 4] + \rho(2,2,0)[6 - 7] \geq 0$, which implies $\rho(2,1,0) \geq \rho(2,2,0)$.

By symmetry, $q_2 = 2$ at $h_2(2)$ requires $\rho(1,2,0) \geq \rho(2,2,0)$.                □

## References

Arce M., D.G., 1997. Correlated strategies as institutions. Theory and Decision 42, 271–285.

Aumann, R.J., 1974. Subjectivity and correlation in randomized strategies. Journal of Mathematical Economics 1, 67–97.

Aumann, R.J., 1987. Correlated equilibria as an expression of Bayesian rationality. Econometrica 55, 1–18.

Barrett, S.A., 1994. Self-enforcing international environmental agreements. Oxford Economic Papers 46, 878–894.

Barrett, S.A., 1998. On the theory of diplomacy of environmental treaty-making. Environmental and Resource Economics 11, 317–333.

Brennan, G., Buchanan, J.M., 1985. The Reason of Rules: Constitutional Political Economy. Cambridge Univ. Press, Cambridge.

Cornes, R., 1993. Dyke maintenance and other stories: some neglected types of public goods. Quarterly Journal of Economics 108, 259–271.

Cornes, R., Sandler, T., 1984. Easy riders, joint production, and public goods. Economic Journal 94, 580–598.

Cornes, R., Sandler, T., 1996. The Theory of Externalities, Public Goods, and Club Goods. 2nd edn. Cambridge Univ. Press, Cambridge.

Farman, J., Gardiner, B., Shanklin, J., 1985. Large losses of total ozone in Antarctica reveal seasonal $CIO_x /NO_x$ interaction. Nature 315, 207–210.

Frohlich, N., Oppenheimer, J.A., 1978. Modern Political Economy. Prentice-Hall, Englewood Cliffs, NJ.

Hirshleifer, J., 1983. From weakest-link to best-shot: the voluntary provision of public goods. Public Choice 41, 371–386.

International Institute for Applied Systems Analysis (IIASA), 1993. RAINS model version 6.1 (computer program). IIASA, Laxenburg, Austria

Kaul, I., Grunberg, I., Stern, M.A. (Eds.), 1999. Global Public Goods: International Cooperation in the Twenty-First Century. Oxford Univ. Press, New York.

Kindleberger, C.P., 1981. Dominance and leadership in the international economy: exploitation, public goods, and free rides. International Studies Quarterly 25, 242–254.

Mueller, D.C., 1989. Public Choice II. Cambridge Univ. Press, Cambridge.

Murdoch, J.C., Sandler, T., 1997. The voluntary provision of a pure public good: the case of reduced CFC emissions and the Montreal Protocol. Journal of Public Economics 63, 331–349.

Murdoch, J.C., Sandler, T., Sargent, K., 1997. A tale of two collectives: sulfur versus nitrogen oxides emission reduction in Europe. Economica 64, 281–301.

Olson, M., 1965. The Logic of Collective Action. Harvard Univ. Press, Cambridge, MA.

Olson, M., Zeckhauser, R., 1966. An economic theory of alliances. Review of Economics and Statistics 48, 266–279.

Peck, J., Shell, K., 1991. Market uncertainty: correlated and sunspot equilibria in imperfectly competitive economies. Review of Economic Studies 58, 1011–1029.

Rawls, J., 1971. A Theory of Justice. Harvard Univ. Press, Cambridge.

Sandler, T., 1997. Global Challenges: An Approach to Environmental, Political, and Economic Problems. Cambridge Univ. Press, Cambridge.

Sandler, T., 1998. Global and regional public goods: a prognosis for collective action. Fiscal Studies 19, 221–247.

Sandler, T., Murdoch, J.C., 1990. Nash–Cournot or Lindahl behavior?: An empirical test for the NATO allies. Quarterly Journal of Economics 105, 875–894.

Sandler, T., Sargent, K., 1995. Management of transnational commons: coordination, publicness, and treaty formation. Land Economics 71, 145–162.

Schotter, A., 1981. The Economic Theory of Social Institutions. Cambridge Univ. Press, Cambridge.

Schweizer, U., 1990. Calculus of consent: a game-theoretic perspective. Journal of Institutional and Theoretical Economics 146, 28–54.

Skyrms, B., 1996. Evolution of the Social Contract. Cambridge Univ. Press, Cambridge.

Wärneryd, K., 1993. Anarchy, uncertainty, and the emergence of property rights. Economics and Politics 5, 1–14.

Young, O.R., 1989. International Cooperation: Building Regimes for Natural Resources and the Environment. Cornell Univ. Press, Ithaca, NY.

Young, O.R., 1991. Political leadership and regime formation: on the development of institutions in international society. International Organization 45, 281–308.

Young, O.R., 1998. Creating Regimes: Arctic Accords and International Governance. Cornell Univ. Press, Ithaca, NY.

Young, O.R., Demko, G.J., 1996. Improving the effectiveness of international environmental governance systems. In: Young, O.R., et al. (Eds.), Global Environmental Change and International Governance. Dartmouth Univ. Press, Hanover, Chapter 10.

Zacher, M.W., 1996. Governing Global Networks: International Regimes for Transportation and Communications. Cambridge Univ. Press, Cambridge.