5

Co-ordination and Leadership in the Unregulated Common Property: Some Lessons from Game Theory

In the previous chapter, describing the various advantages of small-group settings, we have drawn attention to the fact that the prisoner's dilemma is not necessarily an appropriate representation of the payoff structure obtaining in situations of unregulated common property. We now want to delve into this issue by discussing different game forms that may be more relevant in this respect, particularly the so-called chicken game, the co-ordination game, and heterogeneous games that combine the payoff structures of the PD and these two games. Note that so far, the analysis has been essentially focused on *appropriation* problems understood as collective action problems concerned with excluding potential beneficiaries and allocating the subtractable flow of an existing common property resource. Another category of collective action problems may arise: these are the *provision* problems which are to do with the process of creating a resource, maintaining or improving its production capabilities, or avoiding its destruction. The latter type of problems are obviously akin to the well-known *public goods* problems (Ostrom, 1990: 30–3; Ostrom *et al.*, 1993).

5.1 Unilateral Contribution

Let us first consider a kind of situation in which everybody agrees that something is to be done but the problem is who will actually do it. In the real world, indeed, one encounters many situations in which a collective problem is solved through unilateral action. Olson has clearly such situations in mind when he discusses the possibility of groups being *privileged*. A *privileged* group is defined as a group so small 'that each of its members, or at least some one of them, has an incentive to see that the collective good is provided, even if he has to bear the full burden of providing it himself' (Olson, 1965: 50). This happens when the collective good (or some quantity of it) can be obtained at a cost sufficiently low in relation to its benefit that it pays any individual in the relevant group to provide it all by himself. In other words, the total gain resulting from the production of a collective good or from the avoidance of a collective bad is so large in relation to the total cost that an individual's share of the aggregate gain would exceed the total cost. Such a situation may also obtain when the group is composed of individuals of unequal size or extent of interest in the collective good. According to Olson, in *heterogeneous* groups, the greater the interest in the collective good of any

		Player B: C	Player B: D
Player A	C	8, 8	6, 10
	D	10, 6	2, 2

FIG. 5.1. A 2 × 2 symmetrical chicken game

single member, the greater the likelihood that this 'large' member will get such a significant proportion of the total benefit from the collective good that he will gain from seeing that the good is provided, even if he has to pay all of the cost himself [ibid. 33–4]. As we shall see below, however, the case of heterogeneity may cover a wide variety of different situations which need to be carefully specified. Indeed, the meaning behind the notion of interest has to be clarified in order to be able to predict which type of resource user (the rich or the poor) is more likely to contribute.

Collective action in privileged groups

Let us begin by examining the situation obtaining in small homogeneous groups. This situation can be portrayed as a two-person chicken game, that is, as a game in which there are two (Nash) equilibria in pure strategy and in each of these one player co-operates while the other defects. Consider the game with the payoff structure described in Figure 5.1.

The following provision problem can serve as a first illustration. Assume that players A and B are two small-scale marine fishermen who stand threatened by the invasion of a fleet of foreign trawlers in their traditional fishing-zone. The incomes of A and B are presently 10 units. However, to maintain the level of their present catches, they need to have their inshore waters legally protected against the encroachments of the foreign boats. It is assumed that they can succeed in securing that protection from their government provided that they make lobbying efforts (which is the 'public good' to be produced) the total cost of which amounts to 4 units. If both players defect in the sense that they refuse to exercise pressure on their government until it agrees to act in their interests, the competition from the highly efficient foreign trawlers will soon bring their individual incomes down to a mere 2 units (D,D). The total benefit of the lobbying action—that is, the joint payoff (C,C) minus the joint payoff (D,D), or 16–4 = 12—is therefore much larger than its total cost (4). If A and B co-operate in the sense that they agree to share that cost equally between them (C,C), they will both be assured of receiving a payoff of 8 units. Moreover, the lobbying action is so rewarding that it can pay a single player to bear the entire cost of it (D,C, or C,D): the politically active player (the 'sucker') will then get 6 units while the passive player (the free-rider) will of course get more, 10 units in this instance.

From Figure 5.1, it is then easy to see that players have no dominant strategies in this game. What each will decide to do depends on what he expects the other to do. As in a PD game, each player prefers that the other undertakes the lobbying action while he refrains from moving since he will thereby get the maximum possible income (10 units). However, and contrary to the situation obtaining in a PD game, each is willing to take full responsibility for that action if the other refuses to do anything about it. In other words, the consequences of nobody doing the lobby are so disastrous that either of the players would undertake it if the other did not (it is better to be a 'sucker' and to get 6 units of net income than not to be a 'sucker' and get only 2 units).[1]

The problem in the above game is of course that the best symmetric payoff (8,8) cannot be achieved because (C,C) is not an equilibrium outcome: player A would choose to defect if he expected player B to choose to co-operate. However, *with communication*, the players can make a *self-enforcing* plan of action (in the sense that neither player could gain by unilaterally deviating from this plan) that gives them both higher expected utility than what they can get in the absence of communication. In other words, they can benefit, from co-ordinating their actions. Thus, they could agree to toss a coin and then choose (C,D) if it is heads and (D,C) if it is tails. As pointed out by Myerson, this plan of action is self-enforcing, even though the coin toss has no binding impact on the players: player A could not gain by choosing D after heads, since player B is then expected to choose D (Myerson, 1985: 254; 1991: 249–50). In game theory, this plan is known as a *correlated equilibrium*. In our example, it gives each player an expected utility of 8 ($0.5 \times 6 + 0.5 \times 10$). This is a better result than that achieved under the randomized or mixed-strategy equilibrium (7.33) where, in one-ninth of the cases, the worst outcome (D,D) will materialize.

Under a correlated equilibrium, agents endogenously and *non-co-operatively* generate a *co-ordinated* solution which gives them the assurance that collective action will take place: one of them will have to undertake it, but under a scenario agreed on by everybody.

There are many other applications of the chicken game to provision problems, such as maintenance and surveillance of common properties (irrigation systems, pastures, collective fields, hunting-grounds), in particular protection tasks aimed at enforcing exclusive rights against possible intrusion by outsiders, activities of lobbying and political representation, and initiation of collective action. The latter example actually refers to the problem of leadership and is noticeably present in all situations where the production of a public good is envisaged. It should nevertheless be borne in mind that not all the aforementioned activities can be described as chicken games. As we shall see below, this depends on the particular configuration of costs and benefits obtaining in each situation.

[1] There are actually three Nash equilibria of this game: not only (C,D) and (D,C), but also a randomized Nash equilibrium, called a 'mixed strategy' equilibrium in which each player plays his different strategies with a certain, predetermined probability. In the mixed equilibrium of the game considered here, each player co-operates with probability 2/3 and defects with probability 1/3. The expected utility for each is therefore equal to $2/3(2/3 \times 8 + 1/3 \times 6) + 1/3(2/3 \times 10 + 1/3 \times 2) = 7.33$.

		Fisherman B	
		enters	does not enter
Fisherman A	enters	−1, −1	10, 0
	does not enter	0, 10	0, 0

FIG. 5.2. Assignment problem as a 2 × 2 chicken game

Contrary to a widely held view, chicken games may also be suitable for depicting appropriation problems. Consider the case of fishing of lobsters or some other bottom-dwelling species. It is easy to find situations in which the number of (potential) fishermen exceeds the number of locations where enough lobsters can be caught to justify the effort involved. There then arises the problem as to how those fishing locations will be assigned so as to avoid the eruption of harsh conflicts (since two fishermen fishing on the same spot would be doing so at loss). In some fishing communities, such a problem is solved by recognizing the claim of the 'first entrant' into the fishing-ground (see Chapter 10, sect. 2). This solution clearly corresponds to a chicken game as illustrated in Figure 5.2.

Assigning exclusive rights to the first entrant is not the only possible solution to the problem concerned. Assignments by lottery are not infrequently practised which really amount to a correlated equilibrium as defined above (see also Chapter 10).

Extending the chicken game to *N*-player situations is quite straightforward: there are as many Nash equilibria in pure strategies as there are players and, in each of them, a single player contributes while the others freeride. In an infinitely repeated framework, the folk theorem also holds, implying the existence of a plethora of equilibria. Paradoxically perhaps, some of these equilibrium outcomes may be characterized by a large number of 'non-co-operative' rounds: indeed, a player may have an interest in building a reputation of being 'tough' so as induce another player to contribute.

It is noteworthy that, in the above two numerical examples, the equilibrium outcomes in pure strategies are Pareto-optimal. This is not necessarily so, at least in the one-shot game. For the sake of illustration, let us consider the case of two fishermen who must decide how many boats to put out at sea. If they put out one boat each, their payoffs are (4,4). If they put out two boats each, the catch per boat becomes so low that the net profit (payoff) is negative so that the outcome is (−1,−1). On the other hand, if one of the fishermen puts out two boats while the other puts only one boat, the former's payoff jumps to 6 while the latter's payoff is 1. The corresponding payoff matrix is given in Figure 5.3.

In the one-shot version of the above chicken game, there are two Nash equilibria in pure strategies {(1,6), (6,1)} and one in mixed strategy, where putting out one boat is played with a probability of 1/2, and putting out two boats, with a probability of 1/2, yielding an expected payoff of 2.5 to each fisherman. It is therefore evident that the

no. of boats	Fisherman B 1	2
Fisherman A 1	4, 4	1, 6
2	6, 1	−1, −1

FIG. 5.3. A chicken game with a non-(Pareto) optimal equilibrium

Pareto-optimum (4,4) cannot be achieved. In the long finitely repeated version of that game (and *a fortiori* in an infinitely repeated version), however, the Pareto-efficient allocation can be approximately reached if each player obeys the following equilibrium strategy as described in Myerson (1991: 338). Each fisherman puts out only one boat as long as both of them have done so, until the last two rounds. If the two fishermen have always abided by this plan of action, then, in the last two rounds, they both play the mixed equilibrium that yields expected payoffs (2.5, 2.5) at each round. On the other hand, if, in a round before the last two rounds, either fisherman ever deviates from the strategy of putting only one boat at sea, then the two fishermen thereafter play the equilibrium that, at each round, gives payoff 1 to the fisherman who deviated first and payoff 6 to the other one. (If both deviate first at the same round, then let us say that the fishermen act as if fisherman A deviated first). It is a subgame-perfect equilibrium for both fishermen to behave according to this scenario in the T-round finitely repeated game (for any positive integer T). Furthermore, this equilibrium gives an expected average payoff per round of $((T - 2) \times 4 + 2 \times 2.5)/T$, which gets close to 4 when T is large.

It is important to note that there are also equilibria of such finitely repeated games that are worse for both players than any equilibrium of the one-shot game, but of course are not worse than the minimax value (for more details, see Myerson, 1991: 339). One can therefore conclude that, contrary to a widely held view, repetition does not always improve the average performance of the group.

The above results actually illustrate an important theorem of game theory which has been proved by Benoit and Krishna (1985) following which, if a strategic-form game has multiple equilibria (when it is not repeated) that give two or more different payoffs to each player, then, under general conditions, the average payoffs in subgame-perfect equilibria of long finitely repeated versions of this game are very close to *any* average payoffs attainable in the *infinitely* repeated versions of the game (Myerson, 1991: 338). (Bear in mind that the folk theorem applies in the latter case.) This theorem points up the essential difference that exists between prisoner's dilemmas and other game forms and the consequent danger, frequently encountered in the so-called 'collective action' literature, of confining the analysis to the former as though the theoretical propositions applying to it could be easily generalized to other games. More specifically, while in PD games, the results obtained under infinitely repeated

		Player B (poor)	
		C	D
Player A (rich)	C	15, 3	13, 5
	D	17, 1	2, 0

FIG. 5.4. A 2 × 2 asymmetrical chicken game

versions radically differ from those obtained in finite versions, this is not true of other game forms.

Collective action in heterogeneous groups

Let us now turn to situations where asymmetric payoffs prevail, reflecting economic inequalities within the small group. As pointed out above, Olson then predicts that the richer party will necessarily contribute because he has a higher stake. To examine the validity of this prediction, let us start by considering the chicken game presented in Figure 5.4 in which it is evident that player A (the richer player) has a lot to lose if (D,D) occurs.

As is typical of any 2 × 2 chicken game, there are two equilibria in pure strategies leading to an asymmetric outcome (D,C) or (C,D), and an equilibrium in mixed strategy where all outcomes have a positive probability of occurrence. It cannot therefore be taken for granted, as Olson seems to do, that the richer party necessarily plays C in equilibrium. This is all the more so if we assume that the richer party is superior not only in economic terms but also in terms of power. As a matter of fact, in the kind of situation contemplated here, one obvious form which the power of the rich can take on is his ability to make a move prior to the poorer and weaker party. If this is so, the appropriate tool for analysing the decision problems of both parties is no more the strategic form game used above but a two-step chicken game best described in extensive form. In a two-player extensive game, decision-making is sequential so that one player (the follower) can observe what the other (the leader) has done before deciding his own action. In Figure 5.5, we have translated the payoff matrix described in Figure 5.4 into such an extensive form.

Player B will make his decision to play C or D only after having observed what player A has done. His best strategy is clearly 'play C if A has played D, and play D if A has played C'. The strength of player A lies in the fact that, owing to his privileged position as the first mover, he is able to take into consideration B's best reactions while making his own decision in the first stage of this sequential game. In this case, it is obvious that he will choose to play D in the expectation that player B will choose C. The outcome (D,C) is therefore the unique subgame-perfect Nash equilibrium. Note carefully however that, for the above result to obtain, it is necessary that the first move

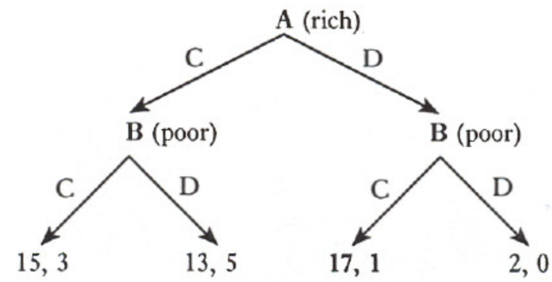

FIG. 5.5. A 2 × 2 asymmetrical chicken game with sequential moves

by the richer party is perceived by the poorer party as an unmistakable signal about the intention of the former. This is precisely the ability of the richer party to commit himself to a particular plan of action and to signal it in an unambiguous way to the weaker party that constitutes the source of his power. If such power imbalance did not exist, one could argue that the above representation is unsuitable to describe the decision problem under concern. This is because, if power is more or less evenly distributed, the position of the first mover cannot be clearly established: one player can give a signal to the other to which the latter may respond by a counter-signal, to which the former again responds, etc. The analysis of such games, which involve continuous time, lies beyond the scope of the present work.

Even in the absence of any explicit demonstration of power by the richer party, the sheer poverty of the poorer party may create a situation in which he is forced to contribute. This may most obviously happen when the latter's subsistence is threatened by, say, the neglect or the disappearance of a common property. Consider the following problem: a land area is being irrigated through a surface canal system and the proper maintenance of some water control infrastructure is crucial for the satisfactory performance of the system. The land in the command area is unequally distributed among two persons: the rich and the poor. Yet, contrary to the rich who possesses a large portfolio of assets including the irrigated parcels, the poor has no other wealth besides his tiny plots of irrigated land. In such circumstances, it is a serious possibility that the rich is less mindful of the proper maintenance of the irrigation system. We moreover assume that maintenance works can be effectively carried out by a single player alone. That this is not a completely unrealistic assumption is attested by the observation made by Yoder (1986) about the Thambesi irrigation system in Nepal. As a matter of fact, we are told that maintenance carried out prior to the monsoon rains requires 'only four to five hours of work with all the members participating', as a result of which it is possible for only a few farmers to keep the whole system going (Yoder (1986), quoted from Ostrom and Gardner, 1993: 97).

Fitting the above set of assumptions is the payoff matrix depicted in Figure 5.6.

Two Nash equilibria in pure strategies (and none in mixed strategy owing to the infinitely negative payoff accruing to the poor in case of universal neglect) characterize the above game. This being said, there is good ground to believe that, given the peculiar circumstances highlighted above, the two 'pure' Nash equilibria are not

		Poor	
		maintains	neglects
Rich	maintains	15, 3	13, 5
	neglects	17, 1	2, −∞

FIG. 5.6. A 2 × 2 asymmetrical chicken game with a subsistence constraint

equally plausible. Indeed, one of them (neglects, maintains), is likely to emerge as a 'focal' point because, the poor, being subsistence-minded, is not ready to incur the least risk that the system ends up remaining unmaintained. In fact, what this means is that, given the binding nature of the subsistence constraint, the poor tends to adopt a maximin strategy which is justifiable in rational terms. In particular, if there is the slightest chance of the rich being irrational, mindless, or prone to making 'mistakes', the poor is eager to ensure his livelihood by choosing the safe outcome, which is personally to carry out the maintenance work. (In game-theoretical terms, the (neglects, maintains) outcome is the only trembling-hand equilibrium.)

In actual fact, it is not difficult to find, in the real world, situations in which the poorer segments of the population have a vital interest in the preservation of common properties, while the rich do not have that concern because they have available to them significant *exit* options. Here lies another source of power in the hands of the rich.

The array of possibilities may be further enlarged if we allow for the fact that the situation of the poor may be so desperate that he cannot be forced to contribute. Consider the following appropriation problem. There are two players who can choose between a conservationist strategy, C, or a more destructive strategy, D, regarding the exploitation of a common property resource. One of them is very poor, so that he cannot afford to follow *alone* the conservationist strategy, C. (For instance, his concern about sheer survival in the short run may lead him to heavily discount the future benefits of such a strategy and to overweigh the present costs, an issue which has already been addressed in Chapter 3.) Such is not the case, however, with the other player who, being richer, can safely adopt a more conservative use of the resource. This situation is depicted in Figure 5.7.

Evidently, the above game is not a chicken game. The poor has a dominant strategy, which is to defect, while the rich has a payoff structure characteristic of a chicken game. The only equilibrium outcome is (13,5), in which the poor, despite his weaker economic position, cannot be coerced into being the single contributor. Such a result (which is Pareto-inferior) also holds in any finitely repeated version of the game, as was observed with the PD game, with which it shares the feature of yielding a unique equilibrium. (The theorem of Benoit and Krishna (1986) does not apply.)

An important lesson to be drawn from the latter example is that a player does not enjoy a decisive advantage over another player who is hard-pressed by a survival

		Poor	
		C	D
Rich	C	15, 4	13, 5
	D	17, –1	2, 2

FIG. 5.7. A 2 × 2 asymmetrical heterogeneous game with a subsistence constraint

		Poor	
		C	D
Rich	C	15, 4	16, 2
	D	14, 2	2, 2

FIG. 5.8. A 2 × 2 asymmetrical game with a norm of participation

constraint. This arises from the fact that, in some circumstances in which he has actually no choice, the poor may credibly precommit himself to non-co-operation. In this case, power, understood as the ability to precommit to non-co-operation, appears to be on the unexpected side.

A last form of power is worth mentioning, which is perhaps more subtle or disguised than those analysed above. Under this form, the stronger party assumes a leadership role in devising and putting into effect a system of sanctions (in the form of payoff transfers) that punishes all free–riders, including himself. This allows him to bring pressure to bear on the weak party so as to make him share the burden of collective action even though the latter has much less interest than the former in the success of this action. In other words, the strong party agrees to bind himself to co-operation in order to bring the weak party to co-operate. One can think of a norm of participation according to which everybody, rich or poor, must join the collective action irrespective of whether there is much or little to gain from participating. By referring to such a norm, the rich and powerful can thus drive the poor and weak to take actions that entail large benefits for the former and small (or even negative, as in the game below) ones for the latter (for example, all landowners, both large and small, participate to an equal extent to the construction of an irrigation canal). This situation is illustrated by the game portrayed in Figure 5.8, which is actually a modified version of the game depicted in Figure 5.7 in which a player pays a fine of 3 units to the other player whenever he free–rides on the other's effort. (We return in Chapter 8 to the role of sanctions as a co-operation-enforcing mechanism.)

It is easy to verify that the game in Figure 5.8 is virtuous in the sense that universal co-operation is the only equilibrium outcome. What needs to be stressed is that in this instance such an outcome has been achieved through a subtle manipulation. As a matter of fact, the stronger party has resorted to a genuine strategy of delusion. By threatening himself with sanctions for any act of free–riding on the weaker party's effort, he conceals the fact that he has no interest in defecting: clearly, the threat is empty as far as he is concerned.

That power is effectively exercised in the above instance is evident from the fact that, were the poor allowed to do it, he would vote against establishing the aforedescribed norm of participation and the payoff transfers associated with it. As a matter of fact, he earns a payoff of 4 units after this norm has been laid down whereas he could earn 5 units in the initial situation. The rich will of course vote in the opposite way since the change in the payoff matrix brings him a payoff gain of 2 units (he earns 15 instead of 13 units).

Note carefully that the problem of whether or not to establish the norm of participation with the attendant payoff transfers is quite distinct from the question as to whether, once established, the punishment system involved is self-enforcing. It is easy to see that, if the right to punish is vested with either the rich or the poor, punishment will not be meted out whenever the agent (whether rich or poor) has to punish himself. If, however, the punishment mechanism is actuated by a neutral external agent or any kind of automatic device set for that purpose, it will be unfailingly enforced (provided that, whenever he punishes, the punisher gets a reward, say ε, from the guilty player). The corresponding game is the three-stage game described in Figure 5.9 in which we have assumed that the rich makes the first move and there are three players including the punishing agent. The latter must choose between two strategies: punish (denoted by P) and abstain from punishing (denoted by N).

Knowing that the punishing agent always punishes (he always plays P rather than N), the poor has an interest in co-operating if the rich has done so in the first round of the game whereas he is indifferent between co-operating and defecting if the rich has previously defected. Being aware of this optimal reaction of the poor conditional

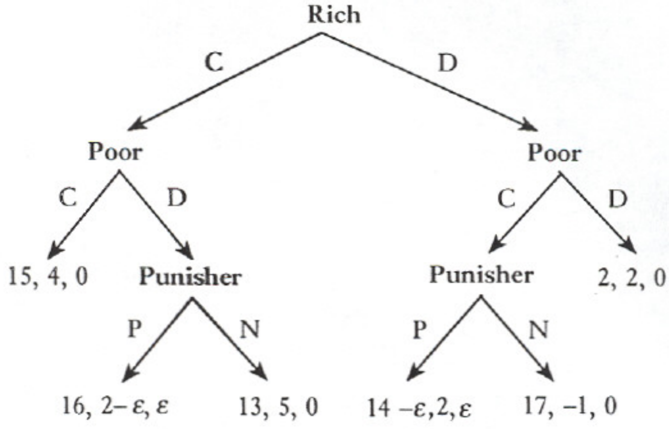

FIG. 5.9. A three-stage sequential asymmetrical game with a punishing player

on his own previous move, and knowing that he will be punished if the (D,C) outcome materializes, the rich will choose to co-operate, since he will thereby get 15 payoff units instead of either $(14 - \varepsilon)$ or 2 units. As a consequence, the co-operative outcome (C,C) is a subgame-perfect equilibrium.

Incentive dilution in a chicken game

As we have seen in the previous chapter, the size of the group may be an important determinant of the success of collective action. That the incentive for the members of a group to act in a collectively rational way tends to become thinner as the number of members increases is a proposition generally understood in the framework of the prisoner's dilemma. What we now want to argue is that, starting from a chicken game, the same dilution process may occur. As the size of the group increases, the structure of the game modifies itself from a chicken game into that of a prisoner's dilemma, and, as a result, the collectively rational outcome which was initially established degenerates into the Pareto-dominated non-co-operative outcome.

Let us illustrate this result with the help of a simple example. A group of n members can create a public good which yields a benefit of b to each of them at a collective cost of $c(n)$, where c is an increasing function of n, but $(c(n))/n$ is decreasing with n. Each member of the group is free to contribute or not. If he contributes, he shares the cost $c(n)$ with the m other willing contributors so that each of them incurs individually a cost of $(c(n))/(m + 1)$. If not, he simply free–rides on the public good produced by others, enjoying its benefits without incurring any cost. We assume that $b - (c(n))/n$ is always positive, so that creating the public good is collectively rational. However, we also assume that there exists n^*, with $n^* > 0$, such that, if $n < n^*$, $b - c(n) \geq 0$ and, if $n > n^*$, $b - c(n) \leq 0$. Figure 5.10 portrays the choice problem facing player i.

As can be readily seen from the figure, as long as n is relatively small ($n \leq n^*$ so that $b > c(n)$), the payoff structure given above defines an N-player version of a chicken game. However, when the size of the group increases, $c(n)$ rises up to a point where $b - c(n)$ becomes negative: the payoff structure of the above game turns into a PD and defecting becomes a dominant strategy. In other words, in the simple example described here, when the group is small, the cost of creating the public good is also small

Pay-off to player i if the number of other players contributing is

		$n-1$	$n-2$	$n-3$	. . .	0
Player i	contributes	$b-\dfrac{c(n)}{n}$	$b-\dfrac{c(n)}{n-1}$	$b-\dfrac{c(n)}{n-2}$		$b-c(n)$
	does not contribute	b	b	b		0

FIG. 5.10. Endogenous transformation of a chicken game into a PD game

so that everyone is ready to pay for it, even if he is alone to do so. However, as the size of the group increases, the cost involved also increases, while the individual benefits remain unchanged. Consequently, financing the public good on a voluntary basis becomes problematic when the group reaches a certain size. This is essentially the same argument as that put forward by Olson to underline the advantage of small ('privileged') over large groups, even though he presented it in the converse way (as the size of the group increases, the total cost of producing the public good is constant but the share of the benefits accruing to each individual declines).

5.2 Co-ordinated contributions

An important class of problems that arise in connection with the management of common property resources requires symmetric and co-ordinated actions to be overcome. Examples abound both in appropriation and in provision problems. For instance, in a fishery where the use of dynamite is an available technical option, it is obvious that self-restraint must be practised by everybody if the destruction of the fishing-ground is to be avoided. Protection of the breeding-grounds gives rise to the same problem. To quote examples from other sectors, restricted use of fire for the clearing of agricultural lands or management of water control infrastructures (including control of soil salinity and water-logging problems through sub-surface drainage) also require co-ordinated actions. Important issues of provision, such as steep-slope management and anti-erosion control in mountainous terrain, programmes of pest control, or certain surveillance actions requiring a critical amount of effort (e.g. guarding coastal fishing-grounds against the encroachments of mechanized boats) obviously belong to the above class of problems.

The one-shot assurance game

The game form suitable for representing this kind of situation is known as the assurance game (see Sen, 1967, 1973, 1985; Runge, 1981, 1984*b*, 1986; Dasgupta, 1988; Taylor, 1987; see also Ullmann-Margalit, 1977: 41; Collard, 1978: 12–13; 36–44, 80–9; Field, 1984: 699–700; Levi, 1988: ch. 3). In this game, a minimal effort must be contributed by all players if they are to receive any benefit from their own action.

To return to a familiar example, consider the case in which two fishermen must independently decide whether to put one or two boats at sea for the catching of fish. Let us assume that their payoffs for the various possible outcomes are as given in Figure 5.11.

The important point to note is that, contrary to what obtains in a PD game, the net payoff accruing to a player when he freerides on the public good provided by the other player (6 units) is smaller than the net payoff he would receive by co-operating (8 units). Nevertheless, if actors think it best to co-operate with each other, they still find it very unpleasant to be exploited by free–riders: contrary to what is observed in the chicken game, each player prefers mutual defection (where he gets a payoff of 2 units)

	no. of boats	Fisherman B: 1	Fisherman B: 2
Fisherman A	1	8, 8	1, 6
	2	6, 1	2, 2

FIG. 5.11. A fishing assurance game

to being a 'sucker' (which causes him to receive a payoff of only 1 unit). In short, universal co-operation is the most preferred outcome. Then comes generalized free-riding. Least preferred are those outcomes in which a mismatch of actions occurs. This payoff structure actually determines three possible equilibria, two in pure strategies—each fisherman puts out one boat or each fisherman puts out two boats—and one in mixed strategy. The Pareto-optimal outcome (each fisherman puts one boat out to sea), is only one of the two equilibria in pure strategies. Which equilibrium will be selected actually depends on prior expectations regarding the other's intended action.

Clearly, therefore, the best policy for each party depends on what he thinks the other will do. In actual fact, optimal choice for each fisherman is to put out only one boat if the probability that the other fisherman will choose the same strategy is assessed by him to be in excess of 1/3, and his optimal choice is to put out two boats if this probability is less than 1/3. Denoting by p the probability that the other fisherman puts out one boat, the value of 1/3 is obtained by solving the following equation:

$$8p + 1(1 - p) = 6p + 2(1 - p)$$

which establishes the condition for each fisherman to be indifferent between putting out one boat and putting out two boats to sea. (As is implicit from the above equation, the equilibrium in mixed strategy is such that each fisherman puts out one boat with probability 1/3 and two boats with probability 2/3.)

Thus, there is no certainty that the game will equilibrate at the more favourable of the three (Nash) equilibrium points. It is noteworthy, however, that players need not have complete assurance that others will also co-operate to adopt the same strategy: probabilities significantly smaller than 1 may provide sufficient incentive for co-operation. Still, the possibility exists that the worst equilibrium outcome will emerge *even though* the assumption of common knowledge implies that *each player knows that the other also prefers the co-operative outcome.*[2] This is because there is a genuine trust

[2] Curiously, Taylor rules out this possibility on the grounds that, since both players prefer the co-operative outcome to the mutual defection outcome, 'neither will expect the latter to be the outcome, so the unique Pareto-optimal outcome will result'. The assurance game is consequently deemed to be 'unproblematic' (Taylor, 1987: 39–40). For the reason explained in the text, Taylor's argument is unacceptable. In effect, it comes down to denying the fundamental fact that the 2 × 2 AG comprises three (Nash) equilibria.

problem, that is, a problem of assurance regarding the other person's intended action. Thus, A may know that B would prefer joint co-operation, yet he entertains the fear that B, even though he has corresponding knowledge about his own preference, will choose the maximin strategy ('defect') due to mistrust in what he will himself eventually decide to do. And B can reason in the same way with respect to A's presumed behaviour. *The trust problem is clearly reciprocal since it is basically a problem of mutual expectations*: A may fear that B will abstain from co-operating not because B prefers to free–ride but because B's expectations about his own (A's) behaviour may be pessimistic, and vice versa for B *vis-à-vis* A.

Now, if some form of rudimentary co-ordination device such as pre-play communication (say, in the form of 'cheap talk') is allowed and if the signals sent by the players are interpretable in an unambiguous way, co-operation or joint contribution by both players is much more likely to arise because the players then have the opportunity to reassure one another and to form optimistic expectations about their mutual behaviours.[3] What is worth emphasizing is that the nature of interactions in small groups is highly conducive to pre-play communication and, therefore, if both players' profile is that of an AG player, the Pareto-superior outcome is very likely to be established even in this one-shot game. (Remember that we have reached the same conclusion, *but for repeated games*, when we analysed situations structured as PD.)

Leadership in co-ordination problems

The uncertainty surrounding the players' decisions in a co-ordination problem is overcome as soon as either of the two players can take the initiative in the game with a view to signalling to the other his intention to co-operate. In game-theoretical terms, a particular way of representing the possibility of leadership is by specifying a two-stage assurance game. When the game is played in such a fashion, co-operation by both players is certain to occur: indeed, knowing that the other player will follow suit, the leader has an incentive to make a co-operative move. In other words, by co-operating in the first stage of the game, the leader does not incur the least risk of being 'exploited' by the follower. The outcome (co-operate, co-operate) is clearly a subgame-perfect equilibrium. This is illustrated in Figure 5.12 in which the same payoffs as those assumed in Figure 5.11 have been represented in an extensive form.

If a pure problem of distrust (such as is implicit in the assurance game) can be easily surmounted as soon as one of the players can send a signal or make a first move to the effect that he is determined to co-operate, then, *a fortiori*, the same problem is solved when the game can be repeated. Assume, for instance, that one of the players follows a cautious strategy (start by defecting and, thereafter, co-operating only if the other player has co-operated in the previous round). The other player's best reply to that strategy is obviously not to replicate it but, instead, to start by co-operating (say,

[3] Note that this is precisely the crucial role which Runge ascribes to institutions: to co-ordinate individuals' expectations so as to enable them to co-operate (Runge, 1981, 1984*a*, 1984*b*, 1986).

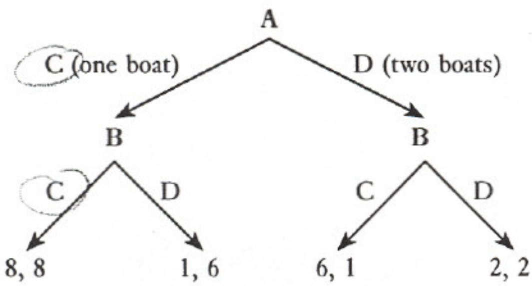

FIG. 5.12. A sequential assurance game

because he follows a strategy of unconditional co-operation) so as to trigger an uninterrupted chain of universal co-operation. Clearly, the cautious strategy is not a Nash equilibrium strategy. However, a 'bad' strategy such as one of unconditional defection is a best reply to itself and therefore supports a Nash equilibrium. (This obviously follows from the fact that, if AG players like best to co-operate, they do not want to be 'exploited'.) What needs to be stressed is that such a strategy is not subgame-perfect since, if by mistake a player co-operates, the other player's best response to that mistake is to co-operate, thus deviating from his Nash equilibrium path. To put it in another way, the commitment of one player to unconditional defection is not credible. Notice that the possibility of a co-operative outcome in such a repeated game is an application of the aforementioned folk theorem and its extension by Benoit and Krishna (1985). Just to give a simple example, a repeated assurance game underlies the observation that in lobster fisheries molesting another fisherman's trap is rarely done because by refraining from doing so a fisherman improves the chances that his own traps will not be molested (Sutinen, Rieser, and Gauvin, 1990: 341).

Threshold effects and freeriding in N*-player co-ordination games*

An interesting feature that arises in connection with co-ordination problems is the existence of threshold effects. As a matter of fact, in many cases, a collective action can bear fruit only if the number of contributors reaches a critical size. To analyse such kinds of situation, we need to consider an N-player assurance game. Let us assume that a given public good (say, the maintenance and management of an irrigation system) yields individual benefits to each member of a group equal to $b(m)$, where m stands for the number of voluntary contributors. Each contributor incurs a fixed cost of c units and, therefore, the total cost for the group is equal to $c \times m$. The choice facing player i can then be represented as in Figure 5.13.

First assume that both $(\partial b(m))/\partial m$ and $(\partial^2 b(m))/\partial m^2$ are positive, implying increasing returns to provision of the public good. Assume also that $b(1) - c < 0$, so that if no other player contributes to the public good, player i also chooses not to contribute. Yet, there exists a critical size m^* such that $b(m) - c > b(m^* - 1)$ or $c < b(m^*) - b(m^* - 1)$: once a certain number, m^*, of other players agree to contribute, player i

Pay-off to player i if the number of other players contributing is

		$n-1$	$n-2$	$n-3$	. . .	0
Player i	contributes	$b(n)-c$	$b(n-1)-c$	$b(n-2)-c$		$b(1)-c$
	does not contribute	$b(n-1)$	$b(n-2)$	$b(n-3)$		0

FIG. 5.13. A N-player assurance game

has an incentive to follow suit since the cost of individual contribution is less than the marginal individual benefit of that contribution. It is evident that, since $(\partial^2 b(m))/\partial m^2 > 0$, if $b(m^*) - c > b(m^* - 1)$, then $b(j) > b(j-1) + c$, $\forall j > m^*$. Therefore, as long as at least m^* other players contribute, player i prefers to co-operate rather than free ride.

In the above game, there are two Nash equilibria in pure strategies. The first equilibrium is characterized by universal defection: given that no one else contributes, player i has no incentive to undertake the collective action alone (we are therefore not in a chicken game). The second equilibrium is characterized by the fact that the collectively optimal level of the public good is provided: everybody contributes to that equilibrium. To avoid falling into the 'bad' equilibrium, a subgroup of players may decide to undertake the collective action in concert, regardless of what the others do. Here lies an important rationale for leadership and the function of the leader consists of mobilizing a sufficient number of contributors rather than showing the good example as assumed in the previous subsection.

It deserves to be noted that an interesting problem which can be raised within the framework of an N-player assurance game is actually a limit case of that analysed above, namely the case in which $b(j) = 0$, $\forall j < n$ and $b(n) > c$. In other words, the collective action can succeed or the public good can be provided only if everyone participates; if only a single agent defects, the public good disappears. The protection of an endangered species or of a breeding-ground illustrates such a possibility that perfectly fits with the description of what an assurance game is about.

Let us now consider the case where there are decreasing returns to scale in the provision of the public good: $(\partial^2 b(m))/\partial m^2$ is negative. In this case, there again exists a critical number of contributors, m^*, below which no individual player has any incentive to contribute. Yet, there now also exists an upper threshold number of contributors, say m^{**}, beyond which the individual marginal benefit of contributing falls short of cost c. The two Nash equilibria in pure strategies are easy to identify: the 'bad' equilibrium in which nobody contributes and a 'nice' equilibrium in which just m^{**} players contribute while the others defect. As long as the size of the group, n, is small (below m^{**}), everyone participates in the collective action under the 'nice' equilibrium. However, in large groups whose size exceeds the threshold m^{**}, the public good is only *partially* produced by a subgroup of players and the amount provided is not Pareto-optimal. It is actually less than the collectively rational amount which would require m^o contributors, with $m^o = \text{argmax}(nb(m^o) - m^o c)$. The collectively rational (co-operative) outcome requires that the collective marginal benefit is equal to the marginal cost c, that is $n(\partial b(m^o))/\partial m^o = c$. It is to be compared to the individually rational (Nash) outcome, m^{**}, which is by definition such that $(\partial b(m^{**}))/\partial m^{**} = c$. Bearing in mind the assumption of decreasing returns to public-good provision, it is evident that $m^o > m^{**}$.

In the latter circumstances (the group is large and $n > m^{**}$), a fraction of the players does not contribute in equilibrium and freeride on the others' efforts. Of course, the wider the gap between the size of the group, n, and the equilibrium threshold number of contributors, m^{**}, the larger the proportion of freeriders. In actual fact, the problem facing the players resembles that of an N-player chicken game, in which the Nash equilibrium would be suboptimal.

In community settings, a large proportion of such freeriders may cause serious tensions to arise. The community may possibly overcome these tensions, however. Thus, it may resort to a co-ordinated solution which has the effect of rotating over time the burden of contributions among the various agents. One option here is to use a correlated equilibrium solution in which contributors are selected through a lottery mechanism. It may also, at a given point of time, ensure that contributors with respect to a given collective action are allowed to abstain from participating in other collective actions so as to distribute equally the costs of public-good provision over a series of different activities. If the above kind of solutions are not applied, an exclusionary process is likely to ensue. This is apparently the case referred to by Ostrom and Gardner (1993) when analysing the Thambesi irrigation system in Nepal. Here, as pointed out earlier, maintenance of the headworks can be carried out by a limited number of the water users and, in particular, the work can be done by head-enders alone. The implication of this situation is that tail-enders may find themselves in a low bargaining position whenever important matters are to be discussed (Ostrom and Gardner, 1993: 97–9).

5.3 Heterogeneous Situations with PD, AG, and CG Players

In real-world settings, groups may not be homogeneous as we have assumed so far. This certainly applies to communities with respect to the management of local-level natural resources. It is indeed often observed that members of a particular user group behave differently because they do not derive the same benefits from a given action. This may be due to a variety of reasons, including differential endowments, different characteristics in terms of the technique employed and the pattern of use of the resource concerned (think of nomadic herders and sedentary agriculturalists), different social identities, different exit possibilities, varying perceptions of the stake involved in resource preservation, etc.

In game-theoretical terms, we will say that, in this case, encounters are heterogeneous in the sense that different *types* of players have to deal with each other. The type of a player is characterized by a particular payoff vector, which may be known or not by the other players. In the following, attention will be focused on heterogeneous games in which players with a payoff structure characteristic of the assurance game face players with a payoff structure characteristic of the prisoner's dilemma. These games are especially interesting because they portray a situation that has much relevance in many human encounters, namely that in which people who do not like to 'exploit' others meet with opportunists. The question that arises in such games is theoretically rich, in so far as it is not a priori clear who among the 'fair' players and the opportunists will determine the final outcome. Before turning to these games, however, mention will be made of two other kinds of heterogeneous encounters. First, we will consider a game in which the two players have an AG payoff structure, yet the benefits accruing to them are not identical. Second, a game in which a player with a chicken game (CG) structure encounters a player with an AG structure will be analysed.

Encounters between two different AG players

Let us assume that the two players who meet in an one-shot game have an AG payoff structure, implying that both of them have no incentive to free–ride on the other's efforts. However, player A has a greater interest in joint co-operation than player B, as illustrated in Figure 5.14.

As usual, there are three Nash equilibria: (C,C), (D,D), and the mixed strategy in which the probability that A plays C is equal to 1/2 and the probability that B plays C is 1/4. Which of these equilibria will emerge depends on the expectations that the players hold about the likelihood that the other co-operates. Assuming that they both hold the same expectation, p, both players co-operate if $p > 1/2$ and defect if $p < 1/4$. Clearly, there exists a range, $1/4 < p < 1/2$, in which A co-operates while B abstains from doing so. Such an outcome, however, is not an equilibrium (player A will not accept to be 'exploited' by player B). If it may arise, it is actually because there exists an inverse relationship between the size of the payoff accruing to the player in case of joint co-operation and the degree of trust required to prompt the player to co-operate.

		Player B	
		C	D
Player A	C	8, 6	1, 5
	D	5, 1	2, 2

FIG. 5.14. A 2 × 2 asymmetrical assurance game

		AG player C	AG player D
CG player	C	10, 4	8, 4
	D	12, 0	0, 2

FIG. 5.15. A CG player meets an AG player

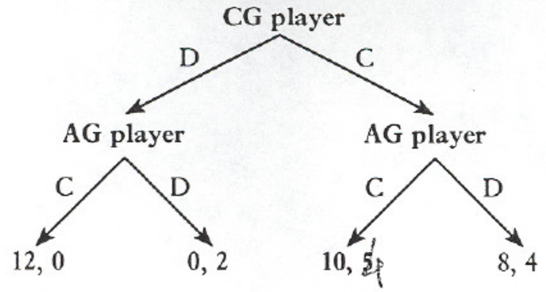

FIG. 5.16. A CG player meets an AG player and leads the sequential game

Encounters between AG and CG players

An interesting situation emerges when an AG player faces a CG player, since it allows us to realize the importance of leadership in determining the equilibrium outcome. To start with, consider the one-shot game with simultaneous moves represented in Figure 5.15.

It is easy to check that, in this game, no equilibrium in pure strategy exists. There is only one equilibrium in mixed strategy, where the probability of the CG player playing C is 2/3 and the probability of the AG player playing C is 4/5.

Equilibria in pure strategies are nevertheless possible as soon as the game is played sequentially. Yet, which equilibrium will arise depends on which player is in the first-mover position. Assume first that the CG player is the leader, as in the sequential game described in Figure 5.16.

It is immediately apparent from Figure 5.16 that joint co-operation will occur: it is in the interest of the CG player to start by co-operating so as to induce the AC player to follow suit. Indeed, if the CG player makes a non-co-operative first move, he is sure to bring about a situation of mutual defection, which he wants absolutely to avoid. In other words, when a party with a leadership role is keen that a collective action is undertaken, but preferably not by himself, whereas the other party tends to follow the leader's behaviour, but prefers bilateral co-operation to bilateral free-riding, joint co-operation will be established. This happy outcome entirely depends on the fact that the leader has a CG payoff structure. Indeed, had the leadership roles been inverted

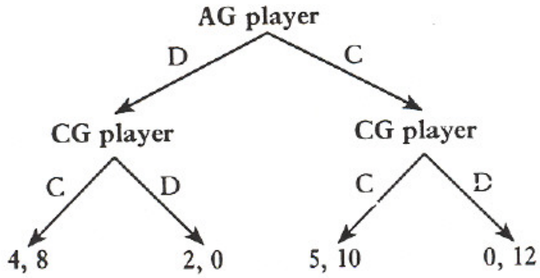

FIG. 5.17. An AG player meets a CG player and leads the game

(the leader is the AG player and the follower is the CG player), universal co-operation would be prevented from arising. This is a straightforward conclusion from Figure 5.17 where the game is led by the AG player.

As is evident from the figure, the collective action will also be undertaken but only by one of the players: having the right to the initial move, the AG player uses the advantage of knowing that the other player has a CG payoff structure to force him to bear the whole cost of this action. Notice carefully, however, that the AG player as a leader is unable to bring about the outcome which he best prefers (universal co-operation) since by co-operating he would incite the follower to defect. This frustration would not have occurred if both players had a CG payoff structure: forcing the other player to co-operate by defecting in the first place is then the ideal outcome which each player wishes for.

An interesting feature which emerges from any encounter between an AG player and a CG player is that both players have an interest in granting leadership to the latter: indeed, the outcome of the first game (Figure 5.16) dominates the outcome of the second game (Figure 5.17). This means that, in a more complex game in which the players would be invited to select the leader before deciding sequentially whether to co-operate or not, the unique subgame-perfect equilibrium path is as follows: the players select the CG player as the leader, thereafter this player co-operates and, in the final stage, the AG player responds by co-operating too. The lesson from such a three-stage game is that, by binding himself to the leadership position, the CG player commits himself to co-operation.

An example which illustrates the aforedescribed situation can again be borrowed from studies of irrigation management. Consider once more a situation in which water users are divided into two subgroups according to whether they are head-enders or tail-enders. Head-enders have a CG payoff structure since they are keen that maintenance of the water control infrastructure is undertaken, but would very much prefer that tail-enders do the work alone (something which may be technically possible, as we have pointed out in the case of the Thambesi irrigation system). On the other hand, tail-enders who are at a locational disadvantage entertain the fear that they may be excluded from decision processes that affect the flow of water reaching their fields (see above): this is why they are eager to participate in maintenance works alongside head-enders, yet would not like to be 'suckers' if head-enders refrain from such partici-

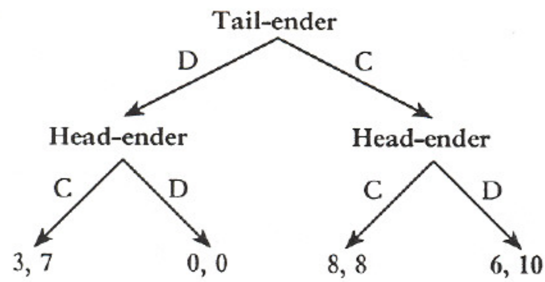

FIG. 5.18. Head-enders deal with tail-enders in a CG-AG encounter

pation (they have an AG payoff structure). In these circumstances, as argued above, head-enders have an incentive to take the leadership so as to associate tail-enders in the maintenance works. This situation seems to fit rather well with the experience of the Pithuwa irrigation project reviewed by Ostrom and Gardner (1993: 105–6).

Another, more plausible situation arises when tail-enders have a mixed CG–AG payoff structure in the following sense: if head-enders participate in the maintenance works, they want to join in order to avoid the aforementioned negative spillover effect but, if head-enders abstain from such participation, they prefer to undertake these works alone rather than leaving the system to fall into decay. If, moreover, as assumed in Figure 5.18, the tail-enders' payoff is higher when they are 'suckers' (6 units) than when they free–ride on the head-enders' maintenance efforts (3 units)—because the price to be paid in terms of loss of reliable access to water is high when participation in these efforts is shunned—the following result obtains: whether the game is played simultaneously or sequentially and whether, in the last case, leadership is exercised by tail-enders or head-enders, the equilibrium outcome is characterized by participation on the part of tail-enders and defection on the part of head-enders. In other words, even though they have the first move, the tail-enders are unable to take advantage of the chicken game structure of the head-ender' payoffs to make them co-operate, owing to the latter's critical control over the supply of water.

Note that, if the tail-ender's payoff when they are 'suckers' is 2 units instead of 6 units or, if their payoff while freeriding on the head-enders' efforts is 6.5 instead of 3 units, it is easy to see that what has just been said is no longer true: if they have the first move, they are now in a position to force head-enders to undertake (alone) the maintenance works. In other words, if the cost of free–riding on maintenance efforts in terms of loss of reliable access to water is not too high for the tail-enders and if they hold a leadership position (two rather implausible assumptions), their leverage allows them to impose the cost of maintenance on the head-enders.

Encounters between AG and PD players in small groups

Let us now turn to the important set of situations in which AG players encounter PD players. There are numerous relevant cases which are worth considering. In order to

		PD player	
		C	D
AG player	C	2, 2	1, 5
	D	1, –1	0, 0

FIG. 5.19. A PD player meets an AG player

help the reader to follow the arguments, these cases are presented in increasing order of complexity.

1. To start with, let us examine the simple one-shot two-player game with perfect information in which an AG player meets a PD player. That in this case co-operation cannot occur is immediately evident from the payoff matrix depicted in Figure 5.19.

The non-co-operative outcome (D,D) is the only Nash equilibrium in this game. Note that defection is a dominant strategy for the PD player as a result of which there is no equilibrium in mixed strategy.

2. What would happen if such a game were to be played in a sequential manner? The answer to that question is rather straightforward. If the AG player is in the first-mover position, he will be prompted to defect since he anticipates that the PD player defects in any event. In contrast, and rather unexpectedly, the reverse outcome obtains in the case where the PD player takes the lead: as a matter of fact, knowing that his opponent responds to co-operation by co-operation and to defection by defection, the PD leader has an incentive to start by co-operating. Because (D,D) is the subgame-perfect equilibrium of the former sequential game and (C,C) that of the latter, were both players allowed to choose their leader, both of them would concur in selecting the PD player. Note the similarity between this conclusion and that reached when analysing CG–AG player encounters.

3. If the game is (finitely) repeated rather than being played sequentially, mutual defection appears as the only possible equilibrium as observed in the finitely repeated PD game. Reasoning by backwards induction makes this result clear. As a matter of fact, since the AG player knows that his opponent has a PD payoff structure, he can infer that the latter will surely defect in the last round of the game. His best reply is therefore also to defect in this last round. Being aware that the AG player is going to defect, the PD player has no incentive to build a reputation of 'co-operator' in the round (t – 1) and, as a consequence, he will also defect in that round. Since he knows that things will turn out that way, the AG player defects in the same round too. This reasoning can be pursued backwards till the very first round of the game.

4. If the game is played in infinite time (or if its length is finite but indeterminate), the folk theorem applies and joint co-operation is a possible (subgame-perfect) equilibrium.

AG against PD player

	C	D
C	2, 2	−1, 3
D	1, −1	0, 0

AG against AG player

	C	D
C	2, 2	−1, 1
D	1, −1	0, 0

FIG. 5.20. A 2 × 2 game with one-sided asymmetric information

5. By relaxing the assumption of perfect information, we can raise the question as to whether co-operation becomes a possible equilibrium outcome in a finitely repeated game. This question is worth raising bearing in mind the result achieved by Kreps and his associates in the framework of a finitely repeated PD game (see above). Let us first consider the case where uncertainty about the payoff structure of the other player is one-sided asymmetric. More precisely, we assume the first player has a PD payoff structure, and this is common knowledge, while the type of the second player is not known with certainty: the first player assigns probability p to the possibility of the second player being an AG player and probability $(l - p)$ to the possibility that he is a PD player. We know from Kreps *et al.* (1982: 251) that in these circumstances defection will occur throughout the whole game: indeed, if the other player is a PD player, we know that defection is the only possible equilibrium and we have just shown that, if a PD player meets repeatedly an AG player, defection also occurs. There is therefore no reason to expect that a co-operative equilibrium can be generated when there is a doubt about whether the second player has an AG or a PD payoff structure.

6. The next case to consider is that in which the first player is an AG player, and this is common knowledge, while there is doubt about the payoff structure of the second player. Is co-operation more likely to emerge in those more favourable conditions? The answer is a conditional 'yes'. More precisely, co-operation by both players till the last stage of the game may occur if the expectation held by the player with a certain AG payoff structure that the opponent is also an AG player exceeds a certain level. In case this expectation falls below that level, universal defection occurs throughout the game. To see this, consider a three-period game in standard form in which the Row player is known with certainty to be an AG player while there is doubt about whether the Column player is a PD or an AG player. The payoffs pertaining to the two possible kinds of encounters are given in Figure 5.20.

Thus, for instance, by defecting while his opponent co-operates, the PD player gets a payoff of 3 which is more than his payoff when both co-operate. By contrast, in the same circumstances, the AG player gets a payoff of only 1 which is less than his payoff when both co-operate. Let us now assume that an AG player follows the brave reciprocity strategy consisting of starting by co-operating and, thereafter, co-operating as long as the opponent is in 'good standing'. As for the PD player, he starts by co-operating, thereafter mimics what the opponent has done in the previous round till the last round where he defects. The question then is whether these two strategies are the best replies to one another, which would imply that mutual co-operation starts from the first round and continues till the last round when the Column player defects while the Row player co-operates till the very end of the game.

To begin with, consider which payoffs the Column player would earn, *if he is of the AG type*, by following various possible strategies in his encounters with the Row player who is known to be of the AG type and follows the aforedescribed strategy. If he plays a strategy of brave reciprocity (note that similar strategies such as tit for tat or unconditional co-operation also lead to the (C,C,C) sequence of actions), he gets a total payoff over the three periods equal to 6 (2 + 2 + 2). If he plays a strategy whereby, against the brave reciprocity strategy of Row, he co-operates in the first two rounds and defects in the last one, his payoff amounts to 5 (2 + 2 + 1). If he co-operates in the first round and defects in the last two rounds, he earns 3 (2 + 1 + 0) while, if he defects from beginning to end, he earns only 1 (1 + 0 + 0). It is therefore evident that unconditional co-operation dominates the other three strategies. In other words, if Column is of the AG type and Row is of the same type and follows a strategy which consists of starting by co-operating and, thereafter, mimicking what the opponent has done in the previous round, then Column's best reply to the latter is to co-operate throughout the whole game. Since the Row player has adopted the above strategy, he will also co-operate from the beginning to the end of the game. Mutual co-operation therefore occurs till the game ends. There is actually nothing surprising in this result which has already been accounted for at an earlier stage of our analysis.

If Column is of the PD type, his total payoffs while playing various possible strategies against Row are as follows: 6 (2 + 2 + 2) if he plays a co-operative strategy leading to the (C,C,C) sequence of actions; 7 (2 + 2 + 3) if he plays the *fake strategy* whereby he co-operates as his opponent co-operates, and defects in the last round, then revealing his true type; 5 (2 + 3 + 0) if his strategy leads to a (C,D,D) sequence of moves; and 3 (3 + 0 + 0) if he plays unconditional defection. The fake strategy dominates the other strategies available to Column.

Let us now turn to the Row player in order to check whether the brave reciprocity strategy can be a best reply to brave reciprocity (or similar strategies) played by Column if of the AG type and to the fake strategy if of the PD type. Let p be the probability Row assigns to the possibility that Column is of the AG type and $(1-p)$ the probability that Column is of the PD type. If Row plays brave reciprocity, his payoff is 2 for the first period, 2 again for the second period, and $(p \times 2 + (1-p)(-1))$ for the third period, amounting to a total payoff of $3p + 3$ for the three periods together. If, instead, he plays a *safe strategy* whereby he follows the strategy of brave reciprocity except in the last round when he ensures himself against being a 'sucker' by defecting, his payoff is $2 + 2 + (p \times 1 + (1-p)0) = p + 4$. If he plays a strategy leading to the (C,D,D) sequence of actions, he earns $2 + 1 + 0 = 3$ while, if he plays unconditional defection, he gets only 1 $(1 + 0 + 0)$. Clearly, the latter two strategies are dominated by the first two. Whether the first strategy dominates the second strategy or is dominated by it hinges upon the value of p, that is, upon the expectation of Row regarding the payoff identity of Column. In more exact terms, brave reciprocity dominates the safe strategy if $p > 1/2$.

Row will therefore co-operate all throughout the game if he believes there is more than 50 per cent chance that Column is of the same type as himself, otherwise he will stop co-operating after two periods. Since p is common knowledge—Column knows Row's expectation regarding his own (Column's) payoff structure—if $p > 1/2$, Column will continue to co-operate till the very end of the game in case he is of the AG type, and till the last period in case he is of the PD type. If $p < 1/2$, on the other hand, Column knows that Row will defect in the last (third) round of the game and, as a result, he has no incentive to co-operate in the second round to maintain a reputation of 'co-operator'. Applying the argument backwards, it is easy to see that, in these conditions, co-operation unravels and universal defection occurs from beginning to end. To sum up, either Row's expectation regarding the chance that Column is of the AG type is sufficiently high, and universal co-operation is sure to occur till at least the last stage of the game, or this expectation is too low and universal defection occurs throughout the game.

Clearly, co-operation is not doomed to failure because groups are heterogeneous in the sense that there is a non-negligible proportion of potential opportunists. As we have seen above and will continue to see in the three following points, co-operation is a serious possibility when expectations are favourable to it.

7. In the two foregoing points, we have only considered situations of one-sided asymmetric information. It is tempting to examine now whether co-operation is a possible outcome when the imperfection of information is two-sided, that is, when the two players entertain mutual doubts about their respective payoff structure. An important—but largely neglected (see, however, Gibbons, 1992: 226)—result obtained by Kreps and his associates in their aforementioned, celebrated article (1982) is that extension of uncertainty about payoffs to the two players may increase the chance of co-operation. Remember that, as seen under point 4 above, co-operation is impossible when one player is of the PD type and doubts whether the other player is of the AG or the PD type. What Kreps *et al.* show, however, is that when the two players are of the PD type but believe that their opponent might perhaps be of the AG type, there can exist an equilibrium in which both players co-operate until the last few stages of the game (the end-game is rather complex). Yet, it deserves to be emphasized that this game admits (subgame-perfect Nash) equilibria in which long-run co-operation does not ensue. Co-operation actually requires a 'boot-strapping' operation (since there is obviously a trust problem): even if each side is certain that the other has an AG payoff structure, co-operation ensues only if each side hypothesizes that the other side will co-operate (Kreps *et al.*, 1982: 251).

To see this possibility of co-operation when there is two-sided uncertainty about payoff structures, let us again use our simple three-period framework. Pay-offs are assumed to be the same as in Figure 5.19. In the mind of Row, Column might be of the AG, rather than PD, type, an eventuality to which he assigns a probability p. On the other hand, Column entertains the hypothesis that Row is an AG player (with probability q) rather than a PD player (with probability $(1 - q)$). What we want to show is whether and under which conditions the two aforeprescribed strategies ('start by co-operating and thereafter mimic what the opponent has done in the previous round', till the last stage of the game for the PD player and till the end of the game for the AG player) can be best replies to each other.

In actual fact, part of the preparatory work required to answer that question has already been done in the previous point while considering the decision problem of Row. Bear in mind, indeed, that Row's best strategies, *when he is of the AG type*, are a strategy of brave reciprocity, which yields him a total payoff of $3p + 3$ over the three periods, and the safe strategy, which yields a payoff of $p + 4$. The former strategy dominates the latter if $p > 1/2$. *When Row is of the PD type*, on the other hand, his payoffs are as follows:

$2 + 2 + [p \times 2\ (1 - p)(-1)\ ] = 3p + 3,$	if he plays brave reciprocity;
$2 + 2 + [p \times 3 + (1 - p)0] = 3p + 4,$	if he plays the fake strategy;
$2 + 3 + 0 = 5,$	if he plays the (C,D,D) sequence of moves;
$3 + 0 + 0 = 3,$	if he plays unconditional defection.

The strategies of unconditional defection and of brave reciprocity are clearly dominated. Whether the fake strategy is superior to the other (which leads to the (C,D,D) sequence of moves) depends on the value of p: the former dominates if $p > 1/3$. It is therefore apparent that, if Row expects with a probability higher than 1/3 that Column is of the AG type, he will co-operate till, at least, the last stage of the game. If this probability is higher than 1/2 and he is himself of the AG type, Row will even co-operate till the end of the game.

Exactly the same reasoning can be made with respect to Column. If Column expects with a probability higher than 1/3 ($q > 1/3$) that Row is of the AG type, he has an incentive to co-operate, at least till the last round of the game. We can conclude that, if expectations of both players regarding the chance that the opponent is of the AG type exceed 1/3, co-operation till at least the last stage of the game is an equilibrium outcome. If this expectation is higher than 1/2, both Row and Column will co-operate till the end of the game provided that they are of the AG type. If, say, the expectation of one player is more pessimistic and falls below the threshold level of 1/3, this is sufficient to destroy co-operation. Indeed, the opponent then knows that the pessimistic player is going to defect from as early as the second round—since the (C,D,D) sequence of moves then dominates the fake strategy—and, therefore, he himself has no incentive to co-operate in the second round nor actually in the first round (since it is of no use for him to build up a reputation of 'co-operator'). The pessimistic player, aware of this calculation made by his opponent, will also defect in the initial round. Universal defection occurs throughout the game.

8. We will now extend the above analysis to games with many players. To keep things as simple as possible, consider a three-player game that is played over only two periods. The three players are uncertain about the payoff structure of the other two

		C,C[1]	C,D[2]	D,D[3]
Player 3	co-operates	**4** (4)	**3** (1)	**–1** (–1)
	defects	**2** (5)	**1** (3)	**0** (0)

FIG. 5.21. A three-player game with asymmetric information

players; more specifically they entertain doubts about whether the other players are of the AG or PD type.

Let us consider the decision problem faced by player 3 as it is depicted in Figure 5.21.

All players have a probability q of being AG and a probability $(1 - q)$ of being PD. If player 3 is of the AG type, he faces the payoff numbers written in bold characters. For instance, if he defects while at least another player co-operates, he is less well-off than if he co-operates. If both other players defect, he prefers to defect too because he does not want to be a 'sucker'. In contrast, if player 3 is of the PD type, his payoffs are those indicated between brackets: defection is then a dominant strategy.

(i) Let us first assume that, if of the AG type, a player adopts a *strategy of harsh punishment*. In this case, he starts by co-operating and thereafter defects if one of the other two players defected in the previous round. Otherwise, he co-operates. Now, if a player is of the PD type, he follows a *fake strategy* (he mimics being an AG player by co-operating in the first round, and continues to co-operate as long as the other two players co-operate till the last round when he defects). The question is: are these two strategies best replies to one another?

To proceed with the analysis, we begin by examining the situation in which player 3 is of the AG type. If he plays the harsh punishment strategy, his total payoff over the two periods is:

$$\underbrace{4}_{\text{period 1}} + \underbrace{[q^2 4 + q(1-q)3 + (1-q)q3 + (1-q)^2(-1)]}_{\text{period 2}} = -3q^2 + 8q + 3$$

This payoff is obviously identical to that which he would obtain were he to follow either an *unconditional co-operation strategy* or a *soft-punishment strategy*, since, in actual fact, he cannot know the other players' types by observing their first period's moves (since PD players fake till the last round). By *soft-punishment strategy*, we mean a strategy whereby he continues to co-operate as long as at least one other player has co-operated in the previous round (or, to put it in another way, he defects only if all other players have defected). We will return later to this particular strategy. To counter the difficulty that he will know the other players' types only in the last round, player 3 may choose to play a *safe strategy* (he starts by co-operating and defects in the last round):

$$\underbrace{4}_{period\ 1} + \underbrace{[q^2 2 + q(1-q) + (1-q)q + (1-q)^2(0)]}_{period\ 2} = 2q + 4$$

If he plays other strategies (implying such sequences of actions as (D,D) or (D,C)), the payoffs will obviously be lower than when he plays the above two strategies. Whether the harsh-punishment strategy yields a higher payoff than the safe strategy obviously depends on the value of the probability q. More specifically, the former is superior to the latter if: $-3q^2 + 6q - 1 > 0$, implying that $q > 0.18$.

Consider now the alternative situation in which player 3 is of the PD type. If he plays the fake strategy, he gets the following total payoff:

$$\underbrace{4}_{period\ 1} + \underbrace{[q^2 5 + q(1-q)3 + (1-q)q3 + (1-q)^2(0)]}_{period\ 2} = -q^2 + 6q + 4$$

If he, instead, plays *unconditional defection strategy*, he gets:

$$\underbrace{5}_{period\ 1} + \underbrace{[q^2(0) + q(1-q)(0) + (1-q)q(0) + (1-q)^2(0)]}_{period\ 2} = 5$$

If he plays other strategies, implying in particular a co-operative move in the last round, his payoffs will obviously be lower than when he plays the above two strategies. Moreover, the fake strategy dominates unconditional defection if $-q^2 + 6q - 1 > 0$, implying that $q > 0.17$.

Note carefully that the critical value of q that induces an AG player to reject the safe strategy is actually greater than the value required to prompt a PD player to use the fake strategy, thereby making the latter condition redundant. We can therefore conclude that, if q, the probability that a player is of the AG type, is greater than 0.18, then the best reply to the harsh-punishment strategy adopted by AG players is faking for the PD player, and vice versa. This is an important result in so far as it shows that, even if in the one-period game the dominant strategy of a PD player is to defect, he may have an incentive, in a two-period game, to behave 'co-operatively', as though he were an AG player, till the second round of the game. This result can be extended to more periods: if his expectation that the other players are of the AG type is sufficiently high, the PD player has an incentive to start by co-operating and thereafter continue to co-operate as long as these other players co-operate, till the last round of the game when he defects. It is noteworthy that the critical values of q obtained in games that stretch over, say, three periods are precisely the same as those obtained in the two-period case. Finally, it should be emphasized that, as the above example shows, the critical values of q need not be very high. This obviously hinges upon the fact that, in this example, defection is not very rewarding for a PD player.

(ii) Let us now investigate the possibility of the AG players adopting a soft punishment strategy. In these circumstances, the PD players know that their defection may not necessarily be retaliated in the next round by a non-co-operative move of the AG players. This obviously depends on what the other PD players choose to do. Consider first the decision problem faced by player 3 if he is of the AG type. For a reason explained above, when opposed to a fake strategy, the payoffs associated with different strategies are exactly the same as those obtained under a harsh-punishment strategy. In particular, the soft-punishment strategy is superior to the safe strategy if $q > 0.18$. If player 3 is, instead, of the PD type, the fake strategy yields the following total payoff:

$$\underbrace{4}_{period\ 1} + \underbrace{[q^2 5 + q(1-q)3 + (1-q)q3 + (1-q)^2(0)]}_{period\ 2} = -q^2 + 6q + 4$$

The payoff resulting from unconditional defection is:

$$\underbrace{5}_{period\ 1} + \underbrace{[q^2 5 + q(1-q)3 + (1-q)q3 + (1-q)^2(0)]}_{period\ 2} = -q^2 + 6q + 5$$

Strategies that imply a co-operative move in the last round are clearly inferior. It is immediately apparent that playing unconditional defection is always more rewarding than playing the fake strategy. As a result, with a soft-punishment strategy, it is impossible that all types of players *always* co-operate in the first round. In the above, we have assumed that the other players, if of the PD type, start by co-operating and defect in the second round. We now have to check whether this is really the most sensible strategy for such players given that the third player, when PD, replies by always defecting. To carry out this check, let us examine whether unconditional defection is the best strategy for *all* the PD players simultaneously. The payoff obtained by player 3 when he always defects against the other players who, if of the PD type, are also unconditional defectors, is the following:

$$\underbrace{q^2(5+5)}_{2\ AG} + \underbrace{(1-q)^2(0+0)}_{2\ PD} + \underbrace{2q(1-q)(3+0)}_{1\ AG\ and\ 1\ PD} = 4q^2 + 6q$$

If, instead, he plays the fake strategy, he gets:

$$\underbrace{q^2(4+5)}_{2\ AG} + \underbrace{(1-q)^2(-1+0)}_{2\ PD} + \underbrace{2q(1-q)(1+3)}_{1\ AG\ and\ 1\ PD} = 10q - 1$$

As can easily by seen, unconditional defection always dominates the fake strategy. This, however, is a result that pertains to a border case since the quadratic equation, $4q^2 + 6q = 10q - 1$, has a unique root equal to 0.5. When q is just equal to 50 per cent, player 3 is thus indifferent between the two strategies whereas, for all other values of q, he prefers unconditional defection. By altering the payoffs given in Figure 5.21, it is possible to construct a more general case in which the fake strategy is the best reply of the third player, if PD, to unconditional defection by other PD players and the soft-punishment strategy by the AG players. (Presumably, there is an interval for q such that PD players will adopt a mixed strategy which consists of randomizing between the fake strategy and unconditional defection and such that AG players prefer soft punishment to harsh punishment.) To conclude the analysis based on the payoff matrix given in Figure 5.21, there still remains the question as to whether the soft-punishment strategy is the best reply of an AG player to the unconditional defection strategy adopted by the PD players. To see this, let us consider the payoffs which would accrue to an AG player when he, alternatively, chooses to play soft punishment, harsh punishment, or a strategy of cautious reciprocity (start by defecting and co-operate only if at least one other player has co-operated in the first round). The payoffs associated with these strategies are, respectively:

$$\underbrace{q^2(4+4)}_{2\text{ AG}} + \underbrace{(1-q)^2(-1+0)}_{2\text{ PD}} + \underbrace{2q(1-q)(3+3)}_{1\text{ AG and }1\text{ PD}} = -5q^2 + 14q - 1,$$

$$\underbrace{q^2(4+4)}_{2\text{ AG}} + \underbrace{(1-q)^2(-1+0)}_{2\text{ PD}} + \underbrace{2q(1-q)(3+1)}_{1\text{ AG and }1\text{ PD}} = -q^2 + 10q - 1,$$

$$\underbrace{q^2(2+4)}_{2\text{ AG}} + \underbrace{(1-q)^2(0+0)}_{2\text{ PD}} + \underbrace{2q(1-q)(1+(-1))}_{1\text{ AG and }1\text{ PD}} = 6q^2$$

From a comparison of the above payoffs, it is evident that the harsh-punishment strategy is dominated by the soft-punishment strategy. On the other hand, the strategy of cautious reciprocity is superior to the latter when the probability of meeting AG players is very low (below 0.08 approximately).

To conclude, there are plausible conditions, implying a sufficient probability of meeting other players of the AG type, under which AG players follow a strategy of soft punishment while PD players unconditionally defect.

Encounters between AG and PD players in large groups

Let us turn to another type of situation where the number of players is significantly large. In such a situation, members meet anonymously, they cannot remember the exact course of actions followed in the past by any particular player, yet past aggregate outcomes are observable and remembered. In these conditions, agents have no incentive to build up a 'good' reputation and, therefore, to play strategically has not the same meaning as when group size is restricted. To proceed with the analysis of such games, let us first consider the payoff matrices described in Figure 5.22: the first one gives the benefits accruing to an AG player, when the proportion of players who co-operate varies from 0 to 100 per cent while the second one gives the benefits accruing to a PD player in the same circumstances.

The argument behind this example is the following. In an N-person game, the gains from co-operation and defection for each actor obviously depend on the proportion of people who actually co-operate (or defect) in the entire group. The gains which both AG- and PD-type players derive from co-operation decrease when the proportion of co-operating members in the group declines. Yet such gains are higher for AG players than for PD players for any given proportion of co-operators in the group. On the contrary, the gains from defection are always smaller for AG players than for PD

Pay-offs for a AG-player	Proportion of co-operators in the group 100%	80%	60%	40%	20%	0%
C	20	13	6	−1	−8	−15
D	6	6	6	6	6	6

Pay-offs for a PD-player	Proportion of co-operators in the group 100%	80%	60%	40%	20%	0%
C	10	4	−2	−8	−14	−20
D	30	28	25	21	16	8

FIG. 5.22. Payoffs to AG- and PD-type players according to the proportion of co-operators in a large group

players. Moreover, the latter's gains from defection have a tendency to diminish with the proportion of co-operators in the group: it is more rewarding to free–ride when everyone else co-operates than when only a fraction of the other members co-operate, and the gains from free–riding are at their lowest when defection is generalized.

By contrast, the gains from defection accruing to AG players exhibit a constant pattern even when the proportion of co-operators in the group decreases. This is because two opposite effects are at work when these players defect. On the one hand, there is the above-noted fact that defection is all the less rewarding as the percentage of free–riders in the population increases. But, on the other hand, AG players 'feel bad' about defecting, especially so if they are amidst a large number of co-operating people. Or, to put it in the converse way, the higher the proportion of free–riders in the group, the more they are relieved of their 'bad feelings' since they can justify their 'opportunistic' acts by reference to the fact that many others behave in the same way as they do. Consequently, the net effect of an increase in the proportion of free–riders on the utility payoffs accruing to AG players when they defect cannot be determined on an a priori basis. Here, we have assumed that the two effects exactly counterbalance each other so that these payoffs are left unaffected by changes in the percentage of freeriders in the group.

Furthermore, it is worthy of note that the payoff to AG players when they co-operate and everybody else also co-operates (or when more than 60 per cent of all members co-operate) is higher than the payoff they receive when they are the only ones to defect in the group (6 units): this is a typical reflection of an AG-preference structure. The opposite is of course true of PD players who receive higher payoffs by defecting than by co-operating not only when all other members or a majority of them co-operate but also when few others or even nobody in the group co-operates. Another noteworthy feature is that the payoff received by PD players when they freeride jointly with everybody else (8 units) is smaller than that which they obtain by co-operating jointly with everybody else (10 units), a feature characteristic of a PD game. This, of course, holds *a fortiori* true for AG players.

It is immediately apparent from Figure 5.22 that PD players have a dominant strategy which is to defect. As for AG players, their preferred strategy will obviously depend on their expectations regarding the likely behaviour of the other players. They will choose to co-operate if they expect more than 60 per cent of the group members to co-operate, otherwise they will defect. Thus, for example, if AG players assess the proportion of co-operators in the group to be around one-half, generalized freeriding will take place as both types of players choose to defect. In this kind of situation, the meaningfulness of the concept of trust is evident. In the words of Dasgupta, trust here is to be understood 'in the sense of correct expectations about the *actions* of other people that have a bearing on one's own choice of action when that action must be chosen before one can *monitor* the actions of those others' (Dasgupta, 1988: 51).

The main conclusion that emerges from the above N-players game at this stage is the following: for co-operation to prevail on a large scale in an anonymous society or in a large group, it is not sufficient that a significant majority of people prefer universal co-operation but it must also be the case that these people feel confident enough that their willingness to co-operate is shared by many others too.

Now the question is not only how, or under what conditions, collective action can occur in a large group with the characteristics considered here; the question is also whether the co-operative outcome can be sustained on a large enough scale over time. To answer this last question, more information is needed about the dynamics of expectation formation. In a dynamic setting, indeed, decision by AG players whether or not to co-operate requires continual re-evaluation of the probability that others will also co-operate based on concrete experiences in past rounds. Not only do expectations affect co-operative behaviour but, over time, past co-operative outcomes affect expectations and future actions, though in a way that leaves no room for strategic considerations: a single player's co-operation cannot affect the proportion of co-operators in the group.

In accordance with what has been said above about the observability of past aggregate outcomes, the assumption is made that agents are broadly able to make out *ex post* whether and to what extent the collective action under concern has been successful. This is because they can observe the concrete results that collective action has produced: an irrigation canal has been more or less well maintained; foreign trawlers have been effectively deprived of access to inshore waters; the spawning area for fish has not been encroached upon; no felling of trees or cutting of wood has happened in the forest during forbidden times; little grazing occurred on the collective fields before the date fixed, etc. As is evident from these illustrations, the members of a large group may even be in a position to *approximately* assess the relative number of individuals who have co-operated or defected (yet they are not able to personally identify them).

Let us adopt the following conventions:

P^{AG}	denotes the proportion of AG players in the group;
$P^{PD} = 1 - P^{AG}$	denotes the proportion of PD players in the group;
P^*	denotes the minimum proportion of co-operators required to induce co-operative behaviour among AG players;
P^e_t	denotes the proportion of co-operators whom AG players expect to be present in the group at time $t + 1$; (P^e_0 is therefore the *initial* expectation of AG players which reflects their beliefs about the percentage of group members who will co-operate in the *first* round of the game)
P^a_t	denotes the actual proportion of co-operators in the group at time t.

We know that, if $P^e_0 \geq P^*$, AG players choose to co-operate at the beginning of the game and, as a result, the actual proportion of co-operators equals the proportion of AG players in the group: $P^a_1 = P^{AG}$. On the other hand, if $P^e_0 < P^*$, AG players choose to defect and $P^a_1 = 0$.

We are now ready for a discussion of the dynamics of collective action in a large group where there are two types of players with the preferences depicted in Figure 5.22. Four possibilities can be distinguished. Under the first possibility, we have $P^{AG} \geq P^e_0 \geq P^*$. The AG players co-operate from the beginning of the game, P^e_t is equal to P^{AG} for all t greater than zero, and their willingness to so behave is actually confirmed as more rounds are completed. If P^e_0 is strictly smaller than P^{AG}, these players realize after the first round that the actual proportion of co-operators in the group is higher than what they had initially expected (bear in mind that $P^a_1 = P^{AG}$ since $P^e_0 \geq P^*$). Consequently, their expectations are revised upwards and P^e_t becomes equal to P^{AG} at $t = 1$. If P^e_0 is equal to P^{AG}, AG players discover after the first round that their expectations are fully justified by experience and no change occurs in their expectations. In both cases, collective action is clearly a durable outcome.

The second possibility arises when the following conditions are satisfied: $P^e_0 > P^{AG} \geq P^*$. This is typically the case where AG players are overoptimistic about the likely behaviour of others, yet this does not prevent collective action from being established and sustained. The AG players participate in collective action but they are led to bring down their assessment of the likely proportion of co-operators in the light of the first round's experience.

Such is not the case under the third possibility where the overoptimism of AG players cannot avoid the collective action to suddenly collapse at the second round. This case obtains when we find $P^e_0 \geq P^* > P^{AG}$. The problem obviously arises from the fact that there are now in the group less AG players than required to induce *sustainable* co-operation ($P^{AG} < P^*$). After the first round, AG players choose to discontinue co-operation forever.

The fourth possibility is the most interesting one. It arises when the proportion of co-operators expected by AG players is smaller than the minimum required to induce co-operation among these players, that is, when $P^e_0 < P^* < P^{AG}$. In this case, nobody co-operates in the initial round and nobody will ever be incited to co-operate thereafter. In other words, even though there are actually enough willing co-operators in the (large) group to make co-operation possible, such co-operation fails to emerge because they do not have sufficient confidence in the group's inclination to co-operate. Because it cannot be corrected through a co-ordination mechanism, pessimism turns into a self-fulfilling prophecy. This case illustrates the critical importance of trust for co-operation to be possible in large groups.

Note that, even if there is one fully informed AG player who knows that there are actually enough players like him in the population to sustain co-operation, he will not choose to co-operate in the first round since, given the large size of the group, he is unable to persuade others to change their expectations and modify their behaviour. It would be wrong to think that such a result obtains because this individual player is alone to hold correct expectations. To see this, let us assume that, among AG players, there is a subgroup of players who hold optimistic expectations. These players are called subtype I AG players and are distinguished from another category called subtype II who are pessimistic. By optimists, we mean AG players who believe that the proportion of subtype I players in the population is at least equal to P^*. Pessimists are those AG players for whom the proportion of subtype I AG players is less than P^*.

Two different situations can arise. In a first case, the actual proportion of optimists in the population is higher than P^*. After one round, they realize that they are numerous enough to sustain co-operation, no matter what the pessimists do, and the latter are then led to revise their expectations upwards. From the second round onwards, the pessimists join the optimists in the collective action. The presence of the optimists, to paraphrase Elster, appears as a catalyst for co-operation while the pessimists act as a multiplier on the co-operation of the former (Elster, 1989*a*: 205). In the second case, the *actual* proportion of optimists in the population is lower than the critical level P^*. After one round when the optimists realize that they are not numerous enough to justify co-operation, and are unable to drive the pessimists in the collective action, they stop co-operating: universal defection ensues.

A richer picture of reality obtains when the assumption of two homogeneous subtypes of AG players is relaxed and replaced by the more realistic one that the *degree* of optimism of each player is different and unknown to the others. To put it in another way, the distribution of subtypes (i.e. optimism) among AG players is not known a priori. However, the analysis of such a situation lies beyond the scope of the present work. We shall here restrict ourselves to pointing out the main results which can be intuitively expected from such an analysis. The important point to note is that the revision of expectations now takes place in a gradual way after each round rather than in a discrete manner after the first round only. In a border case, all AG players start by co-operating and continue to co-operate forever since even the pessimists have high enough expectations to give co-operation a try. Experience confirms them in their behaviour. A more general case is when the most optimistic players start by co-operating but it turns out in the initial rounds that their number is too small to make co-operation worth while even for them. If those players are led to revise their expectations downwards, some initially pessimistic players may now be induced to co-operate. In such circumstances, it is impossible to say a priori whether co-operation will spread or gradually unravel. Note that in the latter, general case, the most favourable scenario occurs when co-operation is initiated by the most optimistic AG players, then, after subsequent rounds these players revise downwards their expectations yet still co-operate and they are joined by successive batches of players who were initially less optimistic than themselves.

Let us now return to the case where the AG players are divided between two subgroups. However, instead of assuming that members from subtypes I and II differ in terms of the more or less pessimistic character of their expectations, it is possible to differentiate them in terms of the intensity of their interest in co-operation. More precisely, we may assume that players from subtype I derive a higher payoff from co-operation than players from the other subtype, with the result that the threshold proportion for co-operation is lower for the more co-operation-interested players. Let us denote this assumption by writing $P^{*II} > P^{*I}$. Three interesting cases may be distinguished which lead to results analogous to those obtained in the above analysis of heterogeneous AG players. In a first situation, we have (assuming that players of the two subtypes have *similar expectations*):

$$P^{*I} < P^{e}_{0} < P^{*II} < P^{AGI},$$

where P^{AGI} stands for the proportion of subtype I AG players in the population. Under these conditions, all AG players participate in the collective action after the first round. Players I participate from the very beginning while players II first choose to defect but, as their expectations are being adjusted upwards, concrete experience from the first round gives them enough assurance of others' willingness to co-operate for themselves to join the collective action. This is the virtuous situation in which the more co-operation-inclined players succeed in *anonymously* persuading the less co-operation-inclined (but non-opportunistic) players to participate in collective action. Thanks to this demonstration effect, the former see their payoffs increase once the latter have joined them. *Ex post*, we can reinterpret the utility 'losses' incurred by players I during the first round as the necessary price to pay for dragging more prudent men of goodwill into the production of a public good, and thereby draw higher benefits from their own participation in this effort.

A second interesting situation obtains when the following conditions are satisfied:

$$P^{*I} < P^{e}_{0} < P^{AGI} < P^{*II}.$$

Here, the more co-operation-interested players continuously co-operate but, contrary to what we observed in the previous situation, they are not able to prevent the less co-operation-interested players from defecting. This is because, even though the latter's expectations are adjusted upwards, the threshold proportion P^{*II} will not be crossed. Such a situation is especially unfortunate if

$$P^{AG} > P^{*II},$$

that is, if the proportion of *all* AG players in the group actually exceeds that required to induce co-operation among the less co-operation-interested players.

There then remains the third, vicious case where even players I's willingness to co-operate unravels. This case is observed when

$$P^{AGI} < P^{*I} < P_0^e < P^{*II},$$

which conditions can also be satisfied when $P^{AG} > P^{*II}$. Players I start by co-operating but, as players II do not join hands with them, the actual proportion of co-operators (P^{AGI}) is too small to incite even the former to sustain their co-operative efforts.

5.4 Conclusion

Clearly, situations which can arise in field settings are of a much wider variety than what the tragedy of the commons implies. In the previous chapter, emphasis was laid on the fact that even within the PD framework repetition can possibly get people out of the non-co-operative equilibrium trap. In this chapter, it has been argued that this framework, although useful to account for many field situations which have really developed into the kind of tragedy envisioned by Hardin, is nevertheless too narrow to describe a whole range of other situations. Depending on the characteristics of the resource and the technique used as well as on various features of user groups (their size, their rate of discount of future income and the importance of their subsistence constraints, their exit possibilities, etc.), problems of resource exploitation may or may not be adequately described as PD games. Thus, such problems of resource management may well entail co-ordination or chicken game-like problems, or a mixture of different payoff structures. In this new perspective, the focus of the analysis is no more on the irresistible tendency of individuals to overexploit the commons. It is being shifted to human encounters involving problems of trust, leadership, co-ordination, group identity, and homogeneity or heterogeneity of group members.

A particularly striking result obtains in heterogeneous encounters with sequential moves in which the first agent has an AG payoff structure while the second agent has a CG, an AG, or even a PD payoff structure. If the second type of agent can assume leadership, co-operation will automatically ensue but the reverse is not true except in the case where both the leader and the follower happen to have an AG payoff structure. Clearly, the payoff profile of the leader matters a lot and, in a rather paradoxical way, co-operation is better ensured if 'nice' people do not occupy the leadership position.

Leadership does not necessarily refer to the ability to make the first move in a sequential decision-making process. It can also mean the ability to mobilize a sufficient number of people for enterprises requiring co-ordinated efforts. If such leadership is not present in these situations, collective action may not occur even though every agent would actually like to co-operate with the others.

The discussion about situations structured like asymmetric chicken games has shown the importance of precising the nature of power in order to be able to predict who, between the rich and the poor, are more likely to bear the cost of producing a public good (or preventing a public 'bad') in this kind of situation. Power can take various forms. It may be reflected in the ability to make a credible commitment to non-co-operation in the first stage of a sequential decision-making process. Or, it may have its source in exit possibilities that are not available to the other agents. Or again, it may express itself in the ability to lay down social norms that drive everybody to co-operate, irrespective of individual interests in the public good. Sheer poverty can, however, confer leverage upon the poor if the latter are so hard-pressed by subsistence constraints that they are not capable of producing the public good alone. Yet, even in this case, the third way of exercising power (imposing norms of participation) can enable the rich to transform the situation partly to their advantage. Note, moreover, that in situations involving co-ordination problems but where the efforts of the whole group are not required, power can manifest itself in the ability to exclude people from collective action, thereby preventing them from fully participating in the management of community affairs.

Regarding group size, it bears emphasis that the central conclusions reached at the end of Chapter 4 continue to hold true and are even reinforced when allowance is made for non-PD payoff structures. Thus, as the size of the group increases, due to incentive dilution a chicken game degenerates into a prisoners' dilemma with the result that no contribution, whether unilateral or universal, is made towards producing collective CPR infrastructures or no effort towards following use-restraining rules. Also, the fact that limited group size favours continuous interactions and easy observability and memorization of each other's actions proves to be a decisive factor in explaining the emergence of co-operation. In particular, when PD players coexist with AG players, it may be in the interest of the former to conceal their freerider type by co-operating till the last (few) stages of the game. This is not possible in large groups since the agents' co-operative moves cannot be interpreted by the others in a way conducive to universal co-operation. As a result, when numerous actors are involved, each of them tends to consider others' behaviour as a datum which he is unable to influence (Buchanan, 1975: 66).

In the previous chapter, the feasibility of pre-play communication in small-group settings has been emphasized. This aspect of the problem of collective action assumes special relevance when agents operate within an AG payoff structure. As a matter of fact, if such agents are able to signal to the others their predisposition to co-operate and their aversion to being 'exploited', the Pareto-superior equilibrium is very likely to be established and sustained. This is all the more true if the feeling of sameness or togetherness permeates the culture of the small group.