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:

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

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:

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

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

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

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:

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

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

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

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 $P^{*\mathrm{II}} > P^{*\mathrm{I}}$. 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*):

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

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:

$$P^{*\mathrm{I}} < P_0^e < P^{\mathrm{AGI}} < P^{*\mathrm{II}}.$$

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 $P^{*\mathrm{II}}$ will not be crossed. Such a situation is especially unfortunate if

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

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

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

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.