# 11

# Collective-Action Games

THE GAMES AND STRATEGIC SITUATIONS considered in the preceding chapters have usually included only two or three players interacting with each other. Such games are common in our own academic, business, political, and personal lives and so are important to understand and analyze. But many social, economic, and political interactions are strategic situations in which numerous players participate at the same time. Strategies for career paths, investment plans, rush-hour commuting routes, and even studying have associated benefits and costs that depend on the actions of many other people. If you have been in any of these situations, you likely thought something was wrong—too many students, investors, and commuters crowding just where you wanted to be, for example. If you have tried to organize fellow students or your community in some worthy cause, you probably faced frustration of the opposite kind—too few willing volunteers. In other words, multiple-person games in society often seem to produce outcomes that are not deemed satisfactory by many or even all of the people in that society. In this chapter, we will examine such games from the perspective of the theory that we have already developed. We present an understanding of what goes wrong in such situations and what can be done about it.

In the most general form, such many-player games concern problems of **collective action**. The aims of the whole society or collective are best served if its members take some particular action or actions, but these actions are not in the best private interests of those individual members. In other words, the socially optimal outcome is not automatically achievable as the Nash equilibrium of the game. Therefore, we must examine how the game can be modified to lead to the optimal outcome or at least to improve on an unsatisfactory Nash equilibrium. To do so, we must first understand the nature of such games. We find that they come in three forms, all of them familiar to you by now: the prisoners' dilemma, chicken, and assurance games. Although our main focus in this chapter is on situations where numerous individuals play such games at the same time, we build on familiar ground by beginning with games between just two players.

## 1 COLLECTIVE-ACTION GAMES WITH TWO PLAYERS

Imagine that you are a farmer. A neighboring farmer and you can both benefit by constructing an irrigation and flood-control project. The two of you can join together to undertake this project, or one of you might do so on your own. However, after the project has been constructed, the other automatically benefits from it. Therefore, each is tempted to leave the work to the other. That is the essence of your strategic interaction and the difficulty of securing collective action.

In Chapter 4, we encountered a game of this kind: three neighbors were each deciding whether to contribute to a street garden that all of them would enjoy. That game became a prisoners' dilemma in which all three shirked; our analysis here will include an examination of a more general range of possible payoff structures. Also, in the street-garden game, we rated the outcomes on a scale of 1 to 6; when we describe more general games, we will have to consider more general forms of benefits and costs for each player.

Our irrigation project has two important characteristics. First, its benefits are **nonexcludable**: a person who has not contributed to paying for it cannot be prevented from enjoying the benefits. Second, its benefits are **nonrival**: any one person's benefits are not diminished by the mere fact that someone else is also getting the benefit. Economists call such a project a **pure public good**; national defense is often given as an example. In contrast, a pure *private* good is fully excludable and rival: nonpayers can be excluded from its benefits, and if one person gets the benefit, no one else does. A loaf of bread is a good example of a pure private good. Most goods fall somewhere on the two-dimensional spectrum of varying degrees of excludability and rivalness. We will not go any deeper into this taxonomy, but we mention it to help you relate our discussion to what you may encounter in other courses and books.[1]

[1] Public goods are studied in more detail in textbooks on *public economics* such as those by Jonathan Gruber, *Public Finance and Public Policy*, 4th ed. (New York: Worth, 2012), Harvey Rosen and Ted Gayer, *Public Finance*, 9th ed. (Chicago: Irwin/McGraw-Hill, 2009), and Joseph Stiglitz, *Economics of the Public Sector*, 3rd ed. (New York: W. W. Norton & Company, 2000).

## A. Collective Action as a Prisoners' Dilemma

The costs and the benefits associated with building the irrigation project depend, as do those associated with all collective actions, on which players participate. In turn, the relative size of the costs and benefits determine the structure of the game that is played. Suppose each of you acting alone could complete the project in 7 weeks, whereas if the two of you acted together, it would take only 4 weeks of time from each. The two-person project is also of better quality; each farmer gets benefits worth 6 weeks of work from a one-person project (whether constructed by you or by your neighbor) and 8 weeks' worth of benefit from a two-person project.

More generally, we can write benefits and costs as functions of the number of players participating. So the cost to you of choosing to build the project depends on whether you build it alone or with help; costs can be written as $C(n)$ where cost, $C$, depends on the number, $n$, of players participating in the project. Then $C(1)$ would be the cost to you of building the project alone. $C(2)$ would be the cost *to you* of building the project with your neighbor; here $C(1) = 7$ and $C(2) = 4$. Similarly, benefits ($B$) from the completed project may vary depending on how many ($n$) participate in its completion. In our example, $B(1) = 6$ and $B(2) = 8$. Note that these benefits are the same for each farmer regardless of participation due to the public-good nature of this particular project.

In this game, each farmer has to decide whether to work toward the construction of the project or not—that is, to shirk. (Presumably, there is a short window of time in which the work must be done, and you could pretend to be called away on some very important family matter at the last minute, as could your neighbor.) Figure 11.1 shows the payoff table of the game, where the numbers measure the values in weeks of work. Payoffs are determined on the basis of the difference between the cost and the benefit associated with each action. So the payoff for choosing Build will be $B(n) - C(n)$ with $n = 1$ if you build alone and with $n = 2$ if your neighbor also chooses Build. The payoff for choosing Not is just $B(1)$ if your neighbor chooses Build, because you incur no cost if you do not participate in the project.

| | | NEIGHBOR | |
|---|---|---|---|
| | | Build | Not |
| YOU | Build | 4, 4 | −1, 6 |
| | Not | 6, −1 | 0, 0 |

**FIGURE 11.1** Collective Action as a Prisoners' Dilemma: Version I

Given the payoff structure in Figure 11.1, your best response if your neighbor does not participate is not to participate either: your benefit from completing the project by yourself (6) is less than your cost (7), for a net payoff of −1, whereas you can get 0 by not participating. Similarly, if your neighbor does participate, then you can reap the benefit (6) from his work at no cost to yourself; this is better for you than working yourself to get the larger benefit of the two-person project (8) while incurring the cost of the work (4), for a net payoff of 4. The general feature of the game is that it is better for you not to participate no matter what your neighbor does; the same logic holds for him. (In this case, each farmer is said to be a **free rider** on his neighbor's effort if he lets the other do all the work and then reaps the benefits all the same.) Thus, not building is the dominant strategy for each. But both would be better off if the two were to work together to build (payoff 4) than if neither builds (payoff 0). Therefore, the game is a prisoners' dilemma.

We see in this prisoners' dilemma one of the main difficulties that arises in games of collective action. Individually optimal choices—in this case, not to build regardless of what the other farmer chooses—may not be optimal from the perspective of society as a whole, even if the society is made up of just two farmers. The **social optimum** in a collective-action game is achieved when the sum total of the players' payoffs is maximized; in this prisoners' dilemma, the social optimum is the (Build, Build) outcome. Nash-equilibrium behavior of the players does not consistently bring about the socially optimal outcome, however. Hence, the study of collective-action games has focused on methods to improve on observed (generally Nash) equilibrium behavior to move outcomes toward the socially best ones. As we will see, the divergence between Nash equilibrium and socially optimum outcomes appears in every version of collective-action games.

Now consider what the game would look like if the numbers were to change slightly. Suppose the two-person project yields benefits that are not much better than those in the one-person project: 6.3 weeks' worth of work to each farmer. Then each of you gets $6.3 - 4 = 2.3$ when both of you build. The resulting payoff table is shown in Figure 11.2. The game is still a prisoners' dilemma and leads to

| | | NEIGHBOR | |
|---|---|---|---|
| | | Build | Not |
| YOU | Build | 2.3, 2.3 | −1, 6 |
| | Not | 6, −1 | 0, 0 |

**FIGURE 11.2** Collective Action as a Prisoners' Dilemma: Version II

the equilibrium (Not, Not). However, when both farmers build, the total payoff for both of you is only 4.6. The social optimum occurs when one of you builds and the other does not, in which case together you get payoff $6 + (-1) = 5$. There are two possible ways to get this outcome. Achieving the social optimum in this case then poses a new problem: Who should build and suffer the payoff of $-1$ while the other is allowed to be a free rider and enjoy the payoff of 6?

## B. Collective Action as Chicken

Yet another variation in the numbers of the original prisoners' dilemma game of Figure 11.1 changes the nature of the game. Suppose the cost of the work is reduced so that it becomes better for you to build your own project if your neighbor does not. Specifically, suppose the one-person project requires 4 weeks of work, so $C(1) = 4$, and the two-person project takes 3 weeks from each, so $C(2) = 3$ (to each); the benefits are the same as before. Figure 11.3 shows the payoff matrix resulting from these changes. Now your best response is to shirk when your neighbor works and to work when he shirks. In form, this game is just like a game of chicken, where shirking is the Straight strategy (tough or uncooperative), and working is the Swerve strategy (conciliatory or cooperative).

If this game results in one of its pure-strategy equilibria, the two payoffs sum to 8; this total is less than the total outcome that both players could get if both of them build. That is, neither of the Nash equilibria provides so much benefit to society as a whole as that of the coordinated outcome, which entails both farmers' choosing to build. The social optimum yields a total payoff of 10. If the outcome of the chicken game is its mixed-strategy equilibrium, the two farmers will fare even worse than in either of the pure-strategy equilibria: their expected payoffs will add up to something less than 8 (4, to be precise).

The collective-action chicken game has another possible structure if we make some additional changes to the benefits associated with the project. As with version II of the prisoners' dilemma, suppose the two-person project is not much better than the one-person project. Then each farmer's benefit from the two-person project, $B(2)$, is only 6.3, whereas each still gets a benefit of $B(1) = 6$

| | | NEIGHBOR | |
|---|---|---|---|
| | | Build | Not |
| YOU | Build | 5, 5 | 2, 6 |
| | Not | 6, 2 | 0, 0 |

**FIGURE 11.3** Collective Action as Chicken: Version I

from the one-person project. We ask you to practice your skill by constructing the payoff table for this game. You will find that it is still a game of chicken—call it chicken II. It still has two pure-strategy Nash equilibria in each of which only one farmer builds, but the sum of the payoffs when both build is only 6.6, whereas the sum when only one farmer builds is 8. The social optimum is for only one farmer to build. Each farmer prefers the equilibrium in which the other builds. This may lead to a new dynamic game in which each waits for the other to build. Or the original game might yield its mixed-strategy equilibrium with its low expected payoffs.

## C. Collective Action as Assurance

Finally, let us change the payoffs of the original prisoners' dilemma case in a different way altogether, leaving the benefits of the two-person project and the costs of building as originally set out and reducing the benefit of a one-person project to $B(1) = 3$. This change reduces your benefit as a free rider so much that now if your neighbor chooses Build, your best response also is Build. Figure 11.4 shows the payoff table for this version of the game. This is now an assurance game with two pure-strategy equilibria: one where both of you participate and the other where neither of you does.

As in the chicken II version of the game, the socially optimal outcome here is one of the two Nash equilibria. But there is a difference. In chicken II, the two players differ in their preferences between the two equilibria, either of which achieves the social optimum. In the assurance game, both of them prefer the same equilibrium, and that is the sole socially optimal outcome. Therefore, achieving the social optimum should be easier in the assurance game than in chicken.

## D. Collective Inaction

Many games of collective action have payoff structures that differ somewhat from those in our irrigation project example. Our farmers find themselves in a situation in which the social optimum generally entails that at least one, if not

| | | NEIGHBOR | |
|---|---|---|---|
| | | Build | Not |
| YOU | Build | 4, 4 | −4, 3 |
| | Not | 3, −4 | 0, 0 |

**FIGURE 11.4** Collective Action as an Assurance Game

both, of them participates in the project. Thus the game is one of collective *action*. Other multiplayer games might better be called games of collective *inaction*. In such games, society as a whole prefers that some or all of the individual players do *not* participate or do *not* act. Examples of this type of interaction include choices between rush-hour commuting routes, investment plans, or fishing grounds.

All of these games have the attribute that players must decide whether to take advantage of some common resource, be it a freeway, a high-yielding stock fund, or an abundantly stocked pond. These collective "inaction" games are better known as common-resource games; the total payoff to all players reaches its maximum when players refrain from overusing the common resource. The difficulty associated with not being able to reach the social optimum in such games is known as the "tragedy of the commons," a phrase coined by Garrett Hardin in his paper of the same name.[2]

We supposed above that the irrigation project yielded equal benefits to both you and your farmer-neighbor. But what if the outcome of both farmers' building was that the project used so much water that the farms had too little water for their livestock? Then each player's payoff could be negative when both choose Build, lower than when both choose Not. This would be yet another variant of the prisoners' dilemma we encountered in Section 1.A, in which the socially optimal outcome entails neither farmer's building even though each one still has an individual incentive to do so. Or suppose that one farmer's activity causes harm to the other, as would happen if the only way to prevent one farm from being flooded is to divert the water to the other. Then each player's payoffs could be negative if his neighbor chose Build. Thus, another variant of chicken could also arise. In this variant, each of you wants to build when the other does not, whereas it would be collectively better if neither of you did.

Just as the problems pointed out in these examples of both collective action and collective inaction are familiar, the various alternative ways of tackling the problems also follow the general principles discussed in earlier chapters. Before turning to solutions, let us see how the problems manifest themselves in the more realistic setting where several players interact simultaneously in such games.

## 2 COLLECTIVE-ACTION PROBLEMS IN LARGE GROUPS

In this section, we extend our irrigation-project example to a situation in which a population of $N$ farmers must each decide whether to participate. Here we make use of the notation we introduced above, with $C(n)$ representing the cost

[2] Garrett Hardin, "The Tragedy of the Commons," *Science*, vol. 162 (1968), pp. 1243–48.

each participant incurs when $n$ of the $N$ total farmers have chosen to participate. Similarly, the benefit to each, regardless of participation, is $B(n)$. Each participant then gets the payoff $P(n) = B(n) - C(n)$, whereas each nonparticipant, or shirker, gets the payoff $S(n) = B(n)$.

Suppose you are contemplating whether to participate or to shirk. Your decision will depend on what the other $(N - 1)$ farmers in the population are doing. In general, you will have to make your decision when the other $(N - 1)$ players consist of $n$ participants and $(N - 1 - n)$ shirkers. If you decide to shirk, the number of participants in the project is still $n$, so you get a payoff of $S(n)$. If you decide to participate, the number of participants becomes $n + 1$, so you get $P(n + 1)$. Therefore, your final decision depends on the comparison of these two payoffs; you will participate if $P(n + 1) > S(n)$, and you will shirk if $P(n + 1) < S(n)$. This comparison holds true for every version of the collective-action game analyzed in Section 1; differences in behavior in the different versions arise because the changes in the payoff structure alter the values of $P(n + 1)$ and $S(n)$.

We can relate the two-person examples of Section 1 to this more general framework. If there are just two people, then $P(2)$ is the payoff to one from building when the other also builds, $S(1)$ is the payoff to one from shirking when the other builds, and so on. Therefore, we can generalize the payoff tables of Figures 11.1 through 11.4 into an algebraic form. This general payoff structure is shown in Figure 11.5.

The game illustrated in Figure 11.5 is a prisoners' dilemma if the inequalities

$$P(2) < S(1), \quad P(1) < S(0), \quad P(2) > S(0)$$

all hold at the same time. The first says that the best response to Build is Not, the second says that the best response to Not also is Not, and the third says that (Build, Build) is jointly preferred to (Not, Not). The dilemma is of type I if $2P(2) > P(1) + S(1)$, so the total payoff is higher when both build than when only one builds. You can establish similar inequalities concerning these payoffs that yield the other types of games in Section 1.

Return now to the multiplayer version of the game with a general $n$. Given the payoff functions for the two actions, $P(n + 1)$ and $S(n)$, we can use graphs to

| | | NEIGHBOR | |
|---|---|---|---|
| | | Build | Not |
| YOU | Build | $P(2)$, $P(2)$ | $P(1)$, $S(1)$ |
| | Not | $S(1)$, $P(1)$ | $S(0)$, $S(0)$ |

**FIGURE 11.5** General Form of a Two-Person Collective-Action Game

help us determine which type of game we have encountered and its Nash equilibrium. We can also then compare the Nash equilibrium to the game's socially optimal outcome.

## A. Multiplayer Prisoners' Dilemma

Take a specific version of our irrigation project example in which an entire village of 100 farmers must decide which action to take. Suppose that the irrigation project raises the productivity of each farmer's land in proportion to the size of the project; specifically, suppose the benefit to each farmer when $n$ people work on the project is $P(n) = 2n$. Suppose also that if you are not working on the project, you can enjoy this benefit and use your time to earn an extra 4 in some other occupation, so $S(n) = 2n + 4$. Remember that your decision about whether to participate in the project depends on the relative magnitudes of $P(n + 1) = 2(n + 1)$ and $S(n) = 2n + 4$. We draw the two separate graphs of these functions for an individual farmer in Figure 11.6, showing $n$ over its full range from 0 to $(N - 1)$ along the horizontal axis and the payoff to the farmer along the vertical axis. If there are currently very few participants (thus mostly shirkers), your choice will depend on the relative locations of $P(n + 1)$ and $S(n)$ on the left end of the graph. Similarly, if there are already many participants, your choice will depend on the relative locations of $P(n + 1)$ and $S(n)$ on the *right* end of the graph.

Because $n$ actually takes on only integer values, each function $P(n + 1)$ and $S(n)$ technically consists only of a discrete set of points rather than a continuous

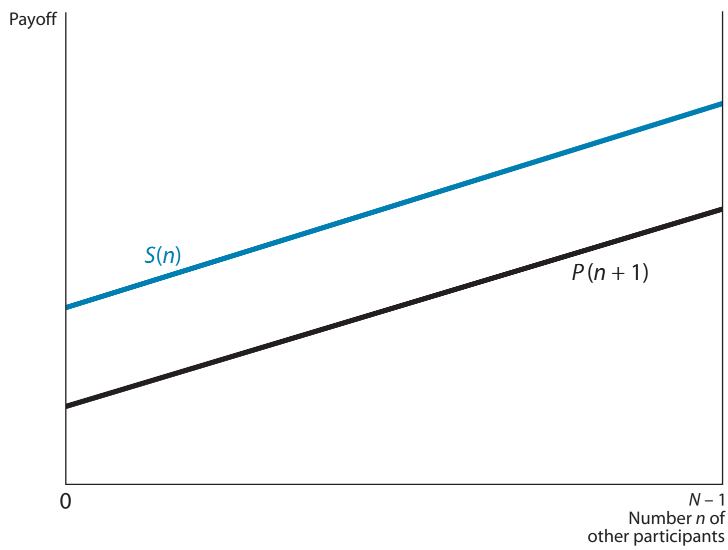



**FIGURE 11.6** Multiplayer Prisoners' Dilemma Payoff Graph

set as implied by our smooth lines. But when $N$ is large, the discrete points are sufficiently close together that we can connect the successive points and show each payoff function as a continuous curve. We also use linear $P(n + 1)$ and $S(n)$ functions in this section to bring out the basic considerations and will discuss more complicated possibilities later.

Recall that you determine your choice of action by considering the number of current participants in the project, $n$, and the payoffs associated with each action at that $n$. Figure 11.6 illustrates a case in which the curve $S(n)$ lies entirely above the curve $P(n + 1)$. Therefore, no matter how many others participate (that is, no matter how large $n$ gets), your payoff is higher if you shirk than if you participate; shirking is your dominant strategy. These payoffs are identical for all players, so everyone has a dominant strategy to shirk. Therefore, the Nash equilibrium of the game entails everyone shirking, and the project is not built.

Note that both curves are rising as $n$ increases. For each action you can take, you are better off if more of the others participate. And the left intercept of the $S(n)$ curve is below the right intercept of the $P(n + 1)$ curve, or $S(0) = 4 < P(N) = 102$. This says that if everyone including you shirks, your payoff is less than if everyone including you participates. Everyone would be better off than they are in the Nash equilibrium of the game if the outcome in which everyone participates could be sustained. This makes the game a prisoners' dilemma.

How does the Nash equilibrium found using the curves in Figure 11.6 compare with the social optimum of this game? To answer this question we need a way to describe the total social payoff at each value of $n$; we do that by using the payoff functions $P(n)$ and $S(n)$ to construct a third function $T(n)$, showing the total payoff to society as a function of $n$. The total payoff to society when there are $n$ participants consists of the value $P(n)$ for each of the $n$ participants and the value $S(n)$ for each of the $(N - n)$ shirkers:

$$T(n) = nP(n) + (N - n)\,S(n).$$

The social optimum occurs when the allocation of people between participants and shirkers maximizes the total payoff $T(n)$, or at the number of participants—that is, the value of $n$—that maximizes $T(n)$. To get a better understanding of where this might be, it is convenient to write $T(n)$ differently, rearranging the expression above to get

$$T(n) = NS(n) - n\,[S(n) - P(n)].$$

This version of the total social payoff function shows that we can calculate it as if we gave every one of the $N$ people the shirker's payoff but then removed the shirker's extra benefit $[S(n) - P(n)]$ from each of the $n$ participants.

In collective-action games, as opposed to common-resource games, we normally expect $S(n)$ to increase as $n$ increases. Therefore, the first term in this expression, $NS(n)$, also increases as $n$ increases. If the second term does not increase too fast as $n$ increases—as would be the case if the shirker's extra benefit, $[S(n) - P(n)]$, is small and constant—then the effect of the first term dominates in determining the value of $T(n)$.

This is exactly what happens with the total social payoff function for our current 100-farmer example. Here $T(n) = n\,P(n) + (N - n)\,S(n)$ becomes $T(n) = n(2n) + (100 - n)\,(2n + 4) = 2n^2 + 200n - 2n^2 + 400 - 4n = 400 + 196n$. In this case, $T(n)$ increases steadily with $n$ and is maximized at $n = N$ when no one shirks.

The large-group version of our two-person example holds the same lesson as above. Society as a whole would be better off if all of the farmers participated in building the irrigation project and $n = N$. But payoffs are such that each farmer has an individual incentive to shirk. The Nash equilibrium of the game, at $n = 0$, is not socially optimal. Figuring out how to achieve the social optimum is one of the most important topics in the study of collective action and one to which we return later in this chapter.

In other situations, $T(n)$ can be maximized for a different value of $n$, not just at $n = N$. That is, society's aggregate payoff could be maximized by allowing some shirking. Even in the prisoners' dilemma case, it is not automatic that the total payoff function is maximized when $n$ is as large as possible. If the gap between $S(n)$ and $P(n)$ widens sufficiently fast as $n$ increases, then the negative effect of the second term in the expression for $T(n)$ outweighs the positive effect of the first term as $n$ approaches $N$; then it may be best to let some people shirk—that is, the socially optimal value for $n$ may be less than $N$. This result mirrors that of our prisoners' dilemma II case in Section 1.

This type of outcome would arise in our village if $S(n)$ were $4n + 4$, rather than $2n + 4$. Then $T(n) = -2n^2 + 396n + 400$, which is no longer linear in $n$. In fact, a graphing calculator or some basic calculus shows that this $T(n)$ is maximized at $n = 99$ rather than at $n = 100$ as was true before. The change to the payoff structure has created an inequality in the payoffs—the shirkers fare better than the participants—which adds another dimension of difficulty to society's attempts to resolve the dilemma. How, for example, would the village designate exactly one farmer to be the shirker?

## B. Multiplayer Chicken

Now we consider some of the other configurations that can arise in the payoffs. For example, when $P(n) = 4n + 36$, so $P(n + 1) = 4n + 40$, and $S(n) = 5n$, the two payoff curves will cross in the figure. This case is illustrated in Figure 11.7. Here, for small values of $n$, $P(n + 1) > S(n)$, so if few others are participating, your choice is to participate. For large values of $n$, $P(n + 1) < S(n)$, so if many others are participating, your choice is to shirk. Note the equivalence of these two statements to the idea in the two-person chicken game that "you shirk if your neighbor works and you work if he shirks." This case is indeed that of chicken. More generally, the chicken case occurs when you are given a choice between two actions, and you prefer to do the one that most others are *not* doing.

We can also use Figure 11.7 to determine the location of the Nash equilibrium of this version of the game. Because you choose to participate when $n$ is small and to shirk when $n$ is large, the equilibrium must be some intermediate value of $n$. Only at that $n$ where the two curves intersect are you indifferent between your two choices. This location represents the equilibrium value of $n$. In our graph, $P(n + 1) = S(n)$ when $4n + 40 = 5n$ or when $n = 40$; that is the Nash equilibrium number of farmers from the village who will participate in the irrigation project.

If the two curves intersect at a point corresponding to an integer value of $n$, then that is the Nash equilibrium number of participants. If that is not the case, then strictly speaking the game has no Nash equilibrium. But in practice, if the current value of $n$ in the population is the integer just to the left of the point of intersection, then one more person will just want to participate, whereas if the current value of $n$ is the integer just to the right of the point of intersection, one person will want to switch to shirking. Therefore, the number of participants will stay in a small neighborhood of the point of intersection, and we can justifiably speak of the intersection as the equilibrium in some approximate sense.

The payoff structure illustrated in Figure 11.7 shows both lines positively sloped, although they don't have to be. It is conceivable that the benefit for each person is smaller when more people participate, so the lines could be negatively sloped instead. The important feature of the chicken collective-action game is that when few people are taking one action, it is better for any one person to

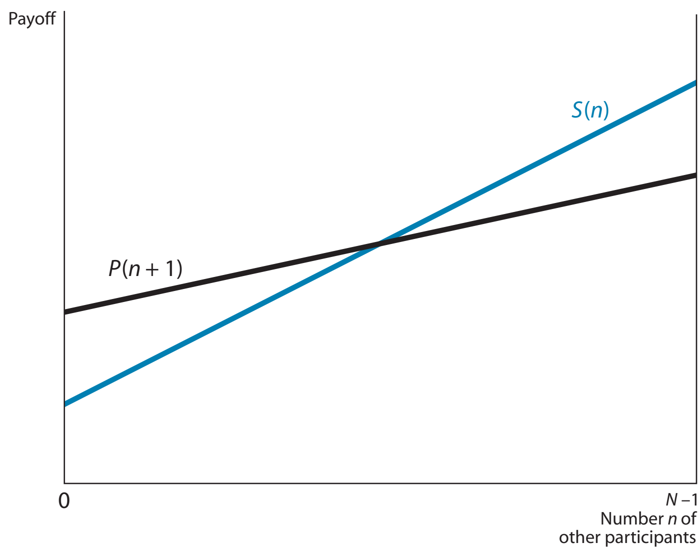



**FIGURE 11.7** Multiplayer Chicken Payoff Graph

take that action; when many people are taking one action, it is better for any one person to take the other action.

What is the socially optimal outcome in the chicken form of collective action? If each participant's payoff $P(n)$ increases as the number of participants increases, and if each shirker's payoff $S(n)$ does not become too much greater than the $P(n)$ of each participant, then the total social payoff is maximized when everyone participates. This is the outcome in our example where $T(n) = 536n - n^2$; total social payoff increases in $n$ beyond the value of $N$ (100 here), so $n = N$ is the social optimum.

But more generally, some cases of chicken will entail social optima in which it is better to let some shirk. If our group of farmers numbered 300 instead of 100, our example here would yield such an outcome. The socially optimal number of participants, found on a graphing calculator or using calculus, would be 268. This is exactly the difference between versions I and II of chicken in our example in Section 1. For an exercise, you may try generating a payoff structure that leads to such an outcome for our village of 100 farmers. In these more general chicken games, the optimal number of participants could even be smaller than that in the Nash equilibrium. We return to examine the question of the social optimum of all of these versions of the game in greater detail in Section 3.

## C. Multiplayer Assurance

Finally, we consider the third possible type of collective-action game, assurance. Figure 11.8 shows the payoff lines for the assurance case, where we suppose that the village farmers get $P(n + 1) = 4n + 4$ and $S(n) = 2n + 100$. Here $S(n) >$

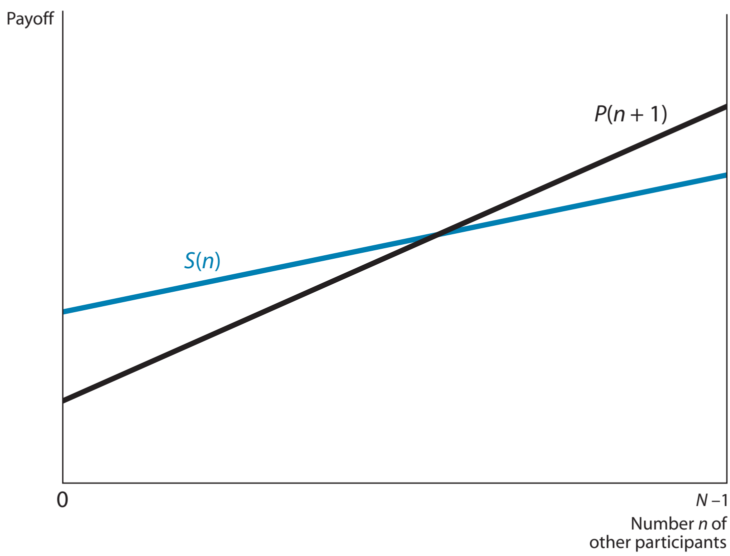



**FIGURE 11.8** Multiplayer Assurance Payoff Graph

$P(n + 1)$ for small values of $n$, so if few others are participating, then you want to shirk, too. But $P(n + 1) > S(n)$ for large values of $n$, so if many others are participating, then you want to participate too. In other words, unlike chicken, assurance is a collective-action game in which you want to make the choice that the others are making.

Except for the labels, the graph in Figure 11.8 looks nearly identical to that in Figure 11.7. The location of the Nash equilibrium depends critically on the labels associated with the two lines, however. In Figure 11.8, for any initial value of $n$ to the left of the intersection, each farmer will want to shirk, and there will be a Nash equilibrium at $n = 0$ where everyone shirks. But the opposite is true to the right of the intersection. In that portion of the graph, each farmer will want to participate, and there will be a second Nash equilibrium at $n = N$.

Technically, there is also a third Nash equilibrium of this game if the value of $n$ at the intersection is an integer value as it is in our example. There we find that $P(n + 1) = 4n + 4 = 2n + 100 = S(n)$ when $n = 48$. Then if $n$ were exactly 48, we would see an outcome in which there were some participants and some shirkers. This situation could be an equilibrium only if the value of $n$ is exactly right. Even then, it would be a highly unstable situation. If any one farmer accidentally joined the wrong group, his choice would alter the incentives for everyone else, driving the game to one of the endpoint equilibria. Those are the two stable Nash equilibria of the game.

The social optimum in this game is fairly easy to see on the graph in Figure 11.8. Because both curves are rising—so each person is better off if more people participate—then clearly the right-hand extreme equilibrium is the better one for society. This is confirmed in our example by noting that $T(n) = 2n^2 + 100n + 10{,}000$, which is increasing in $n$ for all positive values of $n$; thus the socially optimal value of $n$ is the largest one possible, or $n = N$. In the assurance case, then, the socially optimal outcome is actually one of the stable Nash equilibria of the game. As such, it may be easier to achieve than in some of the other cases. The critical question regarding the social optimum, regardless of whether it represents a Nash equilibrium of the underlying game, is how to bring it about.

So far, our examples have focused on relatively small groups of 2 or 100 persons. When the total number of people in the group, $N$, is very large, however, and any one person makes only a very small difference, then $P(n + 1)$ is almost the same as $P(n)$. Thus, the condition under which any one person chooses to shirk is $P(n) < S(n)$. Expressing this inequality in terms of the benefits and costs of the common project in our example—namely, $P(n) = B(n) - C(n)$ and $S(n) = B(n)$—we see that $P(n)$ [unlike $P(n + 1)$ in our preceding calculations] is *always* less than $S(n)$; individual persons will *always* want to shirk when $N$ is very large. That is why problems of collective provision of public projects in a large group almost always manifest themselves as prisoners' dilemmas. But as we have seen, this result is not necessarily true for smaller groups. Neither is it true for large groups in other contexts such as congestion, a case we discuss later in this chapter.

In general, we must allow for a broader interpretation of the payoffs $P(n)$ and $S(n)$ than we did in the specific case involving the benefits and the costs of a project. We cannot assume, for example, that the payoff functions will be linear. In fact, in the most general case, $P(n)$ and $S(n)$ can be any functions of $n$ and can intersect many times. Then there can be multiple equilibria, although each can be thought of as representing one of the types described so far.[3] And some games will be of the common-resource type as well, so when we allow for completely general games, we will speak of two actions labeled $P$ and $S$, which have no necessary connotation of "participation" and "shirking" but allow us to continue with the same symbols for the payoffs. Thus, when $n$ players are taking the action $P$, $P(n)$ becomes the payoff of each player taking the action $P$, and $S(n)$ becomes that of each player taking the action $S$.

## 3 SPILLOVERS, OR EXTERNALITIES

So far, we have seen that collective-action games occur in prisoners' dilemma, chicken, and assurance forms. We have also seen that the Nash equilibria in such games rarely yield the socially optimal level of participation (or restraint). And even when the social optimum is a Nash equilibrium, it is usually only one of several equilibria that may arise. Now we delve further into the differences between the individual (or private) incentives in such games and the group (or social) incentives. We also describe more carefully the effects of each individual's decision on other individuals as well as on the collective. This analysis makes explicit why differences in incentives exist, how they are manifested, and how one might go about achieving socially better outcomes than those that arise in Nash equilibrium.

### A. Commuting and Spillovers

We start by thinking about a large group of 8,000 commuters who drive every day from a suburb to the city and back. As one of these commuters, you may take either the expressway (action P) or a network of local roads (action S). The

[3] Several exercises at the end of this chapter present some examples of simple situations with nonlinear payoff curves and multiple equilibria. For a more general analysis and classification of such diagrams, see Thomas Schelling, *Micromotives and Macrobehavior* (New York: W. W. Norton & Company, 1978), ch. 7. The theory can be taken further by allowing each player a continuous choice (for example, the number of hours of participation) instead of just a binary choice of whether to participate. Many such situations are discussed in more specialized books on collective action, for example, Todd Sandler, *Collective Action: Theory and Applications* (Ann Arbor: University of Michigan Press, 1993), and Richard Cornes and Todd Sandler, *The Theory of Externalities, Public Goods, and Club Goods,* 2nd ed. (New York: Cambridge University Press, 1996).

local-roads route takes a constant 45 minutes, no matter how many cars are going that way. The expressway takes only 15 minutes when uncongested. But every driver who chooses the expressway increases the time for every other driver on the expressway by 0.005 minutes (about one-quarter of a second).

Measure the payoffs in minutes of time saved—by how much the commute time is less than 1 hour, for instance. Then the payoff to drivers on the local roads, $S(n)$, is a constant $60 - 45 = 15$, regardless of the value of $n$. But the payoff to drivers on the expressway, $P(n)$, depends on $n$; in particular, $P(n) = 60 - 15 = 45$ for $n = 0$, but $P(n)$ decreases by 5/1,000 (or 1/200) for every commuter on the expressway. Thus, $P(n) = 45 - 0.005n$. We graph the two payoff lines in Figure 11.9.

Suppose that initially 4,000 cars are on the expressway; $n = 4{,}000$. With so many cars on that road, it takes each of them $15 + 4{,}000 \times 0.005 = 15 + 20 = 35$ minutes to commute to work; each gets a payoff of $P(n) = 25$ [which is $60 - 35$, or $P(4{,}000)$]. As shown in Figure 11.9, that payoff is better than what local-road drivers obtain. You, a local-road driver, might therefore decide to switch from driving the local roads to driving on the expressway. Your switch would increase by 1 the value of $n$ and would thereby affect the payoffs of all the other commuters. There would now be 4,001 drivers (including you) on the expressway, and the commute time for each would be 35 and 1/200, or 35.005, minutes; each would now get a payoff of $P(n + 1) = P(4{,}001) = 24.995$. This payoff is still higher than the 15 from driving on the local roads. Thus, you have a private incentive to make the switch, because for you, $P(n + 1) > S(n)$ $(24.995 > 15)$.

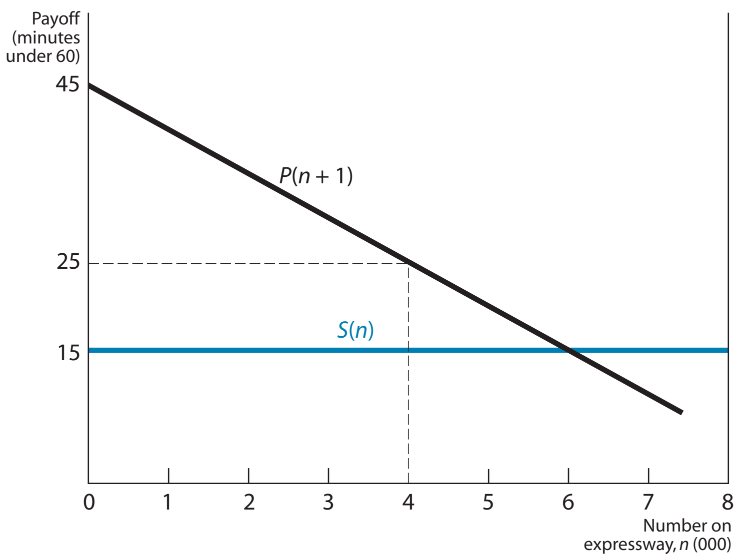



**FIGURE 11.9** Commuting Route-Choice Game

Your switch yields you a *private* gain—because it is privately enjoyed by you—equal to the difference between your payoffs before and after the switch; this private gain is $P(n + 1) - S(n) = 9.995$ minutes. Because you are only one person and therefore a small part of the whole group, the gain in payoff that you receive in relation to the total group payoff is small, or *marginal*. Thus, we call your gain the **marginal private gain** associated with your switch.

But now the 4,000 other drivers on the expressway each take 0.005 of a minute more as a result of your decision to switch; the payoff to each changes by $P(4{,}001) - P(4{,}000) = -0.005$. Similarly, the drivers on the local roads face a payoff change of $S(4{,}001) - S(4{,}000)$, but this is zero in our example. The cumulative effect on all of these other drivers is $4{,}000 \times -0.005 = -20$ (minutes). Your action, switching from local roads to expressway, has caused this effect on the others' payoffs. Whenever one person's action affects others like this, it is called a **spillover effect**, **external effect**, or **externality**. Again, because you are but a very small part of the whole group, we should actually call your effect on others the *marginal spillover effect*.

Taken together, the marginal private gain and the marginal spillover effect are the full effect of your switch on the group of commuters, or the overall marginal change in the whole group's or the whole society's payoff. We call this the **marginal social gain** associated with your switch. This "gain" may actually be positive or negative, so the use of the word *gain* is not meant to imply that all switches will benefit the group as a whole. In fact, in our commuting example, the overall marginal social gain is $9.995 - 20 = -10.005$ (minutes). Thus, the overall social effect of your switch is bad; the social payoff is reduced by a total of just over 10 minutes.

## B. Spillovers: The General Case

We can describe the effects we observe in the commuting example more generally by returning to our total social payoff function, $T(n)$, where $n$ represents the number of people choosing P, so $N - n$ is the number of people choosing S. Suppose that initially $n$ people have chosen P and that one person switches from S to P. Then the number choosing P increases by 1 to $(n + 1)$, and the number choosing S decreases by 1 to $(N - n - 1)$, so the total social payoff becomes

$$T(n + 1) = (n + 1)\, P(n + 1) + [N - (n + 1)]\, S(n + 1).$$

The increase in the total social payoff is the difference between $T(n)$ and $T(n+1)$:

$$\begin{aligned} T(n + 1) - T(n) &= (n + 1)\, P(n + 1) + [N - (n + 1)]\, S(n + 1) - n\, P(n) + (N - n)\, S(n) \\ &= [P(n + 1) - S(n)] + n\, [P(n + 1) - P(n)] \\ &\quad + [N - (n + 1)]\, [S(n + 1) - S(n)] \end{aligned} \tag{11.1}$$

after collecting and rearranging terms.

Equation (11.1) describes mathematically the various different effects of one person's switch from S to P that we saw earlier in the commuting example. The equation shows how the marginal social gain is divided into the marginal change in payoffs for the subgroups of the population.

The first of the three terms in Eq. (11.1)—namely, $[P(n + 1) - S(n)]$—is the marginal private gain enjoyed by the person who switches. As we saw above, this term is what drives a person's choice, and all such individual choices then determine the Nash equilibrium.

The second and third terms in Eq. (11.1) are just the quantifications of the *spillover effects* of one person's switch on the others in the group. For the $n$ other people choosing P, each sees his payoff change by the amount $[P(n + 1) - P(n)]$ when one more person switches to P; this spillover effect is seen in the second group of terms in Eq. (11.1). There are also $N - (n + 1)$ (or $N - n - 1$) others still choosing S after the one person switches, and each of these players sees his payoff change by $[S(n + 1) - S(n)]$; this spillover effect is shown in the third group of terms in the equation. Of course, the effect that one driver's switch has on the time for any one driver on either route is very small, but, when there are numerous other drivers (that is, when $N$ is large), the full spillover effect can be substantial.

Thus, we can rewrite Eq. (11.1) for a general switch of one person from either S to P or P to S as:

Marginal social gain = marginal private gain + marginal spillover effect.

For an example in which one person switches from S to P, we have

$$\begin{aligned}\text{Marginal social gain} &= T(n + 1) - T(n),\\ \text{Marginal private gain} &= P(n + 1) - S(n), \text{ and}\\ \text{Marginal spillover effect} &= n[P(n + 1) - P(n)] + [N - (n + 1)]\,[S(n + 1) - S(n)].\end{aligned}$$

**USING CALCULUS FOR THE GENERAL CASE** Before examining some spillover situations in more detail to see what can be done to achieve socially better outcomes, we restate the general concepts of the analysis in the language of calculus. If you do not know this language, you can omit the remainder of this section without loss of continuity; if you do know it, you will find the alternative statement much simpler to grasp and to use than the algebra employed earlier.

If the total number $N$ of people in the group is very large—say, in the hundreds or thousands—then one person can be regarded as a very small, or infinitesimal, part of this whole. This allows us to treat the number $n$ as a continuous variable. If $T(n)$ is the total social payoff, we calculate the effect of changing $n$ by considering an increase of an infinitesimal marginal quantity $dn$, instead of a full unit increase from $n$ to $(n + 1)$. To the first order, the change in payoff is $T'(n)dn$, where $T'(n)$ is the derivative of $T(n)$ with respect to $n$. Using the expression for the total social payoff,

$$T(n) = nP(n) + (N - n)\,S(n),$$

and differentiating, we have

$$\begin{aligned} T'(n) &= P(n) + nP'(n) - S'(n) + (N - n)S'(n) \\ &= [P(n) - S(n)] + nP'(n) + (N - n)S'(n). \end{aligned} \tag{11.2}$$

This is the calculus equivalent of Eq. (11.1). $T'(n)$ represents the marginal social gain. The marginal private gain is $P(n) - S(n)$, which is just the change in the payoff of the person making the switch from S to P. In Eq. (11.1), we had $P(n + 1) - S(n)$ for this change in payoff; now we have $P(n) - S(n)$. This is because the infinitesimal addition of $dn$ to the group of the $n$ people choosing P does not change the payoff to any one of them by a significant amount. However, the total change in their payoff, $nP'(n)$, is sizable and is recognized in the calculation of the spillover effect [it is the second term in Eq. (11.2)] as is the change in the payoff of the $(N - n)$ people choosing S [namely, $(N - n)\,S'(n)$], the third term in Eq. (11.2). These last two terms constitute the marginal-spillover-effect part of Eq. (11.2).

In the commuting example, we had $P(n) = 45 - 0.005n$, and $S(n) = 15$. Then with the use of calculus, we see that the private marginal gain for each driver who switches to the expressway when $n$ drivers are already using it is $P(n) - S(n) = 30 - 0.005n$. Because $P'(n) = -0.005$ and $S'(n) = 0$, the spillover effect is $n \times (-0.005) + (N - n) \times 0 = -0.005n$, which equals $-20$ when $n = 4{,}000$. The answer is the same as before, but calculus simplifies the derivation and helps us find the optimum directly.

## C. Commuting Revisited: Negative Externalities

A negative externality exists when the action of one person *lowers* others' payoffs; it imposes some extra costs on the rest of society. We saw this in our commuting example, where the marginal spillover effect of one person's switch to the expressway was negative, entailing an extra 20 minutes of drive time for other commuters. But the individual who changes his route to work does not take the spillover—the externality—into account when making his choice. He is motivated only by his own payoffs. (Remember that any guilt that he may suffer from harming others should already be reflected in his payoffs.) He will change his action from S to P as long as this change has a positive marginal *private* gain. He is then made better off by the change.

But society would be better off if the commuter's decision were governed by the marginal *social* gain. In our example, the marginal social gain is negative ($-10.005$), but the marginal private gain is positive (9.995), so the individual driver makes the switch even though society as a whole would be better off if he did not do so. More generally, in situations with negative externalities, the marginal social gain will be smaller than the marginal private gain due to the existence of the negative spillover effect. Individuals will make decisions based on a cost-benefit calculation that is the wrong one from society's perspective. As a result, individual persons will choose actions with negative spillover effects more often than society would like them to do.

We can use Eq. (11.1) to calculate the precise conditions under which a switch will be beneficial for a particular person versus for society as a whole. Recall that if $n$ people are already using the expressway and another driver is contemplating switching from the local roads to the expressway, he stands to gain from this switch if $P(n + 1) > S(n)$, whereas the total social payoff increases if $T(n + 1) - T(n) > 0$. The private gain is positive if

$$\begin{aligned} 45 - (n + 1) \times 0.005 &> 15 \\ 44.995 - 0.005n &> 15 \\ n &< 200\,(44.995 - 15) = 5{,}999, \end{aligned}$$

whereas the condition for the social gain to be positive is

$$\begin{aligned} 45 - (n + 1) \times 0.005 - 15 - 0.005n &> 0 \\ 29.995 - 0.01n &> 0 \\ n &< 2{,}999.5. \end{aligned}$$

Thus, if given the free choice, commuters will crowd onto the expressway until there are almost 6,000 of them, but all crowding beyond 3,000 reduces the total social payoff. Society as a whole would be best off if the number of commuters on the expressway were kept down to 3,000.

We show this result graphically in Figure 11.10; this figure replicates Figure 11.9 with the addition of marginal private and social gain lines. The two lines indicating $P(n + 1)$ and $S(n)$ meet at $n = 5{,}999$; that is, at the value of $n$ for which $P(n + 1) = S(n)$ or for which the marginal private gain is just zero. Everywhere to the left of this value of $n$, any one driver on the local roads calculates that he gets a positive gain by switching to the expressway. As some drivers make this switch, the numbers on the expressway increase—the value of $n$ in society rises as was the case in our example in Section 3.A. Conversely, to the right of the intersection point (that is, for $n > 5{,}999$), $S(n) > P(n + 1)$; so each of the $(n + 1)$ drivers on the expressway stands to gain by switching to the local road. As some do so, the numbers on the expressway decrease and $n$ falls. From the left of the intersection, this process converges to $n = 5{,}999$ and, from the right, it converges to 6,000.

If we had used the calculus approach, we would have regarded 1 as a very small increment in relation to $n$ and graphed $P(n)$ instead of $P(n + 1)$. Then the intersection point would have been at $n = 6{,}000$ instead of at 5,999. As you can see, it makes very little difference in practice. What this means is that we can call $n = 6{,}000$ the Nash equilibrium of the route-choice game when choices are

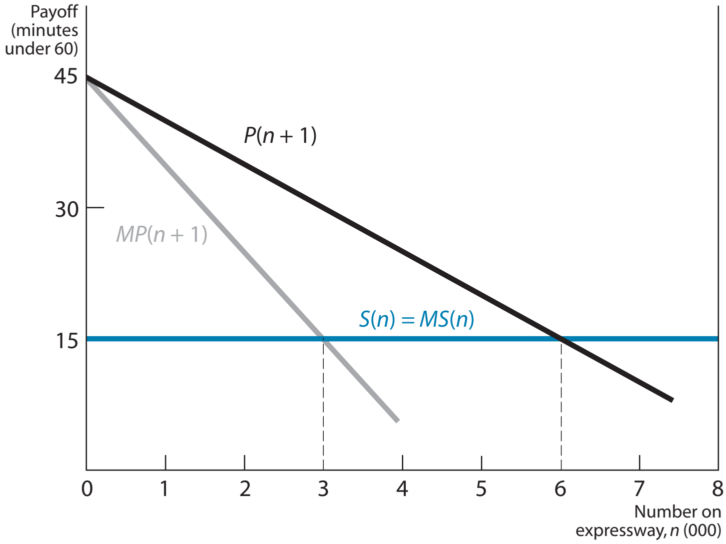



**FIGURE 11.10** Equilibrium and Optimum in Route-Choice Game

governed by purely individual considerations. Given a free choice, 6,000 of the 8,000 total commuters will choose the expressway, and only 2,000 will drive on the local roads.

But we can also interpret the outcome in this game from the perspective of the whole society of commuters. Society benefits from an increase in the number of commuters, $n$, on the expressway when $T(n + 1) - T(n) > 0$ and loses from an increase in $n$ when $T(n + 1) - T(n) < 0$. To figure out how to show this on the graph, we express the idea somewhat differently; we rearrange Eq. (11.1) into two pieces, one depending only on $P$ and the other depending only on $S$:

$$\begin{aligned} T(n+1) - T(n) &= (n+1)\,P(n+1) + [N-(n+1)]\,S(n+1) - nP(n) - [N-n]\,S(n) \\ &= \{P(n+1) + n[P(n+1) - P(n)]\} \\ &\quad - \{S(n) + [N-(n+1)][S(n+1) - S(n)]\}. \end{aligned}$$

The expression in the first set of braces is the effect on the payoffs of the set of commuters who choose P; this expression includes the $P(n + 1)$ of the switcher and the spillover effect, $n[P(n + 1) - P(n)]$, on all the other $n$ commuters who choose P. We call this the marginal social payoff for the P-choosing subgroup, when their number increases from $n$ to $n + 1$, or $MP(n + 1)$ for short. Similarly, the expression in the second set of braces is the marginal social payoff for the S-choosing subgroup, or $MS(n)$ for short. Then, the full expression for $T(n + 1) - T(n)$ tells us that the total social payoff increases when one person switches from S to P (or decreases if the switch is from P to S) if $MP(n + 1) > MS(n)$. The total social payoff decreases when one person switches from S to P (or increases when the switch is from P to S) if $MP(n + 1) < MS(n)$.

Using our expressions for $P(n+1)$ and $S(n)$ in the commuting example, we have

$$MP(n+1) = 45 - (n+1) \times 0.005 + n \times (-0.005) = 44.995 - 0.01n$$

while $MS(n) = 15$ for all values of $n$. Figure 11.10 includes graphs of the relations $MP(n+1)$ and $MS(n)$. Note that the $MS(n)$ coincides with $S(n)$ everywhere because the local roads are never congested. But the $MP(n+1)$ curve lies below the $P(n+1)$ curve. Because of the negative spillover, the social gain from one person's switching to the expressway is less than the private gain to the switcher.

The $MP(n+1)$ and $MS(n)$ curves meet at $n = 2{,}999$, or approximately 3,000. To the left of this intersection, $MP(n+1) > MS(n)$, and society stands to gain by allowing one more person on the expressway. To the right, the opposite is true, and society stands to gain by shifting one person from the expressway to the local roads. Thus, the socially optimal allocation of drivers is 3,000 on the expressway and 3,000 on the local roads.

If you wish to use calculus, you can write the total payoff for the expressway drivers as $nP(n) = n(45 - 0.005n) = 45n - 0.005n^2$. Then $MP(n+1)$ is the derivative of this with respect to $n$—namely, $45 - 0.005 \times 2n = 45 - 0.01n$. The rest of the analysis can proceed as before.

How might this society achieve the optimum allocation of its drivers? Different cultures and political groups use different systems, each with its own merits and drawbacks. The society could simply restrict access to the expressway to 3,000 drivers. But how would it choose those 3,000? It could adopt a first-come, first-served rule, but then drivers would race each other to get there early and waste a lot of time. A bureaucratic society could set up criteria based on complex calculations of needs and merits as defined by civil servants; then everyone will undertake some costly activities to meet these criteria. In a politicized society, the important "swing voters" or organized pressure groups or contributors may be favored. In a corrupt society, those who bribe the officials or the politicians may get the preference. A more egalitarian society could allocate the rights to drive on the expressway by lottery or could rotate them from one month to the next. A scheme that lets you drive only on certain days, depending on the last digit of your car's license plate, is an example. But such a scheme is not so egalitarian as it seems, because the rich can have two cars and choose license-plate numbers that will allow them to drive every day.

Many economists prefer a more open system of charges. Suppose each driver on the expressway is made to pay a tax $t$, measured in units of time. Then the private benefit from using the expressway becomes $P(n) - t$, and the number $n$ in the Nash equilibrium will be determined by $P(n) - t = S(n)$. (Here, we are ignoring the tiny difference between $P(n)$ and $P(n+1)$, which is possible when $N$ is very large.) We know that the socially optimal value of $n$ is 3,000. Using the expressions $P(n) = 45 - 0.005n$ and $S(n) = 15$, and plugging in 3,000 for $n$, we find that $P(n) - t = S(n)$—that is, drivers are indifferent between the expressway and the local roads—when $45 - 15 - t = 15$, or $t = 15$. If we value time at the minimum wage of about \$5 an hour, 15 minutes comes to \$1.25. This is the tax or toll that, when charged, will keep the numbers on the expressway down to what is socially optimal.

Note that when 3,000 drivers are on the expressway, the addition of one more increases the time spent by each of them by 0.005 minute, for a total of 15 minutes. This is exactly the tax that each driver is being asked to pay. In other words, each driver is made to pay the cost of the negative spillover that he imposes on the rest of society. This "brings home" to each driver the extra cost of his action and therefore induces him to take the socially optimal action; economists say the individual person is being made to **internalize the externality**. This idea, that people whose actions hurt others are made to pay for the harm that they cause, adds to the appeal of this approach. But the proceeds from the tax are not used to compensate the others directly. If they were, then each expressway user would count on receiving from others just what he pays, and the whole purpose would be defeated. Instead, the proceeds of the tax go into general government revenues, where they may or may not be used in a socially beneficial manner.

Those economists who prefer to rely on markets argue that if the expressway has a private owner, his profit motive will induce him to charge just enough for its use to reduce the number of users to the socially optimal level. An owner knows that if he charges a tax $t$ for each user, the number of users $n$ will be determined by $P(n) - t = S(n)$. His revenue will be $tn = n[P(n) - S(n)]$, and he will act in such a way as to maximize this revenue. In our example, the revenue is $n[45 - 0.005n - 15] = n[30 - 0.005n] = 30n - 0.005n^2$. It is easy to see this revenue is maximized when $n = 3{,}000$. But in this case, the revenue goes into the owner's pocket; most people regard that as a bad solution.

## D. Positive Spillovers

Many matters pertaining to positive spillovers or positive externalities can be understood simply as mirror images of those for negative spillovers. A person's private benefits from undertaking activities with positive spillovers are less than society's marginal benefits from such activities. Therefore, such actions will be underutilized and their benefits underprovided in the Nash equilibrium. A better outcome can be achieved by augmenting people's incentives; providing those persons whose actions create positive spillovers with a reward just equal to the spillover benefit will achieve the social optimum.

Indeed, the distinction between positive and negative spillovers is to some extent a matter of semantics. Whether a spillover is positive or negative depends on which choice you call P and which you call S. In the commuting example, suppose we called the local roads P and the expressway S. Then one commuter's switch from S to P will reduce the time taken by all the others who choose S, so this action will convey a positive spillover to them. In another example, consider vaccination against some infectious disease. Each person getting vaccinated reduces his own risk of catching the disease (marginal private gain) and reduces the risk of others' getting the disease through him (spillover). If being unvaccinated is called the S action, then getting vaccinated has a positive spillover effect. If remaining unvaccinated is called the P action, then the act of remaining unvaccinated has a negative spillover effect. This has implications for the design of policy to bring individual action into conformity with the social optimum. Society can either reward those who get vaccinated or penalize those who fail to do so.

But actions with positive spillovers can have one very important new feature that distinguishes them from actions with negative spillovers—namely, **positive feedback**. Suppose the spillover effect of your choosing P is to increase the payoff to the others who are also choosing P. Then your choice increases the attraction of that action (P) and may induce some others to take it also, setting in train a process that culminates in everyone's taking that action. Conversely, if very few people are choosing P, then it may be so unattractive that they, too, give it up, leading to a situation in which everyone chooses S. In other words, positive feedback can give rise to multiple Nash equilibria, which we now illustrate by using a very real example.

When you buy a computer, you have to choose between one with a Windows operating system and one with an operating system based on Unix, such as Linux. As the number of Unix users rises, the better it will be to purchase such a computer. The system will have fewer bugs because more users will have detected those that exist, more application software will be available, and more experts will be available to help with any problems that arise. Similarly, a Windows-based computer will be more attractive the more Windows users there are. In addition, many computing aficionados would argue that the Unix system is superior. Without necessarily taking a position on that matter, we show what will happen if that is the case. Will individual choice lead to the socially best outcome?

A diagram similar to Figures 11.6 through 11.8 can be used to show the payoffs to an individual computer purchaser of the two strategies, Unix and Windows. As shown in Figure 11.11, the Unix payoff rises as the number of Unix users rises, and the Windows payoff rises as the number of Unix owners falls (the number of Windows users rises). As already explained, the diagram is drawn assuming that the payoff to Unix users when everyone in the population is a Unix user (at the point labeled U) is higher than the payoff to Windows users when everyone in the population is a Windows user (at W).

If the current population has only a small number of Unix users, then the situation is represented by a point to the left of the intersection of the two payoff

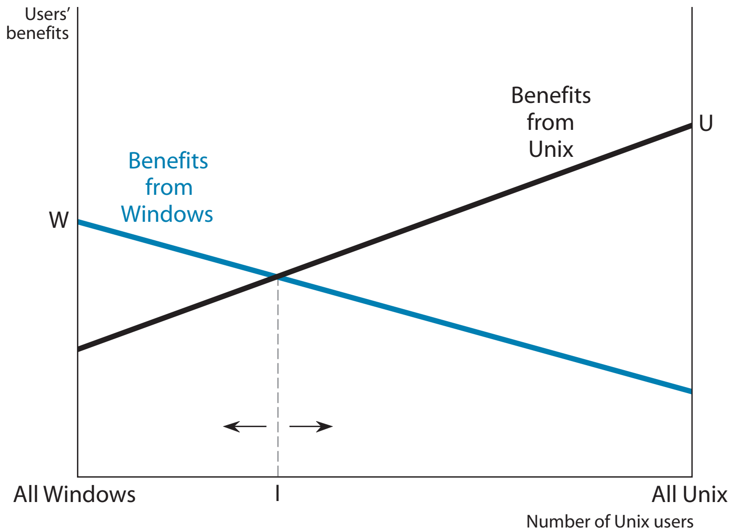



**FIGURE 11.11** Payoffs in Operating-System-Choice Game

lines at I, and each individual user finds it better to choose Windows. When there is a larger number of Unix users in the population, placing the society to the right of I, it is better for each person to choose Unix. Thus, a mixed population of Unix and Windows users is sustainable as an equilibrium only when the current population has exactly I Unix users; only then will no member of the population have any incentive to switch platforms. And even that situation is unstable. Suppose just one person accidentally makes a different decision. If he switches to Windows, his choice will push the population to the left of I, in which case others will have an incentive to switch to Windows, too. If he switches to Unix, the population point moves to the right of I, creating an incentive for more people to switch to Unix. The cumulative effect of these switches will eventually push the society to an all-Unix or an all-Windows outcome; these are the two stable equilibria of the game.[4]

But which of the two stable equilibria will be achieved in this game? The answer depends on where the game starts. If you look at the configuration of today's computer users, you will see a heavily Windows-oriented population. Thus, it seems that because there are so few Unix users (or so many PC users), the world is moving toward the all-Windows equilibrium. Schools, businesses, and private users have become **locked in** to this particular equilibrium as a result of an accident of history. If it is indeed true that Unix provides more benefits to society when used by everyone, then the all-Unix equilibrium should

[4] The term *positive feedback* may create the impression that it is a good thing, but in technical language the term merely characterizes the process and includes no general value judgment about the outcome. In this example, the same positive feedback mechanism could lead to either an all-Unix outcome or an all-Windows outcome; one outcome could be worse than the other.

be preferred over the all-Windows one that we are approaching. Unfortunately, although society as a whole might be better off with the change, no individual computer user has an incentive to make a change from the current situation. Only coordinated action can swing the pendulum toward Unix. A critical mass of individual users, more than I in Figure 11.11, must use Unix before it becomes individually rational for others to choose the same operating system.

There are many examples of similar choices of convention being made by different groups of people. The most famous cases are those in which it has been argued, in retrospect, that a wrong choice was made. Advocates claim that steam power could have been developed for greater efficiency than gasoline; it certainly would have been cleaner. Proponents of the Dvorak typewriter/computer keyboard configuration claim that it would be better than the QWERTY keyboard if used everywhere. Many engineers agree that Betamax had more going for it than VHS in the video recorder market. In such cases, the whims of the public or the genius of advertisers help determine the ultimate equilibrium and may lead to a "bad" or "wrong" outcome from society's perspective. Other situations do not suffer from such difficulties. Few people concern themselves with fighting for a reconfiguration of traffic-light colors, for example.[5]

The ideas of positive feedback and lock-in find an important application in macroeconomics. Production is more profitable the higher the level of demand in the economy, which happens when national income is higher. In turn, income is higher when firms are producing more and are therefore hiring more workers. This positive feedback creates the possibility of multiple equilibria, of which the high-production, high-income one is better for society, but individual decisions may lock the economy into the low-production, low-income equilibrium. The better equilibrium could be turned into a focal point by public declaration—"the only thing we have to fear is fear itself"—but the government can also inject demand into the economy to the extent necessary to move it to the better equilibrium. In other words, the possibility of unemployment due to a deficiency of aggregate demand—as discussed in the supply-and-demand language of economic theory by the British economist John Maynard Keynes in his well-known 1936 book titled *Employment, Interest, and Money*—can be seen from a game-theoretic perspective as the result of a failure to solve a collective-action problem.[6]

[5] Not everyone agrees that the Dvorak keyboard and the Betamax video recorder were clearly superior alternatives. See two articles by S. J. Liebowitz and Stephen E. Margolis, "Network Externality: An Uncommon Tragedy," *Journal of Economic Perspectives*, vol. 8 (Spring 1994), pp. 146–49, and "The Fable of the Keys," *Journal of Law and Economics*, vol. 33 (April 1990), pp. 1–25.

[6] John Maynard Keynes, *Employment, Interest, and Money* (London: Macmillan, 1936). See also John Bryant, "A Simple Rational-Expectations Keynes-type Model," *Quarterly Journal of Economics*, vol. 98 (1983), pp. 525–28, and Russell Cooper and Andrew John, "Coordination Failures in a Keynesian Model," *Quarterly Journal of Economics*, vol. 103 (1988), pp. 441–63, for formal game-theoretic models of unemployment equilibria.

# 4 A BRIEF HISTORY OF IDEAS

## A. The Classics

The problem of collective action has been recognized by social philosophers and economists for a very long time. The seventeenth-century British philosopher Thomas Hobbes argued that society would break down in a "war of all against all" unless it was ruled by a dictatorial monarch, or *Leviathan* (the title of his book). One hundred years later, the French philosopher Jean-Jacques Rousseau described the problem of a prisoners' dilemma in his *Discourse on Inequality.* A stag hunt needs the cooperation of the whole group of hunters to encircle and kill the stag, but any individual hunter who sees a hare may find it better for himself to leave the circle to chase the hare. But Rousseau thought that such problems were the product of civilization and that people in the natural state lived harmoniously as "noble savages." At about the same time, two Scots pointed out some dramatic solutions to such problems. David Hume in his *Treatise on Human Nature* argued that the expectations of future returns of favors can sustain cooperation. Adam Smith's *Wealth of Nations* developed a grand vision of an economy in which the production of goods and services motivated purely by private profit could result in an outcome that was best for society as a whole.[7]

The optimistic interpretation persisted, especially among many economists and even several political scientists, to the point where it was automatically assumed that if an outcome was beneficial to a group as a whole, the actions of its members would bring the outcome about. This belief received a necessary rude shock in the mid-1960s when Mancur Olson published *The Logic of Collective Action.* He pointed out that the best collective outcome would not prevail unless it was in each individual person's private interest to perform his assigned action—that is, unless it was a Nash equilibrium. However, Olson did not specify the collective-action game very precisely. Although it looked like a prisoners' dilemma, Olson insisted that it was not necessarily so, and we have

[7] The great old books cited in this paragraph have been reprinted many times in many different versions. For each, we list the year of original publication and the details of one relatively easily accessible reprint. In each case, the editor of the reprinted version provides an introduction that conveniently summarizes the main ideas. Thomas Hobbes, *Leviathan; or the Matter, Form, and Power of Commonwealth Ecclesiastical and Civil*, 1651 (Everyman Edition, London: J. M. Dent, 1973); David Hume, *A Treatise of Human Nature*, 1739 (Oxford: Clarendon Press, 1976); Jean-Jacques Rousseau, *A Discourse on Inequality*, 1755 (New York: Penguin Books, 1984); Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776 (Oxford: Clarendon Press, 1976).

already seen that the problem can also take the form of a chicken game or an assurance game.[8]

Another major class of collective-action problems—namely, those concerning the depletion of common-access resources—received attention at about the same time. If a resource such as a fishery or a meadow is open to all, each user will exploit it as much as he can, because any self-restraint on his part will merely make more available for the others to exploit. As we mentioned earlier, Garrett Hardin wrote a well-known article on this subject titled "The Tragedy of the Commons." Common-resource problems are unlike our irrigation-project game, in which each person has a strong private incentive to free-ride off the efforts of others. In regard to a common resource, each person has a strong private incentive to exploit it to the full, making everyone else pay the social cost that results from the degradation of the resource.

## B. Modern Approaches and Solutions

Until recently, many social scientists and most physical scientists took a Hobbesian line on the common-resource problem, arguing that it can be solved only by a government that forces everyone to behave cooperatively. Others, especially economists, retained their Smithian optimism. They argued that placing the resource in proper private ownership, where its benefits can be captured in the form of profit by the owner, will induce the owner to restrain its use in a socially optimal manner. He will realize that the value of the resource (fish or grass, for example) may be higher in the future because less will be available, and therefore he can make more profit by saving some of it for that future.

Nowadays, thinkers from all sides have begun to recognize that collective-action problems come in diverse forms and that there is no uniquely best solution to all of them. They also understand that groups or societies do not stand helpless in the face of such problems, and they devise various ways to cope with them. Much of this work has been informed by game-theoretic analysis of repeated prisoners' dilemmas and similar games.[9]

Solutions to collective-action problems of all types must induce individual persons to act cooperatively or in a manner that would be best for the group, even though the person's interests may best be served by doing something else—in particular, taking advantage of the others' cooperative behavior.[10] Humans

[8] Mancur Olson, *The Logic of Collective Action* (Cambridge, Mass.: Harvard University Press, 1965).

[9] Prominent in this literature are Michael Taylor, *The Possibility of Cooperation* (New York: Cambridge University Press, 1987); Elinor Ostrom, *Governing the Commons* (New York: Cambridge University Press, 1990); and Matt Ridley, *The Origins of Virtue* (New York: Viking Penguin, 1996).

[10] The problem of the need to attain cooperation and its solutions are not unique to human societies. Examples of cooperative behavior in the animal kingdom have been explained by biologists in terms of the advantage of the gene and of the evolution of instincts. For more, see Chapter 12 and Ridley, *Origins of Virtue*.

exhibit much in the way of cooperative behavior. The act of reciprocating gifts and skills at detecting cheating are so common in all societies and throughout history, for example, that there is reason to argue that they may be instincts.[11] But human societies generally rely heavily on purposive social and cultural customs, **norms**, and **sanctions** in inducing cooperative behavior from their individual members. These methods are conscious, deliberate attempts to design the game in order to solve the collective-action problem.[12] We approach the matter of solution methods from the perspective of the type of game being played.

A solution is easiest if the collective-action problem takes the form of an assurance game. Then it is in every person's private interest to take the socially best action if he expects all other persons to do likewise. In other words, the socially optimal outcome is a Nash equilibrium. The only problem is that the same game has other, socially worse, Nash equilibria. Then all that is needed to achieve the best Nash equilibrium and thereby the social optimum is to make it a focal point—that is, to ensure the convergence of the players' expectations on it. Such a convergence can result from a social **custom**, or **convention**—namely, a mode of behavior that finds automatic acceptance because it is in everyone's interest to follow it so long as others are expected to do likewise. For example, if all the farmers, herders, weavers, and other producers in an area want to get together to trade their wares, all they need is the assurance of finding others with whom to trade. Then the custom that the market is held in village X on day Y of every week makes it optimal for everyone to be there on that day.[13]

[11] See Ridley, *Origins of Virtue*, ch. 6 and ch. 7.

[12] The social sciences do not have precise and widely accepted definitions of terms such as *custom* and *norm*; nor are the distinctions among such terms always clear and unambiguous. We set out some definitions in this section, but be aware that you may find different usage in other books. Our approach is similar to those found in Richard Posner and Eric Rasmusen, "Creating and Enforcing Norms, with Special Reference to Sanctions," *International Review of Law and Economics*, vol. 19, no. 3 (September 1999), pp. 369–82, and in David Kreps, "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, Papers and Proceedings, vol. 87, no. 2 (May 1997), pp. 359–64; Kreps uses the term *norm* for all the concepts that we classify under different names.

Sociologists have a different taxonomy of norms from that of economists; it is based on the importance of the matter (trivial matters such as table manners are called *folkways*, and weightier matters are called *mores*), and on whether the norms are formally codified as *laws*. They also maintain a distinction between *values* and norms, recognizing that some norms may run counter to persons' values and therefore require sanctions to enforce them. This distinction corresponds to ours between customs, internalized norms, and enforced norms. The conflict between individual values and social goals arises for enforced norms but not for customs or *conventions*, as we label them, or for internalized norms. See Donald Light and Suzanne Keller, *Sociology*, 4th ed. (New York: Knopf, 1987), pp. 57–60.

[13] In his study of the emergence of cooperation, *Cheating Monkeys and Citizen Bees* (New York: Free Press, 1999), the evolutionary biologist Lee Dugatkin labels this case "selfish teamwork." He argues that such behavior is likelier to arise in times of crisis, because each person is pivotal at those times. In a crisis, the outcome of the group interaction is likely to be disastrous for everyone if even one person fails to contribute to the group's effort to get out of the dire situation. Thus, each person is willing to contribute so long as the others do. We will mention Dugatkin's full classification of alternative approaches to cooperation in Chapter 12 on evolutionary games.

One complication remains. For the desired outcome to be a focal point, each person must have confidence that all others understand it, which in turn requires that they have confidence that all others understand. . . . In other words, the point must be common knowledge. Usually, some prior social action is necessary to ensure that this is true. Publication in a medium that is known by everyone to be sufficiently widely read, and discussion in an inward-facing circle so everyone knows that everyone else was present and paying attention, are some methods used for this purpose.[14]

Our analysis in Section 2 suggested that individual payoffs are often configured in such a way that collective-action problems, particularly of large groups, take the form of a prisoners' dilemma. Not surprisingly, the methods for coping with such problems have received the most attention.

The simplest method attempts to change people's preferences so that the game is no longer a prisoners' dilemma. If individuals get sufficient pleasure from cooperating, or suffer enough guilt or shame when they cheat, they will cooperate to maximize their own payoffs. If the extra payoff from cooperation is conditional—one gets pleasure from cooperating or guilt or shame from cheating if, but only if, many others are cooperating—then the game can turn into an assurance game. In one of its equilibria, everyone cooperates because everyone else does, and in the other, no one cooperates because no one else does. Then the collective-action problem is the simpler one of making the better equilibrium the focal point. If the extra payoff from cooperation is unconditional—one gets pleasure from cooperating or guilt or shame from cheating regardless of what the others do—then the game can have a unique equilibrium where everyone cooperates. In many situations, it is not even necessary for everyone to have such payoffs. If a substantial proportion of the population does, that may suffice for the desired collective outcome.

Some such prosocial preferences may be innate, hard wired in a biological evolutionary process. But they are more likely to be social or cultural products. Most societies make deliberate efforts to instill prosocial thinking in children during the process of socialization in families and schools. Growth of such preferences is seen in experiments on ultimatum and dictator games of the kind we discussed in Chapter 3. When these experiments are conducted on children of different ages, very young children behave selfishly. By age eight, however, they develop a significant sense of equality. True prosocial preferences develop gradually thereafter, with some relapses, finally to an adult fair-mindedness. Thus, a

[14] See Michael Chwe, *Rational Ritual: Culture, Coordination, and Common Knowledge* (Princeton: Princeton University Press, 2001), for a discussion of this issue and numerous examples and applications of it.

long process of education and experience instills *internalized norms* into people's preferences.[15]

However, people do differ in the extent to which they internalize prosocial preferences, and the process may not go far enough to solve many collective-action problems. Most people have sufficiently broad understanding of what the socially cooperative action is in most situations, but individuals retain the personal temptation to cheat. Therefore, a system of external sanctions or punishments is needed to sustain the cooperative actions. We call these widely understood but not automatically followed rules of behavior *enforced norms*.

In Chapter 10, we described in detail several methods for achieving a cooperative outcome in prisoners' dilemma games, including repetition, penalties (or rewards), and leadership. In that discussion, we were mainly concerned with two-person dilemmas. The same methods apply to enforcement of norms in collective-action problems in large groups, with some important modifications or innovations.

We saw in Chapter 10 that repetition was the most prominent of these methods; so we focus the most attention on it. Repetition can achieve cooperative outcomes as equilibria of individual actions in a repeated two-person prisoners' dilemma by holding up the prospect that cheating will lead to a breakdown of cooperation. More generally, what is needed to maintain cooperation is the expectation in the mind of each player that his personal benefits from cheating are transitory and that they will quickly be replaced by a payoff lower than that associated with cooperative behavior. For players to believe that cheating is not beneficial from a long-term perspective, cheating should be detected quickly, and the punishment that follows (reduction in future payoffs) should be sufficiently swift, sure, and painful.

A group has one advantage in this respect over a pair of individual persons. The same pair may not have occasion to interact all that frequently, but each of them is likely to interact with *someone* in the group all the time. Therefore, B's temptation to cheat A can be countered by his fear that others, such as C, D, and so on, whom he meets in the future will punish him for this action. An extreme case where bilateral interactions are not repeated and punishment must be inflicted on one's behalf by a third party is, in Yogi Berra's well-known saying, "Always go to other people's funerals. Otherwise they won't go to yours."

But a group has some offsetting disadvantages over direct bilateral interaction when it comes to sustaining good behavior in repeated interactions. The required speed and certainty of detection and punishment suffer as the numbers

[15] Colin Camerer, *Behavioral Game Theory* (Princeton: Princeton University Press, 2003), pp. 65–67. See also pp. 63–75 for an account of differences in prosocial behavior along different dimensions of demographic characteristics and across different cultures.

in the group increase. One sees many instances of successful cooperation in small village communities that would be unimaginable in a large city or state.

Start with the detection of cheating, which is never easy. In most real situations, payoffs are not completely determined by the players' actions but are subject to some random fluctuations. Even with two players, if one gets a low payoff, he cannot be sure that the other cheated; it may have been just a bad draw of the random shock. With more people, an additional question enters the picture: If someone cheated, who was it? Punishing someone without being sure of his guilt beyond a reasonable doubt is not only morally repulsive but also counterproductive. The incentive to cooperate gets blunted if even cooperative actions are susceptible to punishment by mistake.

Next, with many players, even when cheating is detected and the cheater identified, this information has to be conveyed sufficiently quickly and accurately to others. For this, the group must be small or else must have a good communication or gossip network. Also, members should not have much reason to accuse others falsely.

Finally, even after cheating is detected and the information spread to the whole group, the cheater's punishment—enforcement of the social norm—has to be arranged. A third person often has to incur some personal cost to inflict such punishment. For example, if C is called on to punish B, who had previously cheated A, C may have to forgo some profitable business that he could have transacted with B. Then the inflicting of punishment is itself a collective-action game and suffers from the same temptation to "shirk," that is, not to participate in the punishment. A society could construct a second-round system of punishments for shirking, but that in turn may be yet another collective-action problem! However, humans seem to have evolved an instinct whereby people get some personal pleasure from punishing cheaters even when they have not themselves been the victims of this particular act of cheating.[16] Interestingly, the notion that "one should impose sanctions, even at personal cost, on violators of enforced social norms" seems itself to have become an internalized norm.[17]

Norms are reinforced by observation of society's general adherence to them, and they lose their force if they are frequently seen to be violated. Before the

[16] For evidence of such altruistic punishment instinct, see Ernst Fehr and Simon Gächter, "Altruistic Punishment in Humans," *Nature*, vol. 415 (January 10, 2002), pp. 137–40.

[17] Our distinction between internalized norms and enforced norms is similar to Kreps's distinction between functions (iii) and (iv) of norms (Kreps, "Intrinsic Motivation and Extrinsic Incentives," p. 359). Society can also reward desirable actions just as it can punish undesirable ones. Again, the rewards, financial or otherwise, can be given externally, or players' payoffs can be changed so that they take pleasure in doing the right thing. The two types of rewards can interact; for example, the peerages and knighthoods given to British philanthropists and others who do good deeds for British society are external rewards, but individual persons value them only because respect for knights and peers is a British social norm.

advent of the welfare state, when those who fell on hard economic times had to rely on help from family or friends or their immediate small social group, the work ethic constituted a norm that held in check the temptation to slacken one's own efforts and become a free rider on the support of others. As government took over the supporting role and unemployment compensation or welfare became an entitlement, this norm of the work ethic weakened. After the sharp increases in unemployment in Europe in the late 1980s and early 1990s, a significant fraction of the population became users of the official support system, and the norm weakened even further.[18]

Different societies or cultural groups may develop different conventions and norms to achieve the same purpose. At the trivial level, each culture has its own set of good manners—ways of greeting strangers, indicating approval of food, and so on. When two people from different cultures meet, misunderstandings can arise. More important, each company or office has its own ways of getting things done. The differences between these customs and norms are subtle and difficult to pin down, but many mergers fail because of a clash of these "corporate cultures."

Next, consider the chicken form of collective-action games. Here, the nature of the remedy depends on whether the largest total social payoff is attained when everyone participates (what we called "chicken version I" in Section 1.B) or when some cooperate and others are allowed to shirk (chicken II). For chicken I, where everyone has the individual temptation to shirk, the problem is much like that of sustaining cooperation in the prisoners' dilemma, and all the earlier remarks for that game apply here, too. Chicken II is different—easier in one respect and harder in another. Once an assignment of roles between participants and shirkers is made, no one has the private incentive to switch: if the other driver is assigned the role of going straight, then you are better off swerving, and the other way around. Therefore, if a custom creates the expectation of an equilibrium, it can be maintained without further social intervention such as sanctions. However, in this equilibrium, the shirkers get higher payoffs than the participants do, and this inequality can create its own problems for the game; the conflicts and tensions, if they are major, can threaten the whole fabric of the society. Often the problem can be solved by repetition. The roles of participants and shirkers can be rotated to equalize payoffs over time.

Sometimes the problem of differential payoffs in version II of the prisoners' dilemma or chicken is "solved," not by restoring equality but by **oppression** or **coercion**, which forces a dominated subset of society to accept the lower payoff and allows the dominant subgroup to enjoy the higher payoff. In many societies throughout history, the work of handling animal carcasses was forced on

[18] Assar Lindbeck, "Incentives and Social Norms in Household Behavior," *American Economic Review*, Papers and Proceedings, vol. 87, no. 2 (May 1997), pp. 370–77.

particular groups or castes in this way. The history of the maltreatment of racial and ethnic minorities and of women provides vivid examples of such practices. Once such a system becomes established, no one member of the oppressed group can do anything to change the situation. The oppressed must get together as a group and act to change the whole system, itself another problem of collective action.

Finally, consider the role of leadership in solving collective-action problems. In Chapter 10, we pointed out that, if the players are of very unequal "size," the prisoners' dilemma may disappear because it may be in the private interests of the larger player to continue cooperation and to accept the cheating of the smaller player. Here we recognize the possibility of a different kind of bigness—namely, having a "big heart." People in most groups differ in their preferences, and many groups have one or a few who take genuine pleasure in expending personal effort to benefit the whole. If there are enough such people for the task at hand, then the collective-action problem disappears. Most schools, churches, local hospitals, and other worthy causes rely on the work of such willing volunteers. This solution, like others before it, is more likely to work in small groups, where the fruits of their actions are more closely and immediately visible to the benefactors, who are therefore encouraged to continue.

## C. Applications

In her book *Governing the Commons*, Elinor Ostrom describes several examples of resolution of common-resource problems at local levels. Most of them require taking advantage of features specific to the context in order to set up systems of detection and punishment. A fishing community on the Turkish coast, for example, assigns and rotates locations to its members; the person who is assigned a good location on any given day will naturally observe and report any intruder who tries to usurp his place. Many other users of common resources, including the grazing commons in medieval England, actually restricted access and controlled overexploitation by allocating complex, tacit, but well-understood rights to individual persons. In one sense, this solution bypasses the common-resource problem by dividing up the resource into a number of privately owned subunits.

The most striking feature of Ostrom's range of cases is their immense variety. Some of the prisoners' dilemmas of the exploitation of common-property resources that she examined were solved by private initiative by the group of people actually in the dilemma; others were solved by external public or governmental intervention. In some instances, the dilemma was not resolved at all, and the group remained trapped in the all-shirk outcome. Despite this variety, Ostrom identifies several common features that make it easier to solve prisoners' dilemmas of collective action: (1) it is essential to have an identifiable and stable group of potential participants; (2) the benefits of cooperation have to be large enough to make it worth paying all the costs of monitoring and enforcing the rules of cooperation; and (3) it is very important that the members of the group can communicate with each other. This last feature accomplishes several things. First, it makes the norms clear—everyone knows what behavior is expected, what kind of cheating will not be tolerated, and what sanctions will be imposed on cheaters. Next, it spreads information about the efficacy of the detection of the cheating mechanism, thereby building trust and removing the suspicion that each participant might hold that he is abiding by the rules while others are getting away with breaking them. Finally, it enables the group to monitor the effectiveness of the existing arrangements and to improve on them as necessary. All these requirements look remarkably like those identified in Chapter 10 from our theoretical analysis of the prisoners' dilemma and from the observations of Axelrod's tournaments.

Ostrom's study of the fishing village also illustrates what can be done if the collective optimum requires different persons to do different things, in which case some get higher payoffs than others. In a repeated relationship, the advantageous position can rotate among the participants, thereby maintaining some sense of equality over time.

Ostrom finds that an external enforcer of cooperation may not be able to detect cheating or impose punishment with sufficient clarity and swiftness. Thus, the frequent reaction that centralized or government policy is needed to solve collective-action problems is often proved wrong. Another example comes from village communities or "communes" in late-nineteenth-century Russia. These communities solved many collective-action problems of irrigation, crop rotation, management of woods and pastures, and road and bridge construction and repair in just this way. "The village . . . was not the haven of communal harmony. . . . It was simply that the individual interests of the peasants were often best served by collective activity." Reformers of early twentieth-century czarist governments and Soviet revolutionaries of the 1920s alike failed, partly because the old system had such a hold on the peasants' minds that they resisted anything new, but also because the reformers failed to understand the role that some of the prevailing practices played in solving collective-action problems and thus failed to replace them with equally effective alternatives.[19]

The difference between small and large groups is well illustrated by Avner Greif's comparison of two groups of traders in countries around the Mediterranean Sea in medieval times. The Maghribis were Jewish traders who relied on extended family and social ties. If one member of this group cheated another,

[19] Orlando Figes, *A People's Tragedy: The Russian Revolution 1891–1924* (New York: Viking Penguin, 1997), pp. 89–90, 240–41, 729–30. See also Ostrom, *Governing the Commons*, p. 23, for other instances where external, government-enforced attempts to solve common-resource problems actually made them worse.

the victim informed all the others by writing letters. When guilt was convincingly proved, no one in the group would deal with the cheater. This system worked well on a small scale of trade. But as trade expanded around the Mediterranean, the group could not find sufficiently close or reliable insiders to go to the countries with the new trading opportunities.

In contrast, the Genoese traders established a more official legal system. A contract had to be registered with the central authorities in Genoa. The victim of any cheating or violation of the contract had to take a complaint to the authorities, who carried out the investigation and imposed the appropriate fines on the cheater. This system, with all its difficulties of detection, could be more easily expanded with the expansion of trade.[20] As economies grow and world trade expands, we see a similar shift from tightly linked groups to more arm's-length trading relationships and from enforcement based on repeated interactions to that of the official law.

The idea that small groups are more successful at solving collective-action problems forms the major theme of Olson's *Logic of Collective Action* (see footnote 8) and has led to an insight important in political science. In a democracy, all voters have equal political rights, and the majority's preference should prevail. But we see many instances in which this does not happen. The effects of policies are generally good for some groups and bad for others. To get its preferred policy adopted, a group has to take political action—lobbying, publicity, campaign contributions, and so on. To do these things, the group must solve a collective-action problem, because each member of the group may hope to shirk and enjoy the benefits that the others' efforts have secured. If small groups are better able to solve this problem, then the policies resulting from the political process will reflect *their* preferences, even if other groups who fail to organize are more numerous and suffer greater losses than the successful groups' gains.

The most dramatic example of policies reflecting the preferences of the organized group comes from the arena of trade policy. A country's import restrictions help domestic producers whose goods compete with these imports, but they hurt the consumers of the imported goods and the domestic competing goods alike, because prices for these goods are higher than they would be otherwise. The domestic producers are few in number, and the consumers are almost the whole population; the total dollar amount of the consumers' losses is typically far bigger than the total dollar amount of the producers' gains. Political considerations based on constituency membership numbers and economic considerations of dollar gains and losses alike would lead us to expect a consumer victory in this policy arena; we would expect to see at least a push for the

[20] Avner Greif, "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy*, vol. 102, no. 5 (October 1994), pp. 912–50.

idea that import restrictions should be abolished, but we don't. The smaller and more tightly knit associations of producers are better able to organize for political action than the numerous, dispersed consumers.

More than 70 years ago, the American political scientist E. E. Schattschneider provided the first extensive documentation and discussion of how pressure politics drives trade policy. He recognized that "the capacity of a group for organization has a great influence on its activity," but he did not develop any systematic theory of what determines this capacity.[21] The analysis of Olson and others has improved our understanding of the issue, but the triumph of pressure politics over economics persists in trade policy to this day. For example, in the late 1980s, the U.S. sugar policy cost each of the 240 million people in the United States about $11.50 per year for a total of about $2.75 billion, while it increased the incomes of about 10,000 sugar-beet farmers by about $50,000 each, and the incomes of 1,000 sugarcane farms by as much as $500,000 each, for a total of about $1 billion. The net loss to the U.S. economy was $1.75 billion.[22] Each of the unorganized consumers continues to bear his small share of the costs in silence; many of them are not even aware that each is paying $11.50 a year too much for his sweet tooth.

If this overview of the theory and practice of solving collective-action problems seems diverse and lacking a neat summary statement, that is because the problems are equally diverse, and the solutions depend on the specifics of each problem. The one general lesson that we can provide is the importance of letting the participants themselves devise solutions by using their local knowledge of the situation, their advantage of proximity in monitoring the cooperative or shirking actions of others in the community, and their ability to impose sanctions on shirkers by exploiting various ongoing relationships within the social group.

Finally, a word of caution. You might be tempted to come away from this discussion of collective-action problems with the impression that individual freedom always leads to harmful outcomes that can and must be improved by social norms and sanctions. Remember, however, that societies face problems other than those of collective action; some of them are better solved by individual initiative than by joint efforts. Societies can often get hidebound and autocratic, becoming trapped in their norms and customs and stifling the innovation that is so often the key to economic growth. Collective action can become collective inaction.[23]

[21] E. E. Schattschneider, *Politics, Pressures, and the Tariff* (New York: Prentice-Hall, 1935); see especially pp. 285–86.

[22] Stephen V. Marks, "A Reassessment of the Empirical Evidence on the U.S. Sugar Program," in *The Economics and Politics of World Sugar Policies*, ed. Stephen V. Marks and Keith E. Maskus (Ann Arbor: University of Michigan Press, 1993), pp. 79–108.

[23] David Landes, *The Wealth and Poverty of Nations* (New York: W. W. Norton & Company, 1998), ch. 3 and ch. 4, makes a spirited case for this effect.

## 5 "HELP!": A GAME OF CHICKEN WITH MIXED STRATEGIES

In the chicken variant of collective-action problems discussed in earlier sections, we looked only at the pure-strategy equilibria. But we know from Chapter 7 that such games have mixed-strategy equilibria, too. In collective-action problems, where each participant is thinking, "It is better if I wait for enough others to participate so that I can shirk; but then again, maybe they won't, in which case I should participate," mixed strategies nicely capture the spirit of such vacillation. Our last story is a dramatic, even chilling application of such a mixed-strategy equilibrium.

In 1964 in New York City (in Kew Gardens, Queens), a woman named Kitty Genovese was killed in a brutal attack that lasted more than half an hour. She screamed through it all and, although her screams were heard by many people and at least 3 actually witnessed some part of the attack, no one went to help her or even called the police.

The story created a sensation and found several ready theories to explain it. The press and most of the public saw this episode as a confirmation of their belief that New Yorkers—or big-city dwellers or Americans or people more generally—were just apathetic or didn't care about their fellow human beings.

However, even a little introspection or observation will convince you that people do care about the well-being of other humans, even strangers. Social scientists offered a different explanation for what happened, which they labeled **pluralistic ignorance**. The idea behind this explanation is that no one can be sure about what is happening or whether help is really needed and how much. People look to each other for clues or guidance about these matters and try to interpret other people's behavior in this light. If they see that no one else is doing anything to help, they interpret it as meaning that help is probably not needed, and so they don't do anything either. This explanation has some intuitive appeal but is unsatisfactory in the Kitty Genovese context. There is a very strong presumption that a screaming woman needs help. What did the onlookers think—that a movie was being shot in their obscure neighborhood? If so, where were the lights, the cameras, the director, other crew?

A better explanation would recognize that although each onlooker may experience strong personal loss from Kitty's suffering and get genuine personal pleasure if she were saved, each must balance that against the cost of getting involved. You may have to identify yourself if you call the police; you may then have to appear as a witness, and so on. Thus, we see that each person may prefer to wait for someone else to call and hope to get for himself the free rider's benefit of the pleasure of a successful rescue.

Social psychologists have a slightly different version of this idea of free riding, which they label **diffusion of responsibility**. In this version, the idea is that everyone might agree that help is needed, but they are not in direct communication with each other and so cannot coordinate on who should help. Each person may believe that help is someone else's responsibility. And the larger the group, the more likely it is that each person will think that someone else would probably help, and therefore he can save himself the trouble and the cost of getting involved.

Social psychologists conducted some experiments to test this hypothesis. They staged situations in which someone needed help of different kinds in different places and with different-sized crowds. Among other things, they found that the larger the size of the crowd, the less likely was help to come forth.

The concept of diffusion of responsibility seems to explain this finding, but not quite completely. It claims that the larger the crowd, the less likely is any one person to help. But there are more people, and only one person is needed to act and call the police to secure help. To make it less likely that even one person helps, the chance of any one person helping has to decrease sufficiently fast to offset the increase in the total number of potential helpers. To find out whether it does so requires game-theoretic analysis, which we now supply.[24]

We consider only the aspect of diffusion of responsibility in which action is not consciously coordinated, and we leave aside all other complications of information and inference. Thus, we assume that everyone believes the action is needed and is worth the cost.

Suppose $N$ people are in the group. The action brings each of them a benefit $B$. Only one person is needed to take the action; more are redundant. Anyone who acts bears the cost $C$. We assume that $B > C$; so it is worth any one person's while to act even if no one else is acting. Thus, the action is justified in a very strong sense.

The problem is that anyone who takes the action gets the value $B$ and pays the cost $C$ for a net payoff of $(B - C)$, whereas he would get the higher payoff $B$ if someone else took the action. Thus, each person has the temptation to let someone else go ahead and to become a free rider on another's effort. When all $N$ people are thinking thus, what will be the equilibrium or outcome?

If $N = 1$, the single person has a simple decision problem rather than a game. He gets $B - C > 0$ if he takes the action and 0 if he does not. Therefore, he goes ahead and helps.

[24] For a fuller account of the Kitty Genovese story and for the analysis of such situations from the perspective of social psychology, see John Sabini, *Social Psychology*, 2nd ed. (New York: W. W. Norton & Company, 1995), pp. 39–44. Our game-theoretic model is based on Thomas Palfrey and Howard Rosenthal, "Participation and the Provision of Discrete Public Goods," *Journal of Public Economics*, vol. 24 (1984), pp. 171–93. Many purported facts of the story have been recently challenged in *Kitty Genovese: The Murder, the Bystanders, and the Crime that Changed America* by Kevin Cook (New York: W. W. Norton & Company, 2014), but the power and impact of the originally reported story on American thinking about urban crime remains, and it is still a good example for game-theoretic analysis.

If $N > 1$, we have a game of strategic interaction with several equilibria. Let us begin by ruling out some possibilities. With $N > 1$, there cannot be a pure-strategy Nash equilibrium in which all people act, because then any one of them would do better by switching to free ride. Likewise, there cannot be a pure-strategy Nash equilibrium in which no one acts, because *given that no one else is acting* (remember that under the Nash assumption each player takes the others' strategies as given), it pays any one person to act.

There *are* Nash equilibria where exactly one person acts; in fact, there are $N$ such equilibria, one corresponding to each member. But when everyone is making the decision individually in isolation, there is no way to coordinate and designate who is to act. Even if members of the group were to attempt such coordination, they might try to negotiate over the responsibility and not reach a conclusion, at least not in time to be of help. Therefore, it is of interest to examine symmetric equilibria in which all members have identical strategies.

We already saw that there cannot be an equilibrium in which all $N$ people follow the same pure strategy. Therefore, we should see whether there can be an equilibrium in which they all follow the same mixed strategy. Actually, mixed strategies are quite appealing in this context. The people are isolated, and each is trying to guess what the others will do. Each is thinking: Perhaps I should call the police . . . but maybe someone else will . . . but what if they don't . . . ? Each breaks off this process at some point and does the last thing that he thought of in this chain, but we have no good way of predicting what that last thing is. A mixed strategy carries the flavor of this idea of a chain of guesswork being broken at a random point.

So suppose $P$ is the probability that any one person will not act. If one particular person is willing to mix strategies, he must be indifferent between the two pure strategies of acting and not acting. Acting gets him $(B - C)$ for sure. Not acting will get him 0 if none of the other $(N - 1)$ people act and $B$ if at least one of them does act. Because the probability that any one person fails to act is $P$ and because they are deciding independently, the probability that none of the $(N - 1)$ others acts is $P^{N-1}$, and the probability that at least one does act is $(1 - P^{N-1})$. Therefore, the expected payoff to the one person when he does not act is

$$0 \times P^{N-1} + B(1 - P^{N-1}) = B(1 - P^{N-1}).$$

And that one person is indifferent between acting and not acting when

$$B - C = B(1 - P^{N-1}) \quad \text{or when} \quad P^{N-1} = \frac{C}{B} \quad \text{or} \quad P = \left(\frac{C}{B}\right)^{1/(N-1)}.$$

Note how this indifference condition of *one* selected player determines the probability with which the *other* players mix their strategies.

Having obtained the equilibrium mixture probability, we can now see how it changes as the group size $N$ changes. Remember that $C/B < 1$. As $N$ increases from 2 to infinity, the power $1/(N - 1)$ decreases from 1 to 0. Then $C/B$ raised to this power—namely, $P$—increases from $C/B$ to 1. Remember that $P$ is the probability that any one person does not take the action. Therefore, the probability of action by any one person—namely, $(1 - P)$—falls from $1 - C/B = (B - C)/B$ to 0.[25]

In other words, the more people there are, the less likely is any one of them to act. This is intuitively true, and in good conformity with the idea of diffusion of responsibility. But it does not yet give us the conclusion that help is less likely to be forthcoming in a larger group. As we said before, help requires action by only one person. Because there are more and more people, each of whom is less and less likely to act, we cannot conclude immediately that the probability of *at least one* of them acting gets smaller. More calculation is needed to see whether this is the case.

Because the $N$ persons are randomizing independently in the Nash equilibrium, the probability $Q$ that *not even one* of them helps is

$$Q = P^N = \left(\frac{C}{B}\right)^{N/(N-1)}.$$

As $N$ increases from 2 to infinity, $N/(N - 1)$ decreases from 2 to 1, and then $Q$ increases from $(C/B)^2$ to $C/B$. Correspondingly, the probability that *at least one* person helps—namely $(1 - Q)$—decreases from $1 - (C/B)^2$ to $1 - C/B$.[26]

So our exact calculation does bear out the hypothesis: the larger the group, the *less* likely is help to be given at all. The probability of provision does not, however, reduce to zero even in very large groups; instead it levels off at a positive value—namely, $(B - C)/B$—which depends on the benefit and cost of action to each individual.

We see how game-theoretic analysis sharpens the ideas from social psychology with which we started. The diffusion of responsibility theory takes us part of the way—namely, to the conclusion that any one person is less likely to act when he is part of a larger group. But the desired conclusion—that larger groups are less likely to provide help at all—needs further and more precise probability calculation based on the analysis of individual mixing and the resulting interactive (game) equilibrium.

And now we ask, did Kitty Genovese die in vain? Do the theories of pluralistic ignorance, diffusion of responsibility, and free-riding games still play out in the decreased likelihood of individual action within increasingly large cities? Perhaps not. John Tierney of the *New York Times* has publicly extolled the virtues of "urban cranks."[27] They are people who encourage the civility of

[25] Consider the case in which $B = 10$ and $C = 8$. Then $P$ equals 0.8 when $N = 2$, rises to 0.998 when $N = 100$, and approaches 1 as $N$ continues to rise. The probability of action by any one person is $1 - P$, which falls from 0.2 to 0 as $N$ rises from 2 toward infinity.

[26] With the same sample values for $B$ (10) and $C$ (8), this result implies that increasing $N$ from 2 to infinity increases the probability that not even one person helps from 0.64 to 0.8. And the probability that at least one person helps falls from 0.36 to 0.2.

[27] John Tierney, "The Boor War: Urban Cranks, Unite—Against All Uncivil Behavior. Eggs Are a Last Resort," *New York Times Magazine*, January 5, 1997.

the group through prompt punishment of those who exhibit unacceptable behavior—including litterers, noise polluters, and the generally obnoxious boors of society. Such "cranks" are essentially enforcers of a cooperative norm for society. And as Tierney surveys the actions of known "cranks," he reminds the rest of us that "[n]ew cranks must be mobilized! At this very instant, people are wasting time reading while norms are being flouted out on the street. . . . You don't live alone in this world! Have you enforced a norm today?" In other words, we need social norms and some people who have internalized the norm of enforcing norms.

## SUMMARY

Multiplayer games generally concern problems of *collective action.* The general structure of collective-action games may be manifested as a prisoners' dilemma, chicken, or an assurance game. The critical difficulty with such games in any form is that the Nash equilibrium arising from individually rational choices may not be the *socially optimal* outcome—the outcome that maximizes the sum of the payoffs of all the players.

In collective-action games, when a person's action has some effect on the payoffs of all the other players, we say that there are *spillovers,* or *externalities.* They can be positive or negative and lead to individually driven outcomes that are not socially optimal. When actions create negative spillovers, they are overused from the perspective of society; when actions create positive spillovers, they are underused. The additional possibility of *positive feedback* exists when there are positive spillovers; in such a case, the game may have multiple Nash equilibria.

Problems of collective action have been recognized for many centuries and discussed by scholars from diverse fields. Several early works professed no hope for the situation, but others offered up dramatic solutions. The most recent treatments of the subject acknowledge that collective-action problems arise in diverse areas and that there is no single optimal solution. Social scientific analysis suggests that social *custom,* or *convention,* can lead to cooperative behavior. Other possibilities for solutions come from the creation of *norms* of acceptable behavior. Some of these norms are *internalized* in individuals' payoffs; others must be *enforced* by the use of *sanctions* in response to the uncooperative behavior. Much of the literature agrees that small groups are more successful at solving collective-action problems than large ones.

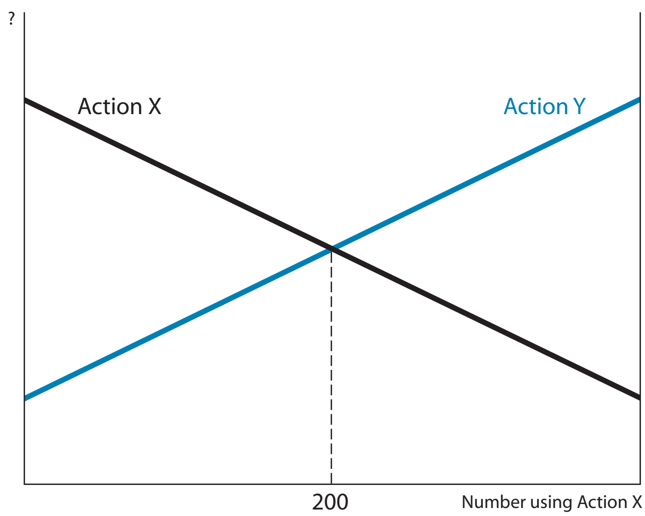
In large-group games, *diffusion of responsibility* can lead to behavior in which individual persons wait for others to take action and *free ride* off the benefits of that action. If help is needed, it is less likely to be given at all as the size of the group available to provide it grows.

## KEY TERMS

**coercion** (449)
**collective action problem** (417)
**convention** (445)
**custom** (445)
**diffusion of responsibility** (454)
**external effect** (433)
**externality** (433)
**free rider** (420)
**internalize the externality** (439)
**locked in** (441)
**marginal private gain** (433)
**marginal social gain** (433)
**nonexcludable benefits** (418)
**nonrival benefits** (418)
**norm** (445)
**oppression** (449)
**pluralistic ignorance** (454)
**positive feedback** (440)
**pure public good** (418)
**sanction** (445)
**social optimum** (420)
**spillover effect** (433)

## SOLVED EXERCISES

S1. Suppose that 400 people are choosing between Action X and Action Y. The relative payoffs of the two actions depend on how many of the 400 people choose Action X and how many choose Action Y. The payoffs are as shown in the following diagram, but the vertical axis is not labeled, so you do not know whether the lines show the benefits or the costs of the two actions.




(a) You are told that the outcome in which 200 people choose Action X is an *un*stable equilibrium. If 100 people are currently choosing

Action X, would you expect the number of people choosing X to increase or decrease over time? Why?

(b) For the graph to be consistent with the behavior that you described in part (a), should the lines be labeled as indicating the *costs* or *benefits* of Action X and Action Y? Explain your answer.

**S2.** A group has 100 members. Each person can choose to participate or not participate in a common project. If $n$ of them participate in the project, then each participant derives the benefit $p(n) = n$, and each of the $(100 - n)$ shirkers derives the benefit $s(n) = 4 + 3n$.

(a) Is this an example of a prisoners' dilemma, a game of chicken, or an assurance game?

(b) Write the expression for the total benefit of the group.

(c) Show, either graphically or mathematically, that the maximum total benefit for the group occurs when $n = 74$.

(d) What difficulties will arise in trying to get exactly 74 participants and allowing the remaining 26 to shirk?

(e) How might the group try to overcome these difficulties?

**S3.** Consider a small geographic region with a total population of 1 million people. There are two towns, Alphaville and Betaville, in which each person can choose to live. For each person, the benefit from living in a town increases for a while with the size of the town (because larger towns have more amenities and so on), but after a point it decreases (because of congestion and so on). If $x$ is the fraction of the population that lives in the same town as you do, your payoff is given by

$$\begin{array}{rl} x & \text{if } \; 0 \leq x \leq 0.4 \\ 0.6 - 0.5x & \text{if } \; 0.4 < x \leq 1. \end{array}$$

(a) Draw a graph like Figure 11.11, showing the benefits of living in the two towns, as the fraction living in one versus the other varies continuously from 0 to 1.

(b) Equilibrium is reached either when both towns are populated and their residents have equal payoffs or when one town—say Betaville—is totally depopulated, and the residents of the other town (Alphaville) get a higher payoff than would the very first person who seeks to populate Betaville. Use your graph to find all such equilibria.

(c) Now consider a dynamic process of adjustment whereby people gradually move toward the town whose residents currently enjoy a larger payoff than do the residents of the other town. Which of the equilibria identified in part (b) will be stable with these dynamics? Which ones will be unstable?

**S4.** Suppose an amusement park is being built in a city with a population of 100. Voluntary contributions are being solicited to cover the cost. Each citizen is being asked to give \$100. The more people contribute, the larger the park will be and the greater the benefit to each citizen. But it is not possible to keep out the noncontributors; they get their share of this benefit anyway. Suppose that when there are $n$ contributors in the population, where $n$ can be any whole number between 0 and 100, the benefit to each citizen in monetary unit equivalents is $n^2$ dollars.

(a) Suppose that initially no one is contributing. You are the mayor of the city. You would like everyone to contribute and can use persuasion on some people. What is the minimum number whom you need to persuade before everyone else will join in voluntarily?

(b) Find the Nash equilibria of the game where each citizen is deciding whether to contribute.

**S5.** Put the idea of Keynesian unemployment described at the end of Section 3.D into a properly specified game, and show the multiple equilibria in a diagram. Show the level of production (national product) on the vertical axis as a function of a measure of the level of demand (national income) on the horizontal axis. Equilibrium is reached when national product equals national income—that is, when the function relating the two cuts the 45° line. For what shapes of the function can there be multiple equilibria? Why might you expect such shapes in reality? Suppose that income increases when current production exceeds current income, and that income decreases when current production is less than current income. In this dynamic process, which equilibria are stable and which ones unstable?

**S6.** Write a brief description of a strategic game that you have witnessed or participated in that includes a large number of players and in which individual players' payoffs depend on the number of other players and their actions. Try to illustrate your game with a graph if possible. Discuss the outcome of the actual game in light of the fact that many such games have inefficient outcomes. Do you see evidence of such an outcome in your game?

## UNSOLVED EXERCISES

**U1.** Figure 11.5 illustrates the payoffs in a general, two-person, collective-action game. There we showed various inequalities on the algebraic payoffs [$p(1)$, etc.] that made the game a prisoners' dilemma. Now you are asked to find similar inequalities corresponding to other kinds of games:

(a) Under what condition(s) on the payoffs is the two-person game a chicken game? What further condition(s) make the game version I of chicken (as in Figure 11.3)?

(b) Under what condition(s) on the payoffs is the two-person game an assurance game?

**U2.** A class with 30 students enrolled is given a homework assignment with five questions. The first four are the usual kinds of problems, totaling to 90 points. But the fifth is an interactive game for the class. The question reads: "You can choose whether to answer this question. If you choose to do so, you merely write 'I hereby answer Question 5.' If you choose not to answer Question 5, your score for the assignment will be based on your performance on the first four problems. If you choose to answer Question 5, then your scoring will be as follows: If fewer than half of the students in the class answer Question 5, you get 10 points for Question 5; 10 points will be added to your score on the other four questions to get your total score for the assignment. If half or more than half of the students in the class answer Question 5, you get $-10$ points; that is, 10 points will be subtracted from your score on the other questions."

(a) Draw a diagram illustrating the payoffs from the two possible strategies, "Answer Question 5" and "Don't Answer Question 5," in relation to the number of other students who answer it. Find the Nash equilibrium of the game.

(b) What would you expect to see happen in this game if it were actually played in a college classroom? Why? Consider two cases: (i) the students make their choices individually with no communication; and (ii) the students make their choices individually but can discuss these choices ahead of time in a discussion forum available on the class Web site.

**U3.** There are two routes for driving from A to B. One is a freeway, and the other consists of local roads. The benefit of using the freeway is constant and equal to 1.8, irrespective of the number of people using it. Local roads get congested when too many people use this alternative, but if not enough people use it, the few isolated drivers run the risk of becoming victims of crimes. Suppose that when a fraction $x$ of the population is using the local roads, the benefit of this mode to each driver is given by

$$1 + 9x - 10x^2.$$

(a) Draw a graph showing the benefits of the two driving routes as functions of $x$, regarding $x$ as a continuous variable that can range from 0 to 1.

(b) Identify all possible equilibrium traffic patterns from your graph in part (a). Which equilibria are stable? Which ones are unstable? Why?

(c) What value of $x$ maximizes the total benefit to the whole population?

**U4.** Suppose a class of 100 students is comparing two careers—lawyer or engineer. An engineer gets take-home pay of \$100,000 per year, irrespective of the numbers who choose this career. Lawyers make work for each other, so as the total number of lawyers increases, the income of each lawyer increases—up to a point. Ultimately, the competition between them drives down the income of each. Specifically, if there are $N$ lawyers, each will get $100N - N^2$ thousand dollars a year. The annual cost of running a legal practice (office space, secretary, paralegals, access to online reference services, and so forth) is \$800,000. Therefore, each lawyer takes home $100N - N^2 - 800$ thousand dollars a year when there are $N$ of them.

(a) Draw a graph showing the take-home income of each lawyer on the vertical axis and the number of lawyers on the horizontal axis. (Plot a few points—say, for 0, 10, 20, . . . , 90, 100 lawyers. Fit a curve to the points, or use a computer graphics program if you have access to one.)

(b) When career choices are made in an uncoordinated way, what are the possible equilibrium outcomes?

(c) Now suppose the whole class decides how many should become lawyers, aiming to maximize the total take-home income of the whole class. What will be the number of lawyers? (If you can, use calculus, regarding $N$ as a continuous variable. Otherwise, you can use graphical methods or a spreadsheet.)

**U5.** A group of 12 countries is considering whether to form a monetary union. They differ in their assessments of the costs and benefits of this move, but each stands to gain more from joining, and lose more from staying out, when more of the other countries choose to join. The countries are ranked in order of their liking for joining, 1 having the highest preference for joining and 12 the least. Each country has two actions, IN and OUT. Let

$$B(i,n) = 2.2 + n - i$$

be the payoff to country with ranking $i$ when it chooses IN and $n$ others have chosen IN. Let

$$S(i,n) = i - n$$

be the payoff to country with ranking $i$ when it chooses OUT and $n$ others have chosen IN.

**(a)** Show that for country 1, IN is the dominant strategy.

**(b)** Having eliminated OUT for country 1, show that IN becomes the dominant strategy for country 2.

**(c)** Continuing in this way, show that all countries will choose IN.

**(d)** Contrast the payoffs in this outcome with those where all choose OUT. How many countries are made worse off by the formation of the union?