

# **Introduction to ‘The Possibility of Cooperation’**

**The Problem of Collective Action**

Michael Taylor

August 1987

# Contents

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| The Prisoners' Dilemma . . . . .                                                 | 12 |
| The problem of collective action . . . . .                                       | 15 |
| Time and the lone exploiter . . . . .                                            | 16 |
| Solutions? Community, states, entrepreneurs, property rights and norms . . . . . | 18 |
| Political entrepreneurs . . . . .                                                | 19 |
| Property rights . . . . .                                                        | 21 |
| Norms . . . . .                                                                  | 23 |
| Plan of the rest of the book . . . . .                                           | 24 |

The most persuasive justification of the state is founded on the argument that, without it, people would not successfully cooperate in realizing their common interests and in particular would not provide themselves with certain public goods : goods, that is to say, which any member of the public may benefit from, whether or not he or she contributes in any way to their provision. The most appealing version of this justification would confine the argument about voluntary cooperation to what are supposed to be the most fundamental public goods: goods (or services) which are thought to be preconditions of the pursuit and attainment of all other valued ends, including less basic public goods, and are therefore desired by *everyone* within the jurisdiction of the state in question.

*The Possibility of Cooperation* is a critique of this justification of the state, and the heart of the critique (chapters 2-4 below) is a detailed study of cooperation in the absence of the state and of other kinds of coercion. (The arguments about public goods provision and the theory of cooperation which make up these chapters can be read – and assessed – independently of the critique of the Hobbesian justification of the state.)

Hobbes's *Leviathan* was the first full expression of this way of justifying the state. The public goods with which he was principally concerned were social order – domestic peace and security – and defence against foreign aggression. Without these, very little else that was worth having could be had. Without internal and external security, there would be not only actual violence but such pervasive uncertainty as to undermine the incentive to invest resources in any projects with delayed returns. But although everyone would prefer the condition of peace and security that mutual restraint ensures to the 'war of all against all' that is the result of everyone pursuing his own interests without restraint, no individual has the incentive, in the absence of the state, to restrain himself. It is therefore rational, says Hobbes, for everyone to institute a government with sufficient power to ensure that everybody keeps the peace.<sup>1</sup>

Many writers who came after Hobbes, including some who professed no sympathy with what they took to be Hobbes's ideas, have taken over the core of his case for the state. Most economists who nowadays write about public goods believe that the failure of people to provide themselves voluntarily with these goods constitutes at least a *prima facie* case for state activity, and most of them presume that the state is the *only* means for remedying this failure.<sup>2</sup> (For nearly all the rest, the remedy is to establish or extend private property rights. I'll comment briefly on this view later in the chapter.)

Of course, many people believe that the state can be justified on further grounds and that it has functions other than that of providing public goods. Certainly, modern states do more than provide such goods. However, the justification I wish to criticize here is common to the arguments of nearly all those who believe that the state is necessary. Its persuasiveness lies in the fact that the state, on this view, exists to further *common* interests, to do what *everybody* wants done. Other arguments – for example, that income redistribution is desirable and can be

---

<sup>1</sup> This is a caricature of Hobbes's argument. In chapter 6, I give a more detailed account, making use of ideas developed in chapters 3–5.

<sup>2</sup> Both of these are true of William J. Baumol's *Welfare Economics and the Theory of the State*, second edition (London: G. Bell, 1965). Much of this book is devoted to the failure of individuals to provide themselves voluntarily with public goods, but I think it is fair to say that 'the Theory of the State' is missing. He is careful to say that, before it is concluded that state action to ensure the supply of public goods is justified, all the *costs* of state action must also be taken into account (p. 22 in the introduction added to the second edition); nevertheless there is a *presumption* that *only* the state could ensure this supply.

brought about only through the intervention of the state — do not appeal to common interests, not, at any rate, in an obvious or uncontroversial way.

In recent years, this argument about the necessity of the state has found new supporters amongst those concerned with the degradation of the environment, the depletion of non-renewable resources and rapid population growth. According to them, people will not *voluntarily* refrain from discharging untreated wastes into rivers and lakes, from hunting whales and other species threatened with extinction, from having ‘too many children’, and so on. Only powerful state action, they say, can solve or avert these problems, which are the consequence of failures to provide public goods (and more generally ‘non-excludable’ goods about which more later). For many environmentalists, some of these public goods are at least as fundamental as peace and security were for Hobbes. Continued failure to provide them will eventually result in ecological catastrophe. Without them, the life of man will not just be ‘solitary, poore, nasty, brutish, and short’; it will be impossible.<sup>3</sup>

There have of course been other responses to the environmental crisis. In particular, some writers, who probably did not think of themselves as anarchists, have come to embrace essentially communitarian anarchist ideas. However, of those who desire on ecological grounds the goal of a social organization along communitarian anarchist lines, there are very few who believe that a transition to such a society can be made without extensive state activity.<sup>4</sup>

As for the members of governments themselves, and indeed of most political parties, especially in industrialized countries, they generally do not recognize that there is or will be an environmental crisis and they believe, not unnaturally, that pollution and resource depletion are problems which can be adequately dealt with by minor modifications within the present institutional framework of whatever country they happen to live in.<sup>5</sup> Generally speaking, the proposed modifications involve an extension of state activity, in the form of state-enforced pollution standards and resource depletion quotas, taxes on industrial pollution, government subsidies and tax credits for the development of pollution control technology, and so on. Most economists who have written on problems of pollution and resource depletion have also confined their discussions to ‘solutions’ of this sort, or otherwise have recommended the extension of private property rights.

Much of what I shall have to say in this book will in fact apply, not just to the voluntary provision of public goods but to ‘collective action problems’, a much larger category. The defining characteristic of a collective action problem, as I shall use this expression, is very roughly that rational egoists are unlikely to succeed in cooperating to promote their common interests. (I will clarify this in a later section.) On this account, as we shall see, the category of collective action problems includes many but not all public goods problems. There is, in particular, a very important class of collective action problems which arise in connection with the use of resources

---

<sup>3</sup> Two examples are William Ophuls, ‘Leviathan or oblivion?’, in Herman E. Daly (ed.), *Toward a Steady-State Economy* (San Francisco: W. H. Freeman, 1973), and Robert L. Heilbroner, ‘The human prospect’, *The New York Review of Books*, 24 January 1974.

<sup>4</sup> An approximate example is *A Blueprint for Survival*, by the editors of *The Ecologist* (Harmondsworth, Middlesex: Penguin Books, 1972; originally published as Vol. 2, No. 1 of *The Ecologist*, 1972). Their goal is not wholly anarchist, but it does include ‘decentralisation of polity and economy at all levels, and the formation of communities small enough to be reasonably self-supporting and self-regulating’. For an anarchist’s account of the necessity of anarchist society on ecological grounds, see Murray Bookchin, ‘Ecology and Revolutionary Thought’, in *Post-Scarcity Anarchism* (Berkeley, California: The Ramparts Press, 1971).

<sup>5</sup> See, for example, Anthony Crosland, *A Social Democratic Britain* (Fabian Tract no. 404, London, 1971), and Jeremy Bray, *The Politics of the Environment* (Fabian Tract no. 412, London, 1972).

to which there is open access — resources, that is, which nobody is prevented from using. These resources need not be public goods, as I will define them shortly. Garrett Hardin's well-known 'tragedy of the commons' concerns resources of this kind.<sup>6</sup>

Hardin asks us to imagine a common, a pasture open to all. The village herdsman keep animals on the common. Each herdsman is assumed to seek to maximize his own gain. As long as the total number of animals is below the carrying capacity of the common, a herdsman can add an animal to his herd without affecting the amount of grazing of any of the animals, including his own. But beyond this point, the 'tragedy of the commons' is set in motion. Asking himself now whether he should add another animal to his herd, he sees that this entails for him a gain and a loss: on the one hand, he obtains the benefit from this animal's yield (milk, meat or whatever); on the other hand, the yield of each of his animals is reduced because there is now overgrazing. The benefit obtained from the additional animal accrues entirely to the herdsman. The effect of overgrazing, on the other hand, is shared by all the herdsman; every one of them suffers a slight loss. Thus, says Hardin, the benefit to the herdsman who adds the animal is greater than *his* loss. He therefore adds an animal to the common. For the same reason, he finds that it pays him to add a second animal, and a third, a fourth and so on. The same is true for each of the other herdsman. The result is that the herdsman collectively bring about a situation in which each of them derives less benefit from his herd than he did before the carrying capacity of the common was exceeded. The process of adding animals may indeed continue until the ability of the common to support livestock collapses entirely.

For similar reasons, many species of fish and whales are hunted without limit and in some cases brought close to extinction : the oceans are like a great common. For similar reasons, too, lakes and rivers are polluted, since each polluter finds that the costs of treating his wastes before discharging them or of modifying his product are too great in comparison with what he suffers from the decline in the quality of the air or water caused by his effluent.

In all these situations, we can say that it is in every individual's interest not to restrain himself (from adding animals to the common, polluting the lake, etc.) but the result of everyone acting without restraint is a state of affairs in which every individual is less well off than he would be if everybody restrained themselves.

In such situations, we might expect people to make an agreement in which they all promised to restrain themselves. However, in the absence of the state (or some other form of coercion), no individual has any greater incentive to abide by the agreement than he had to restrain himself before the agreement was made.

I shall later question whether grazing commons such as those which were part of the European open-field systems or those which were once widespread in pastoral economies — do in fact typically have open access. But certainly there are many such resources to be found in other contexts.

The recent history of the whale 'fisheries' provides a sad example. During the 1950s and 1960s, unlimited killing of blue and fin-back whales, which are the biggest, brought these two species close to extinction. When stocks of blues and fin-backs became very low, the other large species were hunted without limit. In each case the annual harvest far exceeded the maximum sustainable yield, that is, the maximum number which can be replaced each year through reproduction (and the whale hunters knew this). The profitability of whaling declined, and most of the former

---

<sup>6</sup> Garrett Hardin, 'The tragedy of the commons', *Science*, 162 (13 December 1968), 1243–8.

whaling countries were obliged one by one to leave the industry (so that, by 1968, there were only two countries, Japan and the USSR, left in the field). It seems fairly certain that if it were not for the diminished profits from hunting a sparse population, the blue whale and other species would in fact have been hunted to extinction. After the Second World War, the International Whaling Commission was set up by the seventeen countries who were then interested in whaling and was charged with regulating harvests and ensuring the survival of threatened species. Until very recently, this Commission, which has no powers of enforcement, has not been very successful. Its members were often unable to agree to impose the quotas recommended by biologists, or else they could agree only to limits in excess of these recommendations; and when the Commission did decide either to limit harvests or to protect a species completely, the agreement was not always observed by every country.<sup>7</sup>

The provision of public and other non-excludable goods Before embarking on a detailed analysis, we need to define a little more carefully some of the terms that have already been used, and in particular the notions of a 'public good' and a 'collective action problem'.

I shall say that a good or service is a *public good* (or collective good) if it is in some degree indivisible and non-excludable. A good is said to exhibit perfect *indivisibility* or *jointness of supply* (with respect to a given set of individuals, or public) if, once produced, any given unit can be made available to every member of the public, or equivalently if any individual's consumption or use of the good does not reduce the amount available to others.<sup>8</sup> A good is said to exhibit *non-excludability* (with respect to some group) if it is impossible to prevent individual members of the group from consuming it or if such exclusion is 'prohibitively costly' (a notion whose precise definition matters for some purposes but not for mine here).

A perfectly *divisible* good is one that can be divided between individuals. Once any part of it is appropriated by any individual, the same part cannot be made available to others; and once any unit of it is consumed by any individual, the amount available for consumption by others is reduced by the whole of that unit. A loaf of bread and a pot of honey are examples of perfectly divisible goods. A good which is perfectly divisible is called a *private good*. Thus, in order to be public, a good must exhibit some degree of indivisibility or jointness.<sup>9</sup>

---

<sup>7</sup> For a brief account of the overexploitation of whales and various species of fish, see Paul R. Ehrlich and Anne H. Ehrlich, *Population, Resources, Environment*, second edition (San Francisco: W. H. Freeman, 1972), pp. 125–34. See also Frances T. Christy and Anthony Scott, *The Common Wealth in Ocean Fisheries* (Baltimore: Johns Hopkins Press, 1965).

<sup>8</sup> The word 'consumption' should perhaps be used only in connection with private goods, where it has a clear meaning. To speak of 'consuming' national defence, wilderness and radio broadcasts is somewhat strained, but for want of a suitable word to cover a variety of applications, I follow the custom of the economists and retain the word. In many cases, 'consume' means 'use'. Cf. Jean-Claude Milleron, 'Theory of value with public goods: a survey article', *Journal of Economic Theory*, 5 (1972), 419–77, at pp. 422–3.

<sup>9</sup> This follows Samuelson's most recent usage (though I have added the requirement that a public good be also non-excludable). Samuelson had defined a *public good* as one which was consumed *equally* by every individual, so that  $X^1 = X^2 = \dots = x$ , where  $x^i$  is the  $i^{\text{th}}$  individual's consumption of the good and  $x$  is the total amount available; and he defined a *private good* as one which could be divided amongst individuals so that  $x^1 + x^2 + \dots = x$ . See Paul A. Samuelson, 'The pure theory of public expenditure', *Review of Economics and Statistics*, 36 (1954), 387–9. In his 1955 paper, he admitted that these were two pure, polar cases; and most recently he has abandoned these two poles in favour of a 'knife-edge pole' of the pure private good and 'all the rest of the world in the public good domain'. Samuelson, 'Diagrammatic exposition of a theory of public expenditure', *Review of Economics and Statistics*, 37 (1955), 350–6; and 'Pure theory of public expenditure and taxation', in J. Margolis and H. Guitton (eds), *Public Economics* (London: Macmillan, 1969).

A good may be indivisible yet excludable. A road or bridge or park can be provided in this form. Once supplied to one individual, it *can* be made available to others, but it need not be, for it is possible and may not be prohibitively costly to exclude particular individuals. Hence tolls and admission charges can be imposed. Goods like these can be provided in an excludable or non-excludable mode. Indivisibility, then, does not imply non-excludability. Furthermore, divisibility does not entail excludability, although important examples of non-excludable, divisible goods are not easy to come by: economists have suggested such examples as a garden of flowers, whose nectar can be appropriated by individual bees but particular bees cannot be excluded from consumption.

If an individual is not excluded from consumption or use of a public good, it is possible for him to be a *free rider* on the efforts of others, that is, he can consume or use the public good that is provided by others (unless of course everyone else tries to free-ride as well!). Whether or not he will in fact be a free rider is something we have to examine.

Free rider problems (and hence, as we shall see, collective action problems) can arise where there is non-excludability but not indivisibility. In fact non-excludability (or *de facto* non-exclusion) and divisibility (at least in principle) characterize Garrett Hardin's 'commons', or any resource to which there is open access, such as a common fishing ground, a common underground reservoir of oil or water, or the open range on the Great Plains before any property rights were established (including the 'common property' rights that the cattlemen tried to maintain when they formed associations to regulate access and use). I will have a little more to say about such resources in a later section. With some of them, exclusion is possible and economically feasible, but whether or not there is in fact exclusion, consumption or use by one individual reduces the amount available to others and any cutting back on consumption by one individual allows others to consume more.

Most, if not all, public goods interactions are characterized by a certain degree of *rivalness*. It is normally said that a good is *rival* to the extent that the consumption of a unit of the good by one individual decreases the *benefits* to others who consume that same unit. Obviously, in the case of a perfectly divisible good the consumption of a particular unit prevents any other individual from consuming it at all, so that there can be no question of his benefiting from consumption. In this case we might say that the good is perfectly rival. But non-rivalness is not the same thing as indivisibility, as some writers like to say, even though they are usually closely associated. Where there is some degree of divisibility, consumption reduces the *amount* available to others; but where there is some degree of rivalness, consumption reduces the *benefits* to other consumers. An individual's benefit from consumption may not change at all as the amount available for consumption declines, until some threshold of 'crowding' is reached. In fact, although others' consumption usually *lowers* an individual's utility — as is normally the case with congested parks, beaches and roads and with various forms of pollution — some individuals' utilities may *rise* as the number of other consumers increases, at least up to a point; they may, for example, prefer a semicrowded beach or park to an empty one. This brings out the point that rivalness, unlike indivisibility, is strictly speaking a property of individuals (or of their utility functions), not of the goods themselves.

Rivalness is clearly important in the analysis of collective action problems. As we would expect and as we shall see in the next chapter — it plays a crucial role in the analysis of 'size' effects. For just as the alternatives actually available to an individual must change as the number

of people in the group increases if there is some degree of divisibility, so the *utilities* of these alternatives must change with group size if there is rivalness.

I said in an earlier section that social order and national defence are public goods. This needs some qualifying and we are now in a position to do so. National defence can in fact be decomposed (on a first rough cut) into deterrence, which is a pure public good because it is both perfectly indivisible and completely non-excludable, and protection from attack, which is imperfectly indivisible and more or less excludable depending on the form it takes. (And of course the production of the means to these ends produces incidental private goods, including income for shareholders and employees of business firms.)<sup>10</sup> The security of persons and their property which I have taken to be constitutive of social order is similarly the product of a variety of goods and services, which range (in a modern society) from purely private goods like locks and private bodyguards through such services as police forces and law courts to deterrence, which, again, can be purely indivisible — as it would be if the fact that attack was deterred on one individual did not diminish the deterrent effect with respect to other individuals.<sup>11</sup>

Under what circumstances, then, will people cooperate to provide a public good or any non-excludable good which the members of a group have a common interest in providing? The now standard answer to this question (which, however, needs to be qualified, as we shall see) is the one provided by Mancur Olson in his well-known study, *The Logic of Collective Action*. Olson's main contention is that 'the larger a group is, the farther it will fall short of providing an optimal supply of any collective good, and the less likely that it will act to obtain even a minimal amount of such a good. In short, the larger the group, the less likely it will further its common interests.'<sup>12</sup>

There are three arguments in support of this conclusion to be found in Olson's book. Before setting them out, we need some definitions. A group is *privileged* if it pays at least one of its members to provide some amount of the public good unilaterally, that is, to bear the full cost of providing it alone. Any group which is not privileged is said to be *latent*. Where the group is privileged, there is, in Olson's view, a 'presumption' that the public good will be provided; but there should be no such presumption in the case of a latent group. Nevertheless, some latent groups (in Olson's account, which is a little muddy at this point) are sufficiently small that through some sort of strategic interaction amongst their members they may succeed in providing some amount of the public good. (They do not have so many members, says Olson, that an individual contribution to the provision of the public good will go unnoticed by other members.) Such groups are called *intermediate*. The remaining latent groups are so large that this sort of strategic interaction, depending as it does on individual contributions being 'noticeable', is impossible and an individual will contribute only if there is a *selective incentive* to do so, that is, the individual receives a (private) benefit if and only if he contributes and/or incurs a (private) cost if and only if he fails to contribute. Thus, for example, trade unions, which are founded primarily to provide for their members certain public goods such as higher wages and better working conditions, have also had to offer prospective members sickness, unemployment and dispute benefits and other positive selective incentives, and to operate a 'closed shop' which bars non-members from employment.

---

<sup>10</sup> Cf. William Loehr and Todd Sandler (eds), *Public Goods and Public Policy* (Beverly Hills: Sage, 1978), p. 2 and ch. 6.

<sup>11</sup> For a fuller discussion of social order as a public good, see my *Community, Anarchy and Liberty* (Cambridge: Cambridge University Press, 1982), sections 2.1 and 2.3.

<sup>12</sup> Mancur Olson, *The Logic of Collective Action* (Cambridge, Mass.: Harvard University Press, 1965), p. 36.

In deciding whether or not to contribute or participate, the individual compares the cost to him of making his contribution and the benefit *to him* of the additional amount of the public good *provided as a result of his contribution*. The final italicized phrase encompasses both the public good which he himself directly produces or which is funded by his contribution *and* whatever additional public good is provided by the contributions that *others* may make as a result of his contribution (because their contributions are in some way contingent on his). This second component of his benefit may not be forthcoming, because the required interdependence is absent. It is this interdependence which for Olson characterizes 'intermediate' groups.

Now we can state the three arguments which Olson offers in support of his argument that larger groups are less likely than smaller groups to provide any (or an optimal) amount of the public good.<sup>13</sup>

(i) The larger the group, the smaller is each individual's net benefit from the public good.

(ii) The larger the group, the less the likelihood that it will be privileged or intermediate.

(iii) The larger the group, the greater the 'organization costs' of providing the public good (including the costs of communication and bargaining amongst group members and perhaps the costs of creating and maintaining a formal organization).

The last of these claims is the most straightforward. It is also no doubt empirically true, for very many cases.

The second claim is a little less straightforward. How much support it gives to the argument that a public good is more likely to be provided in smaller groups depends on the reliability of Olson's 'presumption' that the public good will be provided in privileged groups and on how likely it is that collective action will be successful in intermediate groups. Olson says that the outcome of interaction in intermediate groups is 'indeterminate'. As for the 'presumption', it is perhaps appropriate only where there is just one individual who is willing to provide the public good unilaterally (and even then there should be no presumption that an optimal amount of it will be provided). But if two or more individuals are so willing, then there could be strategic interaction amongst them ... and the outcome of such interaction is indeterminate.<sup>14</sup> (The game amongst these players may be what is known as a Chicken game, which will be discussed in some detail in the next chapter.) The privileged group is therefore in-effect a group within which there is an intermediate group, that is, a group with a subgroup whose members interact strategically.

The privileged group, it seems to me, is a special case of a group *with at least one subgroup whose members collectively find it worthwhile to provide some amount of the public good by themselves*, that is, a subgroup such that, if all its members cooperated to provide the public good, each of them would be better off than they would be if none of the public good was provided. Again, this does not guarantee that any of the public good will be provided, since normally there will be strategic interaction amongst the members of the subgroup, and there will be strategic behaviour of a different kind — which may also obstruct provision of the public good resulting from the coexistence of *several* such subgroups. A group which is 'privileged' in this generalized sense is also, then, a group within which there is at least one 'intermediate' subgroup.

A final point about claim (ii) is that, as Russell Hardin has observed, there is no necessary connection, and probably a very weak correlation, between the size of a group and whether it is

---

<sup>13</sup> Olson, *The Logic*, p. 48.

<sup>14</sup> Cf. *The Logic*, p. 50, note 70.

privileged (in Olson's or my generalized sense) or intermediate. Privileged groups can be large; groups as small as two can be intermediate or latent.<sup>15</sup>

It is worth emphasizing here parenthetically that it is dangerous to distinguish intermediate and latent groups, as Olson sometimes did and as so many later authors have done, by reference to whether an individual contribution is 'noticeable' or 'perceptible'. Such talk has led a number of writers astray.<sup>16</sup> Individual contributions can be perfectly 'noticeable' in a group which is not privileged and in which there is no strategic interaction, and which therefore fails (in the absence of selective incentives) to provide any of the public good; the failure arises because each individual's contribution, though 'noticeable', brings too little of the public good to be worth the cost of the contribution.

This leaves the first of Olson's three arguments about the effect of increasing group size. As it stands this claim is undecidable. Before we can assess it, we must know what kind of public good is involved and what is held constant as size varies, for, as Hardin says, it is not possible to increase size while holding *everything* else constant.<sup>17</sup> We should, however, hold as many things as possible constant if we are to isolate a pure *size* effect. Now the individual's net benefit can decrease as group size increases because the costs of providing the public good (excluding the organizational costs) increase or the individual's benefits decrease or both. If it is a pure size effect we are looking for, we should count a cost increase as support for Olson's claim only where such an increase is unavoidable. The individual's benefit, on the other hand, decreases with group size only if there is *imperfect jointness* or some degree of *rivalness* or both. If jointness is less than perfect, that is, there is some 'crowding', then the amount available to an individual decreases as the number of consumers increases. If there is rivalness, then, as size increases, the individual's benefits decrease, whether or not the amount actually available to him decreases. (Normally, rivalness is an effect of imperfect jointness. But the two are analytically distinct, and in practice the effects of rivalness can set in, as group size increases, before the effects of imperfect jointness or literal 'crowding' do.)<sup>18</sup>

Olson's first claim in support of the 'size' effect, then, is not necessarily true. It holds only where costs unavoidably increase with size or where there is imperfect jointness or rivalness or both. Most goods, however, exhibit *some* divisibility, and most public goods interactions exhibit *some* rivalness (which is, recall, a property of individual utility functions rather than directly of the good). That is the theoretical position; in practice, we often want to compare groups which differ not only in size but in so many other particulars that this claim is undecidable because isolating a pure size effect is impossible.

There is a more important reason for not pursuing the issue here. The argument here (following Olson and Hardin) assumes that we can simply subtract costs from benefits. This is generally unrealistic (as Olson himself admits<sup>19</sup>). Preferences should instead be represented by indifference maps. This will be done in chapter 2. Further, Olson's whole analysis is entirely static: the individual is supposed in effect to make just one choice, once and for all, of how much to contribute to the public good. But in the real world, most public goods interactions are dynamical. The choice

---

<sup>15</sup> Russell Hardin. *Collective Action* (Baltimore: The Johns Hopkins Press for Resources for the Future, 1982), pp. 41–2.

<sup>16</sup> Including Olson himself, as we shall see when we come to discuss altruism in chapter 5.

<sup>17</sup> Hardin, *Collective Action*, p. 44.

<sup>18</sup> This qualifies the very useful treatment of this issue in *Collective Action*, ch. 3.

<sup>19</sup> Olson, *The Logic*, p. 29, note 46.

of whether to contribute and how much to contribute is a recurring one. There is interaction *over time* between different individuals' choices. And the individual's intertemporal preferences (how much he discounts future relative to present benefits) matter. A dynamical analysis is the subject of chapters 3 and 4.

Olson's *model*, then, is rather unrealistic. Accordingly, not too much weight should be attached to conclusions derived from it, including conclusions about the effects of increases in group size. The size effect which I think should be taken most seriously is *the increased difficulty of conditional cooperation* in larger groups. For, as we shall see, in a dynamical analysis the provision of a public good, or collective action more generally, requires that amongst at least some members of the group there is conditional cooperation. Olson is of little help here, since he does not provide (indeed cannot provide, within his static model) an analysis of conditional cooperation or of any other sort of strategic interaction over time.

To round out this brief discussion of Olson's treatment of the problem of collective action, a comment is in order on his assumptions about incentives and individual motivation. Recall that, according to Olson, only a selective incentive will motivate the member of a large latent group (one that is too large to be intermediate) to contribute to the provision of the public good. (If Olson did not think of individual contributions in such groups as being 'imperceptible' or 'infinitesimal', he would perhaps have said : only the *addition* of a selective incentive will make the difference between contribution and non-contribution. For there is no reason why, upon the introduction of selective incentives, the public good benefit to the individual should drop out of his calculation even though it is very small : it is never so small as to be 'infinitesimal'. ) In fact, says Olson, the public good lobbying efforts of large groups are *byproducts* of organizations which obtain their support by offering selective incentives.<sup>20</sup> But this argument, as several writers have pointed out, though it helps to explain the maintenance of the organization, does not explain its origin.

Now selective incentives can be either positive or negative – providing a benefit to a contributor or imposing a cost on a non-contributor – and they are limited, in Olson's account, to either 'monetary' or 'economic' incentives and 'social' incentives. The *social* incentives essentially derive from the desire for approbation and the dislike of disapprobation, and work through mechanisms like criticism and shaming by friends and associates. Such incentives are effective only in relatively small groups. Hence, a very large group might yet succeed in providing a public good if it has a *federal* structure, for within the local branches or subgroups social incentives can operate to maintain support. (And I would add: if the local branches are small enough for social incentives to be effective, they are probably small enough for conditional cooperation to be sustained, perhaps with the help of the social incentives. More on this in later chapters.)

There are therefore at most four components in the individual's benefit-cost calculations: (i) the benefit to the individual from the increased amount of the public good provided as a result of his contribution; (ii) the cost of his contribution; (iii) the individual's portion of the costs of organization; and (iv) the 'economic' and 'social' benefits and/or costs which operate as selective incentives.

Olson explicitly excludes other types of incentives, including 'psychological' ones, such as 'the sense of guilt, or the destruction of self-esteem, that occurs when a person feels he has forsaken his moral code'.<sup>21</sup> The important reason why (in any explanatory theory) the range of incentives

---

<sup>20</sup> *The Logic*, p. 132.

<sup>21</sup> *The Logic*, p. 61, note 17. But see also p. 160, note 91.

which are assumed to motivate individuals must be limited though this is not among the reasons Olson gives for *his* restriction — is that without such a limitation a rational choice theory such as Olson's is liable to become tautologous. Three important kinds of motivation which Olson in common with nearly all other rational choice theorists — excludes are altruistic motivations (to be discussed in chapter 5), expressive motivations and 'intrinsic' motivation by benefits got in the very act of participating in the provision of the public good as opposed to the benefits which successful provision would bring. The last two give rise to non-instrumental action.<sup>22</sup>

## The Prisoners' Dilemma

It has been widely asserted that individual preferences in public goods interactions and in collective action problems generally are (or usually are) those of a Prisoners' Dilemma game.<sup>23</sup> This game is defined as follows.

Suppose that there are just two individuals (or *players*) and that each of them may choose between two courses of action (or *strategies*). The players are labelled 1 and 2 and the strategies C and D. The two players must choose strategies simultaneously, or, equivalently, each player must choose a strategy in ignorance of the other player's choice. A pair of strategies, one for each player, is called a *strategy vector*. Associated with each strategy vector is a *payoff* for each player. The payoffs can be arranged in the form of a *payoff matrix*. The payoff matrix for the two-person Prisoners' Dilemma which will be studied in this book is:

|          |   | player 2 |      |
|----------|---|----------|------|
|          |   | C        | D    |
| player 1 | C | x, x     | z, y |
|          | D | y, z     | w, w |

where  $y > x > w > z$ . Throughout the book, the usual convention is adopted that rows are chosen by player 1, columns by player 2, and that the first entry in each cell of the matrix is the payoff to player 1 and the second entry is 2's payoff.

<sup>22</sup> On these non-instrumental motivations, see my 'Rationality and revolutionary collective action', in Michael Taylor (ed.), *Rationality and Revolution* (Cambridge: Cambridge University Press, 1987).

<sup>23</sup> The story about two prisoners, which gave the game its name, can be found in R. Duncan Luce and Howard Raiffa. *Games and Decisions* (New York: John Wiley, 1957), p. 95.

Notice first that, since we have assumed  $y > x$  and  $w > z$ , each player obtains a higher payoff if he chooses  $D$  than if he chooses  $C$ , *no matter what strategy the other player chooses*. Thus, it is in each player's interest to choose  $D$ , no matter what he expects the other player to do.  $D$  is said to *dominate*  $C$  for each player.

However, notice now that, if each player chooses his dominant strategy, the outcome of the game is that each player obtains a payoff  $w$ , whereas there is another outcome  $(C, C)$ , which yields a higher payoff to both players, since we have assumed  $x > w$ .

Let us say that an outcome ( $Q$ ) is Pareto-optimal if there is no other outcome which is not less preferred than  $Q$  by any player and is strictly preferred to  $Q$  by at least one player. An outcome which is not Pareto-optimal is said to be Pareto-inferior. Thus, in the two-person Prisoners' Dilemma, the outcome  $(D, D)$  is Pareto-inferior.

If the players could communicate and make agreements, they would presumably both agree to choose strategy  $C$ . But this would not resolve the dilemma, since neither has an incentive to keep the agreement: whether or not he thinks the other player will keep his part of the agreement, it pays him to defect from the agreement and choose  $D$ .

$C$  and  $D$  are the conventional labels for the two strategies in the Prisoners' Dilemma. They stand for Cooperate and Defect. I use them throughout this book, though they are not entirely appropriate: one player may 'Cooperate' (choose  $C$ ) by himself, and he may 'Defect' (choose  $D$ ) even though no agreement has been made from which to defect. In this book, Cooperation and Defection (with capital initials) will always refer to strategies in a Prisoners' Dilemma (or, in chapter 2, in some other game).

If communication between the players is impossible or prohibited, or if communication may take place but agreements are not binding on the players, then the game is said to be *non-cooperative*. The Prisoners' Dilemma is defined to be a non-cooperative game. If it were not, there would be no 'dilemma': the players would obtain  $(C, C)$  as the outcome, rather than the Pareto-inferior outcome  $(D, D)$ . In the situations of interest in this book, communication is generally possible but the players are not constrained to keep any agreements that may be made. It is the possibility of Cooperation (to achieve the outcome  $(C, C)$ ) in the *absence* of such constraint that will be of interest.

As a generalization of this two-person game, an  $N$ -person Prisoners' Dilemma can be defined as follows. Each of the  $N$  players has two strategies,  $C$  and  $D$ , available to him. For each player,  $D$  dominates  $C$ , that is, each player obtains a higher payoff if he chooses  $D$  than if he chooses  $C$ , no matter what strategies the other players choose. However, every player prefers the outcome  $(C, C, \dots, C)$  at which everybody Cooperates to the outcome  $(D, D, \dots, D)$  at which everybody Defects. Thus, as in the two-person game, every player has a dominant strategy but if every player uses his dominant strategy the outcome is Pareto-inferior.

Two-person and  $N$ -person Prisoners' Dilemmas can both be defined in the more general case when any finite number of strategies is available to each player. The generalization, which could be made in several ways, must at least have the characteristic that the predicted outcome is Pareto-inferior. In particular, it could again be stipulated that every player has a strategy which dominates each of the others, and if every player uses his dominant strategy the outcome is Pareto-inferior. I shall not elaborate on this here, as my discussion in this book will mainly be confined to the two-strategy games, though in chapter 2 I shall also consider games in which each individual can choose to contribute a continuously variable amount within some range.

Let us go back now to the ‘tragedy of the commons’. In Garrett Hardin’s account, each individual has in effect a dominant strategy: to add an animal to his herd on the common, to discharge his sewage untreated, to kill as many whales as possible and so on. Each of these corresponds to strategy *D*. The alternative, strategy *C*, is to refrain from doing these things. Hardin assumes, in effect, that *D* yields the highest payoff to each individual, no matter what strategies the other individuals choose (that is, no matter how many of them Cooperate); and he assumes that every individual prefers the mutual Cooperation outcome (*C*, *C*, ..., *C*) to the mutual Defection outcome (*D*, *D*, ..., *D*). In other words, individual preferences are assumed to be those of an N-person Prisoners’ Dilemma.

Russell Hardin has argued explicitly that public goods interaction in sufficiently large groups – in fact ‘the collective action problem’ more generally – can be represented by the N-person Prisoners’ Dilemma.<sup>24</sup> His analysis is as follows. Suppose that each of *N* individuals has the choice (and only the choice) between contributing and not contributing one unit of the cost of producing a non-excludable good (one unit of a *numeraire* private good) and that every unit contributed produces an amount of the public good with benefit *r*. Suppose that each individual’s utility is  $nr/N$  if he does not contribute and  $nr/N - 1$  if he does, where *n* is the total number of units contributed. (Notice that this means that the public good exhibits some rivalness: each individual’s utility declines with increasing *N*, which is the number of individuals who actually consume the good, since nobody is excluded.) Then, if *m* other individuals contribute, an individual’s utility is  $mr/N$  if he does not contribute and  $(m + 1)r/N - 1$  if he does. Thus, the first of these utilities exceeds the second if and only if  $N > r$ , which is independent of *m*. In other words, no matter how many other individuals contribute, it does not pay anyone to contribute as long as the size of the public (*N*) exceeds the ratio of benefits to costs (*r*). When  $N > r$ , the game is an N-person Prisoners’ Dilemma (as defined above): each individual has a dominant strategy, and the outcome which results when everyone chooses his dominant strategy is for everyone less preferable than another outcome. But when  $N < r$ , the dominant strategy for every individual is to contribute and the resulting outcome is the only Pareto-optimal position.

This argument, if correct, would apply also to the public goods with which Hobbes was chiefly concerned, namely domestic peace and security and national defence. I shall indeed show (in chapter 6) that Hobbes assumed men’s preferences in the absence of the state to be those of a Prisoners’ Dilemma game. The remainder of Hobbes’s theory can then be summarized, somewhat crudely, as follows : (a) in the absence of any coercion, it is in each individual’s interest to choose strategy *D*; the outcome of the game is therefore mutual Defection; but every individual prefers the mutual Cooperation outcome ; (b) the only way to ensure that the preferred outcome is obtained is to establish a government with sufficient power to ensure that it is in every man’s interest to choose *C*.

This is the argument which I wish to criticize in this book. But there is one element of the argument which I shall not quarrel with, namely, the analysis of the Prisoners’ Dilemma given above. If individual preferences in the provision of a public good are in fact those of a Prisoners’ Dilemma, then it is quite correct to conclude that the players will not voluntarily Cooperate. To avoid any misunderstanding, I emphasize that the conclusion is correct no matter what the entries in the payoff matrix (which is assumed to be a Prisoners’ Dilemma) actually represent, just as long as it is assumed that each player is concerned only to maximize his own payoff. Of

<sup>24</sup> Russell Hardin, ‘Collective action as an agreeable n-Prisoners’ Dilemma’, *Behavioral Science*, 16 (1971), 472–81.

course, the payoffs may not reflect *all* the incentives affecting the individuals in the situation in question. The conclusion still follows logically; but it is possible to argue that the payoff matrix is a poor description of the relevant real world situation and that in reality the players *do* Cooperate, because the omitted incentives are more important than those reflected in the payoff matrix.

In the next three chapters, the payoffs are assumed *not* to reflect, *inter alia*, (i) incentives due to external coercion, including that applied or threatened by the state or any other external agency or by other members of the group (apart from the tacit threats and offers which may be thought to be embedded in conditional Cooperation — about which more later); (ii) altruistic motivation; and (iii) any ‘internal sanctions’ like guilt, loss of self-respect and so on, which may result from failure to conform to a norm, live up to one’s own ideals, perform one’s duties, or whatever. In chapter 5 I shall begin with a matrix of payoffs which again do not reflect these three classes of incentives, but then I shall consider the effects of assuming that individuals are altruistic (that is, they take account of other players’ payoffs as well as their own in choosing strategies).

The expression ‘voluntary Cooperation’, used occasionally throughout the book, refers to Cooperation chosen only on the basis of the matrix of payoffs (or utilities, where the individual is in some way altruistic); thus, voluntary Cooperation is Cooperation which, amongst other things, is *not* the result of external coercion, including that applied or threatened by the state.

## The problem of collective action

I shall argue in the next chapter that, in many interesting problems of public goods provision, individual preferences at any point in time are *not* those of a Prisoners’ Dilemma. Many other preference structures can arise. These include Chicken and Assurance games, whose two-person payoff matrices are shown below.

|                |          |          |                  |          |          |          |
|----------------|----------|----------|------------------|----------|----------|----------|
|                | <b>C</b> | <b>D</b> |                  |          | <b>C</b> | <b>D</b> |
| <b>C</b>       | 3, 3     | 2, 4     |                  | <b>C</b> | 4, 4     | 1, 2     |
| <b>D</b>       | 4, 2     | 1, 1     |                  | <b>D</b> | 2, 1     | 3, 3     |
| <i>Chicken</i> |          |          | <i>Assurance</i> |          |          |          |

Surely, then, we should not equate ‘the problem of collective action’ with the Prisoners’ Dilemma, as many writers have done even though *some* of these alternative representations of public goods interaction (most notably the Assurance game) do not seem to present ‘problems’ in the sense which I think most people intend by use of the expression ‘collective action problems’. What then do we mean by this expression?

Jon Elster gives a ‘strong definition’ of the collective action problem, which identifies it with the Prisoners’ Dilemma, and a ‘weak definition’ which requires that (i) universal cooperation is preferred to universal non-cooperation by every individual (as in the Prisoners’ Dilemma) and (ii) cooperation is ‘individually unstable’ and ‘individually inaccessible’.<sup>25</sup> There is individual instability if each individual has an incentive to defect from universal cooperation, and there is

<sup>25</sup> Jon Elster, ‘Rationality, morality, and collective action’, *Ethics*, 96 (1985), 136–55. The weak definition, identifying collective action problems with the Prisoners’ Dilemma, is adopted by Elster in ‘Weakness of will and the freerider problem’, *Economics and Philosophy*, 1 (1985), 231–65— but then again he admits that ‘it does not... cover all the cases that intuitively we think of as collective action problems’.

individual inaccessibility if no individual has an incentive to move unilaterally from universal non-cooperation. But then he points out that there are cases in which cooperation is either individually unstable or individually inaccessible but not both — for example Chicken and Assurance games but which nevertheless present collective action problems (though ‘less serious’ ones in the case of Assurance games).

The definition which I think gathers up all the cases that Elster and others are actually concerned with is that a collective action problem exists where rational individual action can lead to a strictly Pareto-inferior outcome, that is, an outcome which is strictly less preferred by every individual than at least one other outcome. The problem with this definition — an unavoidable problem, it seems to me, if one wants to give a general definition that covers *all* the cases one intuitively thinks of as collective action problems — is that it’s not clear in some situations what rationality prescribes (even if we rule out, as I am assuming we should do here, notions of rationality not considered by game theorists). This is true of Chicken games. Any outcome of a Chicken game, including the Pareto-inferior mutual Defection outcome, can be rationalized. Hence, rational action *can* plausibly lead to a Pareto-inferior outcome, so that on my account it is a collective action problem.

Whether Assurance games are collective action problems again depends on what one takes rationality to prescribe. I shall take the view that, if a game has multiple equilibria (as the Assurance game does) but one of them is strictly preferred to all the others by everyone, then the Pareto-preferred one will be the outcome. On this view, rational action in an Assurance game does *not* lead to a Pareto-inferior outcome, so that this game is not a collective action problem.

Since preferences in some public goods interactions are those of an Assurance game, not all such interactions are collective action problems.

In the case of the (one-shot) Prisoners’ Dilemma, rational action unequivocally leads to a Pareto-inferior outcome, so on my account all situations representable as Prisoners’ Dilemmas are collective action problems. So are many other games (some of which will be encountered in the next chapter). Of course, not all of these games (including the Prisoners’ Dilemma and Chicken games) correspond to public goods interactions.

Elster has said that politics is ‘the study of ways of transcending the Prisoners’ Dilemma’.<sup>26</sup> In the light of this discussion of the collective action problem’ (and in anticipation of the discussion of alternatives to the Prisoners’ Dilemma in the next chapter), I think we should be a little more expansive and say that politics is the study of ways of solving collective action problems.

## Time and the lone exploiter

It’s worth noting parenthetically that the degradation of a ‘common’ may *not* be the result of failure to solve a collective action problem. It may occur even where the common has only one user and he acts rationally.

The ‘tragedy of the commons’, on Garrett Hardin’s account, arises because, at any point in time, each individual finds it in his interest to exploit the common (choose strategy *D*) no matter what the others do. The ‘tragedy’ does *not* arise, as some people have written, because each man reasons that ‘*since* the others are going to ruin the common anyway, I may as well exploit it too’.

---

<sup>26</sup> Jon Elster, ‘Some conceptual problems in political theory’, in Brian Barry (ed.), *Power and Political Theory* (London: Wiley, 1976), at pp. 248–9.

(In fact, if the others do *not* exploit the common, if they restrain themselves and choose strategy C, then each individual will find it even more profitable to exploit it than if they do. ) It cannot be said, then, that the common would not be ruined if only *one* individual had access to it ; that if a lake and all its lakeside factories were owned by one man, he would treat his wastes before discharging them into the lake; that if one man had an exclusive right to kill whales, he would see that they did not become extinct.

But surely, it may be said, the sole hunter of whales would not kill them all off, for his whole future livelihood, or at least all his future profits, depends on their survival. Unfortunately, this may not be the case.

Consider a common which one man has exclusive rights to exploit without restraint, and suppose now that at some point in time he is contemplating his whole future course of action with respect to this common. Let us suppose that he divides the future into equal time periods (months, years or whatever) and that in each time period he will receive a payoff. The sequence of payoffs he will receive depends on the course of action he chooses (for example, how many whales he kills in each period). Clearly, what he chooses to do will depend on the *present* value to him of future payoffs. At one extreme he may place no value whatever on any payoff except the one in the time period immediately before him. In this case, the prospect of zero payoffs from some point in the future onwards (as a result of the extinction of the whales, for example) does not trouble him at all. He will act in each time period so as to maximize his payoff in the current time period, and the result may be the ruin of the common.

It is generally assumed that future payoffs are exponentially *discounted* to obtain their present values. In the case when future time is divided into discrete periods, this means that the present value of a payoff  $X_t$  to be made  $t$  time periods from the present is  $X_t a^t$ , where  $a$  is a number such that  $0 < a < 1$  and  $1 - a$  is called the *discount rate*. The higher the discount rate, the lower the present value of future payoffs. If, for example, the individual's discount rate is 0.1 (that is,  $a = 0.9$ ), then a payoff worth 100 units if received now would have a present value of 90 if it were to be received one period hence, 81 if it were to be received two periods hence and so on.

Intuitively, we should expect that if the discount rate is sufficiently high, then, an exploiter who is seeking to maximize present value may eventually and quite 'rationally' ruin the common, even in the absence of other exploiters. The simple mathematics of this are set out by Colin Clark in his study of the exploitation of renewable resources (which, it should be remembered, include atmospheric, soil and water resources as well as such things as whales, fish and bison).<sup>27</sup> In the case when the resource of the common is a biological population, the discount rate which is sufficiently high to result in the extinction of the population will depend above all on the reproductive capacity of the population. (In Clark's model, this is *all* it depends on.)

The ruin of the common by a single individual, though it may be unfortunate, is not a 'tragedy' in Hardin's sense. (In the 'tragedy of the commons', the tragedy resides in the fact that 'rational' action on the part of each individual brings about a state of affairs which nobody wants.) Nor would it be a 'tragedy' if several individuals with similar preferences, including a shared high discount rate, ruined the common together, for this outcome would not be Pareto-inferior for them.

---

<sup>27</sup> Colin Clark, 'The economics of overexploitation', *Science*, 181 (17 August 1973), 630-4.

## Solutions? Community, states, entrepreneurs, property rights and norms

There are, broadly speaking, two sorts of solution to collective action problems, which I will call 'spontaneous' or 'internal' solutions and 'external' solutions. *Internal* solutions neither involve nor presuppose changes in the 'game', that is, in the *possibilities* open to the individuals (which are in part determined by the 'transformation function', specifying how much of the public good can be produced with a given contribution), the individuals' preferences (or more generally *attitudes*), and their *beliefs* (including expectations). *External* solutions, on the other hand, work by changing the game, that is, changing people's possibilities, attitudes or beliefs. The changes do not necessarily originate outside the group of individuals who have the collective action problem. Since individual action is the product directly of the individual's possibilities, attitudes and beliefs, these two exhaust the possible sorts of solution.

It could be said that in the case where an internal 'solution' is forthcoming, there was no 'problem' there to solve. For example, if the 'problem' is correctly modelled as a dynamic game which, though it consists let us say of an iterated Prisoners' Dilemma, is not itself a Prisoners' Dilemma and as a consequence the outcome produced by rational egoists (without any external assistance or other interference) would be mutual cooperation throughout the game, then it could be said that preferences *{including intertemporal preferences}*, etc., are such that there is no collective action *problem*. This would be a perfectly reasonable use of the word problem, but I shall not adopt it here. In fact, I shall take the view that the internal solution is the *basic* one, in two connected senses. It is, first, the only one which is complete in itself. All the external solutions presuppose the prior and/or concurrent solution of other problems, usually (always?) of collective action problems. Many of them, for example, involve the use of threats and offers of sanctions, and the creation and maintenance of the sanction system entail the prior or concurrent solution of collective action problems. (Why, for example, should the rational egoist pay his portion of the taxes that the state requires to maintain its police forces, etc., or why should the individual member of a community go to the trouble of punishing a free rider when he could be a free rider on the sanctioning efforts of others?) The internal solution is basic in a second sense: until we know whether a solution of this kind is possible and what form it will take, we cannot say what work, if any, remains to be done by other putative solutions. Thus, understanding the prospects for and obstacles in the way of an internal solution helps us to see what sorts of external solution are necessary and are likely to emerge in a given context.

External solutions can themselves be divided into two broad categories, which for short I will call centralized and decentralized ; or, better, they can be arrayed along a continuum running from perfectly centralized to perfectly decentralized. Combinations of them are possible normal, in fact. A solution is *decentralized* to the extent that the initiative for the changes in possibilities, attitudes or beliefs that constitute an external solution is dispersed amongst the members of the group; or, the greater the proportion of the group's members involved in solving the collective action problem (e.g. applying sanctions to free riders), the more decentralized the solution. Contrariwise, a solution is *centralized* to the extent that such involvement is concentrated in the hands of only a few members of the group.

Centralized solutions are typified, of course, by *the state*, while decentralized solutions characterize *community*. I have devoted another book to the ways in which a community can provide

itself with public goods without the help of the state and will not reproduce the arguments here.<sup>28</sup> By a community I mean a group of people (i) who have beliefs and values in common, (ii) whose relations are direct and many-sided and (iii) who practise generalized as well as merely balanced reciprocity. The members of such a group of people, or all of its active adult members, can wield with great effectiveness a range of positive and negative sanctions, including the sanctions of approval and disapproval the latter especially *via* gossip, ridicule and shaming. Decentralized solutions can sometimes be effective where there is little community, but the size of the group would still have to be relatively small (as it must be in a community).

External solutions are not necessarily restricted to the use of threats and offers of positive and negative sanctions. These, it is true, work not by altering an individual's preferences among outcomes (properly defined) but by altering his expectations about the actions to be taken by others (and hence the expected utility associated with alternative courses of action). But there are other ways in which an individual's expectations about others' behaviour can be altered and other ways in which he can be got to contribute to a public good, without the use of threats and offers or of force, whether centralized or decentralized. These include persuasion — providing information and arguments about the alternatives, about the consequences of adopting the various courses of action, about others' attitudes and beliefs and so on. Such methods are characteristic of the political entrepreneur, an external solution (relatively centralized, though usually closely combined with decentralized mechanisms) which I shall discuss shortly.

My main concern in this book is with the internal solution — with the possibility of spontaneous cooperation as an alternative to the state. But before turning to this, I want to comment briefly, first, on the role of the political entrepreneur in the solution of collective action problems, and secondly, on the claims made by a number of writers that certain collective action problems can be solved by establishing private property rights and by norms.

## Political entrepreneurs

In what sense do political entrepreneurs or leaders 'solve' collective action problems? In general, to solve or remove a collective action problem he or she must of course change individual preferences (or more generally attitudes), or change beliefs (including expectations) or inject resources (very probably knowledge, or new technology, like guns) into the group so as to make its members' efforts more productive.

Merely offering his services (working to obtain the public good) in exchange for support (subscriptions, food and shelter, or whatever) does not in itself constitute a distinctive solution to the problem. For, in the first place, the entrepreneur's services are themselves a public good, so that supporting him also gives rise to a collective action problem. This includes the case of the politician who in seeking electoral support offers his constituents legislative or other changes they favour. The collective action problem his potential supporters had in obtaining the public goods which such changes would have brought them is replaced by the collective action problem of getting him elected. And secondly, if the entrepreneur gains support by offering selective incentives, as well as by promising to work for the public good, then the solution is precisely the one proposed by Olson himself, in his 'by-product' theory.<sup>29</sup>

<sup>28</sup> Taylor, *Community, Anarchy and Liberty*.

<sup>29</sup> The two points in this paragraph were made by Brian Barry in *Sociologists, Economists and Democracy* (London: Collier-Macmillan, 1970) at pp. 27–39.

In many interesting cases the political entrepreneur may require little or no support from the members of the group whose collective action problem is at issue, because he is supported by (i.e., brings resources from) some external source. He might, for example, in his efforts to solve a *local* collective action problem, be supported by a pre-existing organization (the Communist Party, say, or the Catholic Church). This makes it easier to explain why the local problem is solved (for the members of the local group do not have to produce a 'surplus' to pay or feed the entrepreneur), but it leaves unexplained (a) the production of the resources brought in by the political entrepreneur, which will usually entail that a prior collective action problem — for example the creation and maintenance of an organization — has been solved; and (b) how, even though the (local) group does not have to support the entrepreneur, it now manages to solve a collective action problem that it could not solve without him. If the only difference the entrepreneur makes is the addition of selective incentives to their benefits, then, once again, we do not have a distinctive solution.

But the political entrepreneur is not just 'an innovator with selective incentives',<sup>30</sup> or someone who simply concentrates or centralizes resources. What is perhaps more characteristic of political entrepreneurship is its role in changing beliefs — beliefs about the public good itself, about what others have done and are likely to do and about others' beliefs. Above all, we must remember that most collective action must involve some form of conditional cooperation, for at a minimum an individual would not cooperate if *nobody* else did. And as we shall see (in chapters 3 and 4) conditional cooperation is a very precarious business. It requires amongst other things that the conditional cooperators have information about others' behaviour. The required monitoring can be done by the political entrepreneur.

The entrepreneur can also try to persuade people that their contributions make a big enough difference, either directly or indirectly through their effect on others' behaviour. The second of these might be achieved by persuading people that others' efforts are contingent on theirs.

An organization whose aim is to provide public goods for a very large group might be able to expand its membership and achieve its aims by having its cadres work to solve, through any or all of these entrepreneurial methods, smaller-scale collective action problems for much smaller subgroups. A nationwide movement, for example, may be built upon the success of its cadres in solving local collective action problems and bringing tangible benefits quickly. Samuel Popkin has given an excellent account of activities of this sort in Vietnam, showing how four politico-religious movements (the Catholic Church, the Cao Dai, the Hoa Hao and the Communist Party) won support by having their cadres help the villages, both by providing selective incentives and by facilitating cooperation in the provision of public goods.<sup>31</sup> These private and public goods — with varying degrees of indivisibility and excludability — included the provision of educational opportunities; the creation of insurance and welfare systems; agricultural improvements; the establishment of stock-farm cooperatives; improvements in water storage and irrigation facilities; the creation of local courts to arbitrate disputes; and protection against French courts, marauding notables and local landlords.

---

<sup>30</sup> Olson, *The Logic*, Appendix added in 1971, p. 177.

<sup>31</sup> Samuel L. Popkin, *The Rational Peasant: The Political Economy of Rural Society in Vietnam* (Berkeley: University of California Press, 1979), especially ch. 3.

## Property rights

Many economists, and nearly all those of the 'property rights' school, believe that the solution to free rider problems in public goods provision, and in particular those which would lead to the over-exploitation of a 'common property resource', lies in the establishment of private property rights. Without such rights, the argument goes, every individual has an incentive to intensify his use of the resource because (as we saw in discussing Garrett Hardin's 'tragedy of the commons') although, with each increment in use, every unit of his (and everybody else's) input becomes slightly less productive, this is up to a point outweighed by the marginal return from the increased input. Intensifying use of the resource is continued up to the point where all the 'rent' (income or other return) from the resource has been dissipated. Likewise, the benefits arising from any improvement or renewal or other investment he might make in the resource would be shared by all the users while the costs would be borne by the individual alone. There will therefore be overuse and underinvestment. With the establishment of private property rights, however, the external effects of each individual's actions are 'internalized' : *all* the costs of an increase in use of the resource are borne by the individual, as are all the benefits of investing in its conservation or improvement.

The argument that the 'tragedy of the commons' is the fate of common property resources, and that overuse or underinvestment will be avoided only if common property rights are displaced by private property rights, seems to be positively mocked by the facts. The commons of the European open field system, far from being tragically degraded, were generally maintained in good health during the whole of their lifetimes of many hundreds of years. There is a detailed study of a Swiss alpine village (not, of course, operating an open field system) whose members have for more than five hundred years possessed and used in common various resources, including mountain-side pastures, side by side with privately owned land and other resources and during all this time the productivity of the common land has been maintained and much effort has been invested in its improvement.<sup>32</sup> Contrast with this the treatment, especially in recent decades, of much privately owned land by its very owners: the destruction of vast tracts of rain forest for the sake of a few profitable years of ranching; or the set of practices which together are causing the loss of topsoil from cultivated land through wind and water erosion on such a scale that, according to a recent report, there will be a third less topsoil per person by the end of the century.<sup>33</sup> In parts of Africa, and elsewhere in the world, overexploitation of grazing lands has been caused not by common property arrangements *per se* but by their destruction or disruption.<sup>34</sup> There are, as we saw earlier, perfectly good reasons why the rational private owner or user of a resource might knowingly destroy it; in particular, he might place a very low value on benefits to be derived from the resource in the distant as opposed to the immediate future.

---

<sup>32</sup> Robert McC. Netting, *Balancing on an Alp: Ecological Change in a Swiss Mountain Community* (Cambridge: Cambridge University Press, 1981), especially ch. 3.

<sup>33</sup> Lester Brown and Edward Wolf, *Soil Erosion: Quiet Crisis in the World Economy* (Washington, D.C. : Worldwatch Institute, 1984). According to this report, U.S. farms are losing topsoil at the rate of 1.7 billion tonnes a year. *The New York Times* (10 December 1985) reports that the U.S. Congress looks set to vote to pay farmers to stop farming up to 40 million acres of the worst affected land.

<sup>34</sup> See, for example, Michael H. Glantz (ed.), *Desertification: Environmental Degradation In And Around Arid Lands* (Boulder, Colorado : Westview Press, 1977).

Where do the property rights economists go wrong?<sup>35</sup> In the first place, many of them do not distinguish common property in a resource from *open access* to it. ‘Communal rights’, say Alchian and Demsetz, ‘... means that the use of a scarce resource is determined on a first-come, first-serve basis and persists for as long as a person continues to use the resource’.<sup>36</sup> This is wrong, or at least an abuse of language. If there is open access, then nobody is excluded from using the resource and there is no regulation of the activities of those who do use it. But if there is common ownership or collective control of the resource, then the members of the collectivity, whatever it is, can regulate its use. This is what happened in the European open field system, where the villagers rigorously excluded outsiders from use of the various commons they owned or possessed collectively, and carefully regulated insiders’ use, typically by allotting to individuals ‘stints’ in proportion to their (privately owned) arable holdings and punishing people for infringements. The alpine community described by Netting similarly practised strict external and internal regulation of its commons. So too have countless ‘primitive’ collectivities and peasant villages all over the world.

It is to resources with open access, not to ‘common property resources’, that the property rights economists’ argument about over-exploitation and underinvestment applies. It is not a matter of establishing the right sort of property rights, of moving from collective to private property rights. It is rather a matter (at this stage of the argument at least) of establishing property rights where there were none ; for property entails exclusion, so that where there is open access to a resource, there is no property in it<sup>37</sup>

The property rights economists tend to see only two or three possibilities: open access and private property, to which is sometimes added state ownership. But almost any group of individuals can own or possess property collectively. Historical and contemporary examples are: a family; a wider kin group, such as a matrilineage; all those in a village who also possess land privately; a band; an ethnic group. Where the property rights economists do notice common property rights, they then argue that the costs of negotiating agreements regulating use and, if agreements are forthcoming, the costs of policing them, will be very great, and in this respect, common property rights compare unfavourably with private property rights.<sup>38</sup> But there is no necessary reason why transaction costs of all kinds should in total be greater in the case of common property rights than in the case of private property rights and in the case of the open field system it was in fact the other way round, essentially because of economies of scale in pastoral production.<sup>39</sup>

Finally, the property rights economists, having generally failed to notice common property (as opposed to open access) and to study how individual rights in it are guaranteed, tend to assume that property rights must be enforced by the state.<sup>40</sup> But there can also be decentralized enforcement or maintenance of property rights — both private and common. (The sense of ‘decentralized’

---

<sup>35</sup> The following comments, which are critical of the property rights school’s treatment of the ‘tragedy of the commons’, do not imply a wholesale rejection on my part of the property rights approach.

<sup>36</sup> A. A. Alchian and Harold Demsetz, ‘The property rights paradigm’, *Journal of Economic History*, 33 (1973), 16–27.

<sup>37</sup> See S. V. Ciriacy-Wantrup and Richard C. Bishop, ‘“Common property” as a concept in natural resources policy’, *Natural Resources Journal*, 15 (1975), 713–27.

<sup>38</sup> See, for example, Harold Demsetz, ‘Toward a theory of property rights’, *American Economic Review* (Papers and Proceedings), 57 (1967), 347–59.

<sup>39</sup> Carl J. Dahlman, *The Open Field System and Beyond* (Cambridge: Cambridge University Press, 1980).

<sup>40</sup> See, for example, Eirik G. Furubotn and Svetozar Pejovich, ‘Property rights and economic theory: a survey of recent literature’, *Journal of Economic Literature*, 10 (1972), 1137–62.

intended here is the same as that used in the general remarks made earlier on the solution of collective action problems.) If a collectivity itself is to enforce its members' private property rights or their rights to use the common property, then it must of course be able to wield effective sanctions — unless the property rights are respected as a result of 'spontaneous' conditional cooperation. If the collectivity is a community, then, as we have seen, conditions are conducive to conditional cooperation, and if this fails the community's members have at their disposal a range of effective sanctions. The joint owners of the commons in European open field villages, for example, were communities in the required sense.

Enough has now been said, I think, to show that, insofar as the solution of collective action problems is concerned, nothing new is added by the introduction of property rights *per se*. An individual has property in something only if others forbear from using it, and the forbearance is the result of the threat or offer of *sanctions*, centralized or decentralized (or of conditional cooperation — unless this be reckoned also to involve threats and offers). It is the threats and offers of sanctions (and/or conditional cooperation) that is solving the collective action problem, if it is solved at all. Furthermore, as I remarked in an earlier section, the use of some of these sanctions presupposes the solution to prior collective action problems (for example, the formation and maintenance of a state!).

## Norms

There is, finally the suggestion that norms solve collective action problems. I will comment on this very briefly, for my reaction to it is similar to my view of the suggestion that the introduction of private property rights solves collective action problems, and both follow from the general remarks about solutions to these problems made in an earlier section (though I shall not argue, as some have done, that property rights *are* norms). The view that norms solve collective action problems or more precisely that they solve, amongst other things, the problems inherent in 'generalized PD-structured situations' and coordination problems — has been expounded by Edna Ullman-Margalit.<sup>41</sup> I shall take it that a norm is generally conformed to and is such that nonconformity, when observed, is generally punished. It is unclear whether this is what Ullman-Margalit means by a norm, but in any case it is fairly clear from her discussion of 'PD norms' that it is only 'a norm, *backed by sanctions*' or 'a norm ... *supported by sufficiently severe sanctions*' that is capable of solving Prisoners' Dilemma problem.<sup>42</sup> So norms alone — mere prescriptions for action that people generally conform to — do *not* solve these problems.

If a norm is generally observed simply because it pays the individual to do so (in the absence of sanctions), then there is no (collective action or other) 'problem' to be solved in the first place. This would be the case if the norm had been 'internalized'. I take this expression to indicate that conformity to the norm does not require the application of external sanctions, inducements or any other considerations; as a result of the norm being internalized, the individual *prefers* to conform (without the threat of punishment) or at least has some sort of motivational disposition

---

<sup>41</sup> Edna Ullman-Margalit, *The Emergence of Norms* (Oxford: Clarendon Press, 1977). A 'generalized PD-structured situation ... is one in which the dilemma faced by the ... participants is recurrent, or even continuous' (p. 24); but Ullman-Margalit gives no analysis of iterated games or takes any account of their distinctive problems (so, amongst other things, does not see that cooperation in these situations can occur *without* norms enforced by sanctions). Incidentally, very little of this book actually deals with the emergence of norms; it is mainly taken up with generally informal discussion of some very simple games.

<sup>42</sup> *The Emergence of Norms*, pp. 22 and 28; my emphasis.

to do so. But then, as I say, we would not say that there was a Prisoners' Dilemma or collective action 'problem' to be solved: the individual preferences would not be those of a Prisoners' Dilemma or would not be such as to lead to a collective action problem. Of course, we might nevertheless wish to explain how the norm came to be internalized, or how people came to have such preferences.

If, on the other hand, a norm is generally observed because nonconformity, when noticed, is generally punished, then it is the sanctions that are doing the real work of solving the Prisoners' Dilemma or collective action problem. The sanction system can of course be centralized or decentralized, in the way discussed in an earlier section. And again, it remains to be explained how the system of sanctions itself came into being and is maintained. To this problem, the general point made earlier about sanction systems applies: the maintenance of a system of sanctions itself constitutes or presupposes the solution of another collective action problem. Punishing someone who does not conform to a norm punishing someone for being a free rider on the efforts of others to provide a public good, for example — is itself a public good for the group in question, and everyone would prefer others to do this unpleasant job. Thus, the 'solution' of collective action problems by norms presupposes the prior or concurrent solution of another collective action problem. And as my earlier remarks make clear, this would still be the case if the sanctions were wielded by the state or by a political entrepreneur.

## Plan of the rest of the book

My purpose in this book is to examine the possibility of voluntary cooperation in the provision of public goods and in the solution of other collective action problems, and in doing so and in other ways to raise questions about what I take to be the most persuasive justification of the state. The detailed study of voluntary cooperation which follows (chapters 2, 3 and 4) can be read and evaluated independently of the critique of the liberal theory of the state. Both as a study of cooperation and as a study of the state and its alternatives, it is obviously far from complete; another part of the story is tackled in my *Community, Anarchy and Liberty*, which is complementary to this book.

As a critique of the liberal justification of the state, the argument will be in three stages. First, I argue in chapter 2 that in public goods interactions the individual preferences at any point in time are not necessarily those of a Prisoners' Dilemma game. This is true, I shall argue, of both two-person and N-person games and of cases where strategy sets are continuous as well as those where individuals have only two strategies available to them. It will emerge that important classes of public goods provision problems are better represented by Assurance and especially Chicken games, and in the continuous case by hybrids of these two. In all these games, arguably, if the game is played only once, *some* cooperation is more likely to be forthcoming than in cases for which the Prisoners' Dilemma is the appropriate model.

In the next two chapters (3 and 4), however, I shall assume the worst: that preferences at any point in time are those of a Prisoners' Dilemma game. But I then go on to show that if *time* is introduced and the problem is treated more dynamically, under certain circumstances voluntary Cooperation is rational for each player, even assuming that he seeks to maximize only his own payoff.

My argument here will be cased in terms of the Prisoners' Dilemma *supergame*. This is the game consisting of an indefinite number of iterations of one of the Prisoners' Dilemma games (two-person and  $N$ -person) which were defined earlier. In each *constituent game* (as the repeated game is now called), players choose strategies simultaneously, as before, but they know the strategies chosen by all other players in previous games. Each player discounts future payoffs; his discount rate does not change with time, but discount rates may differ between players. The constituent game is assumed not to change with time. (It would be desirable to relax this last assumption in a more general treatment, and permit the payoff matrix to change with time. See the final section of chapter 4 below.)

The really important difference between the one-shot game and the supergame is that players' strategies can be made interdependent in the latter but not, of course, in the former, since players must choose strategies simultaneously or in ignorance of each other's choices. In the supergame, a player can, for example, decide to Cooperate in each constituent game *if and only if* the other player(s) Cooperated in the previous constituent game. It is on this possibility, the possibility of using *conditional* strategies, that the voluntary Cooperation of all the players turns.

Finally, in chapter 7, I shall raise doubts about the way in which this justification of the state is approached. It is an essential and fundamental feature of the theory I am criticizing that it takes individual preferences as given and fixed. In particular, it is assumed that the state itself has no effect on these preferences. This rules out *ab initio* the possibility, amongst many others, that the state may exacerbate an already existing collective action problem or create such a problem where none existed before: that the state may affect, in other words, the very conditions which are supposed to make it necessary. If preferences may change, especially as a result of the activities of the state itself, it is not at all clear what is *meant* by the desirability of the state.

Criticisms of this sort can of course be levelled against any theory which is founded on assumptions about fixed individual preferences (as most of economic theory and some political theory is); but they are especially important, it seems to me, when the theory purports to justify an institution (like the state) and when the theory is to apply to a very long period of time (as a theory used to justify the state or to explain its origin must do).

I have said that the arguments which are the object of my criticisms in this book have been set out most explicitly by Thomas Hobbes. I shall therefore give (in chapter 6) an exposition of his political theory. My chief reason for devoting to this exposition a rather long chapter later in the book, rather than a short summary at the start of the book where it would otherwise belong, and for making what would otherwise be an unpardonable addition to the considerable critical literature on Hobbes, is that I think it is illuminating to look at these theories in terms of some of the ideas presented in the earlier chapters on the Prisoners' Dilemma and its supergame. I have asserted rather baldly in this informal Introduction that Hobbes's theory is about non-Cooperation in Prisoners' Dilemma games (other writers have made similar assertions, equating Hobbes's theory with, for example, Hardin's analysis of the 'tragedy of the commons'); but the story is more complicated and more interesting than this and deserves a fuller account.

I shall also consider, more briefly, David Hume's political theory. For although it is very similar to Hobbes's theory (despite Hume's objections to what he took to be a fundamental element of Hobbes's theory, the idea of the social contract) and although it is generally less rigorous than Hobbes's version (in *Leviathan*), it does partly supply a deficiency in Hobbes's treatment, namely that it is too static. Hobbes in effect treats only a one-shot Prisoners' Dilemma game, whereas Hume's treatment is more dynamic, with the discounting of future benefits playing an important

role. Also, in Hume, but not in Hobbes, there is explicit recognition of the effects of size, a partial anticipation of Olson's 'logic of collective action'.

Some of the ideas I am interested in here appeared much earlier than *Leviathan* (above all in the *Book of Lord Shang* and the works of Han Fei Tzu which were written in China in the fourth and third centuries BC), but it was Hobbes and Hume who gave the first full, explicit statements of the argument. And in later political theorists the argument is not always explicit, does not stand out boldly and is less precise and less coherent.

In *Leviathan*, Hobbes seems to assume that each man seeks to maximize not merely his own payoff, but also his 'eminence', the difference between his own and other people's payoffs. Hume, on the other hand, assumes that most people are chiefly concerned with their own payoffs but are also possessed of a limited amount of 'benevolence'. In both cases, individuals take *some* account of other individuals' payoffs; I call this 'altruism'. The effects of various sorts of altruism on the outcomes of Prisoners' Dilemma games are treated briefly in chapter 5. Some of the material in that chapter will be of use in the discussion of Hobbes and Hume and also in the final chapter.

The Anarchist Library  
Anti-Copyright



Michael Taylor  
Introduction to 'The Possibility of Cooperation'  
The Problem of Collective Action  
August 1987

<[www.cambridge.org/gb/universitypress/subjects/philosophy/political-philosophy/possibility-cooperation](http://www.cambridge.org/gb/universitypress/subjects/philosophy/political-philosophy/possibility-cooperation)>  
ISBN 9780521339902

Cambridge University Press

**[theanarchistlibrary.org](http://theanarchistlibrary.org)**