Hockey Helmets, Concealed Weapons, and Daylight Saving

A STUDY OF BINARY CHOICES WITH EXTERNALITIES

THOMAS C. SCHELLING
John Fitzgerald Kennedy School of Government
Harvard University

Shortly after Teddy Green of the Bruins took a hockey stick in his brain, *Newsweek* (1969) commented:

> Players will not adopt helmets by individual choice for several reasons. Chicago star Bobby Hull cites the simplest factor: "Vanity." But many players honestly believe that helmets will cut their efficiency and put them at a disadvantage, and others fear the ridicule of opponents. The use of helmets will spread only through fear caused by injuries like Green's–or through a rule making them mandatory . . . One player summed up the feelings of many: "It's foolish not to wear a helmet. But I don't–because the other guys don't. I know that' silly, but most of the players feel the same way. If the league made us do it, though, we'd all wear them and nobody would mind."

The most telling part of the *Newsweek* story is in the declaration attributed to Don Awrey. "When I saw the way Teddy looked, it was an awful feeling . . . I'm going to start wearing a helmet now, and I don't care what anybody says." Viewers of Channel 38 (Boston) know that Awrey does *not* wear a helmet.

Introduction

This paper is about binary choices with externalities. These are either-or situations, not choices of degree or quantity.

An "externality" is present when you care about my choice or my choice affects yours. You may not care, but need to know–whether to pass on left or right when we meet. You may not need to know, but care–you will drive whether or not I drive, but prefer that I keep off the road. You may both care and need to know.

Journal of Conflict Resolution, Vol. 17 No. 3, September 1973,

The literature of externalities has mostly to do with how much of a good or a bad should be produced, consumed, or allowed. Here I consider only the interdependence of choices to do or not to do, to join or not to join, to stay or to leave, to vote yes or no, to conform or not to conform to some agreement or rule or restriction.

Joining a disciplined, self-restraining coalition, or staying out and doing what's natural is a binary choice. If we contemplate all the restraints that a coalition might impose, the problem is multifarious; but if the coalition is there, and its rules have been adopted, the choice to join or not to join is binary. Ratifying a nuclear treaty or confirming a Supreme Court Justice is multifarious until the treaty is drafted or the Justice nominated; there then remains, usually, a single choice.

Paying or not paying your share is an example, as is wearing a helmet in a hockey game. So is keeping your dog leashed, voting yes on ABM, staying in the neighborhood or moving out, boycotting Rhodesia or lettuce, admitting women, deserting from the army or recognizing the People's Republic of China; signing a petition, getting vaccinated, carrying a gun or liability insurance or a tow cable; driving with headlights up or down, riding a bicycle to work, shoveling the sidewalk in front of your house, or going on daylight saving time. The question is not how *much* anyone does but how *many* make one choice or the other.

CONFIGURATIONS

If the number of people is large, the configurations of externalities can be variegated. Everybody's payoff may depend on what each particular individual does: for each among n + 1 individuals, there are 2^n possible environments generated by all the others. The situation is simpler if it has some structure: everybody's payoff may depend only on the choices of people upstream; it may be an additive function of what everyone else does; or everyone may have a "receiving" and a "transmitting strength," and the signal received by anyone is his own receiving strength times the sum of the transmitting strengths of all who broadcast. (People are ranked according to the smokiness of their furnaces and the amount of laundry they hang on the line.)

In some cases, the configuration matters. If everybody needs 100 watts to read by and a neighbor's bulb is equivalent to half one's own, and everybody has a 60-watt bulb, everybody can read as long as he and both his neighbors have their lights on. Arranged in a circle, everybody will keep his light on if everybody else does (and nobody will if his neighbors do not); arranged in a line, the people at the ends cannot read anyway and turn their lights off, and the whole thing unravels.

People can differ in their initial positions: if some cars have direction signals and some do not and there are installation costs, people will be subject to different thresholds. Payoffs may differ by order of choosing: rewards or costs of entry can change as a voting coalition grows or declines. If there are turnaround costs, speculation will matter: one is penalized if the fashion or coalition does not reach critical mass after all; maybe he loses if he defects too soon and has to buy his way back in.[1]

This paper considers only a simple set of situations—those in which people are identically situated both statically and dynamically. Everybody's payoffs, whichever way he makes his choice, depend only on the *number* of people who choose one way or the other. Everybody has the same transmitting and receiving strengths. There is no comparative advantage, no ranking by sensitivity or influence. The payoffs are the same for everybody; and if a fraction of the population chooses one way or the other, it does not matter which individuals comprise the fraction, or in what order they commit themselves to their choices. (Actually, as long as transmitting and receiving strengths are in the same ratio, doubling both for an individual leaves his own payoffs unaffected and makes him the equivalent of two people to all the rest; counting him as a coalition of two lets him fit this restrictive model.)

KNOWLEDGE AND OBSERVATION

If people need to know how others are choosing, it will matter whether or not they can see or find out. I can tell how many people have snow tires if I take a little trouble and look around; it is harder to know how many cars that may pass me in an emergency have tow chains. I have no way of knowing who is vaccinated, unless I ask people to roll up their sleeves; but my doctor can probably find the statistics and tell me. I have a good idea how many people regularly wear ties and jackets to work; but for special ceremonies it is hard to find out, until after I have made my choice, how many people are going black tie, or in sneakers.

Continuous or repeated binary choices, when they are readily visible

1. An intriguing account of complex interdependencies with n = 101 and an almost-binary choice—absence and abstention being possible alternatives—with differential transmitting and receiving strengths, varying degrees of reversibility of choice, incomplete and sometimes manipulated information, small networks of special influence, and nonuniform preferences among the participants, is Harris' (1970) story of the Senate's action on Judge Carswell.

and there are no costs in switching, may allow easy, continuous adjustment to what others are doing. Once-for-all choices are often taken in the dark. Some choices, like resigning in protest, are necessarily visible; some, like loaded guns and vaccination scars, can be revealed or concealed; some, like fouling or not fouling a public pond, may be not only invisible but unrevealable. For discipline and enforcement, it will usually matter whether individual choices or only the aggregates can be monitored. Unless I say otherwise, I shall usually have in mind that people can see and adapt to the choices of others without seriously biased misperceptions; but we should keep in mind that this is a special case, and often an especially easy one to deal with.

What they actually "see and adapt to" is sometimes not the numbers choosing one way or the other but the consequences. While the senator who votes against Judge Carswell probably cares directly about the number of negative votes, the owner of the double-parked automobile is more interested in the safety in numbers than in the numbers themselves. Parents who decline vaccination for their children should be interested in how much safety the vaccination of others provides, not in the numbers themselves, although they may have a more reliable estimate of numbers than of risk. The distinction between numbers per se and their consequences—which it is that one cares about, and which it is that one can observe—is a distinction that ought, in a particular case, to be explicit; but I shall usually speak as though it is the choices themselves that a person can see and that he cares about.

What we have, then, is a population of n individuals, each with a choice between L and R ("Left" and "Right") corresponding to the directions on a horizontal scale or, in an actual choice, the two sides of a road or the two sides of a divided legislature. For any individual, the payoff to a choice of Left or Right depends on how many others in a specified population—for the moment, a finite population—choose Left or Right. It is interesting to work with commensurable payoffs measured in lives, limbs, hours, dollars, or even "utility," so that we can talk about collective totals; it is easy to deform the results and drop back to ordinal relations. So there is a "physical product" interpretation that we can drop when we wish; it allows us to deal with mergers as well as with coalitions.

Prisoner's Dilemma

A good place to begin is the situation familiarly known—in its two-person version—as "prisoner's dilemma." It contains a binary choice

C
(chooses column)

R (chooses row)	1 1	2 -1
	-1 2	0 0

NOTE: Lower-left number in each cell denotes the payoff to R (choosing row), upper-right number the payoff to C (choosing column)

Figure 1.

for each of two people. Each has (1) a *dominant choice:* the same choice is preferred, irrespective of which choice the other person makes. Each has, furthermore, (2) a *dominant preference* with respect to the *other's choice:* his preference for the other person's action is unaffected by the choice he makes for himself. (3) These two preferences, furthermore, go in *opposite* directions: the choice that each prefers to make is not the choice he prefers the other to make. Finally, (4) the strengths of these preferences are such that both are better off making their dominated choices than if both made their dominant choices.

A representative matrix with uniform payoffs for the two individuals is in Figure 1. In that figure, the lower-left number in a cell denotes the payoff to R (choosing row), the upper-right number the payoff to C (choosing column).

The influence of one individual's choice on the other's payoff we can call the *externality.* Then the effect of his own choice on his own payoff can in parallel be called the *internality.* We then describe "prisoner's dilemma" as the situation in which each person has a uniform (dominant) internality and a uniform (dominant) externality, the internality and externality are opposed rather than coincident, and the externality outweighs the internality.

The situation is fairly simple to define.[2] But when we turn to the three-person or multiperson version, the two-person definition is ambig-

2. Not quite: sometimes the situation is further subdivided according to whether or not probabilities or alternating frequencies can be found, for the lower-left and upper-right cells, that offer expected values greater than 1 for both R and C, or greater than 0 for both R and C. Sometimes the definition is allowed to include, sometimes not to include, the limiting cases in which Row's payoffs in one

uous. "The other" equals "all others" when there are but two; with more than two, there are in-between possibilities. We have to elaborate the definition in a way that catches the spirit of prisoner's dilemma, and see whether we then have something distinctive enough to go by a proper name.

EXTENDING THE DEFINITION

There are two main definitional questions. (1) Are the externalities monotonic–is an individual always better off, the more there are among the others who play their dominated strategies? (2) Does the individual's own preference remain constant no matter how many among the others choose one way or the other–does he have a fully dominant choice? Tentatively answering, for purposes of definition, yes to these two questions, and assuming that *only numbers matter* (not identities), and that all payoff rankings are the same for all players, a *uniform multiperson prisoner's dilemma*–henceforth, MPD for short–can be defined as a situation in which:

(1) There are n individuals, each with the same binary choice and the same payoffs.

(2) Each has a dominant choice, a "best choice" whatever the others do. (And the same choice is dominant for everybody.)

(3) Whichever choice an individual makes, his dominant or his dominated, any individual is better off, the more there are among the others who make their dominated choices.

(4) There is some number k, greater than 1, such that, if individuals numbering k or more make dominated choices and the rest do not, those who make dominated choices are better off than if they had all made dominant choices, but, if they number less than k, this is not true. (The uniformity of participants makes k independent of the particular individuals making dominated choices.)

Some other questions occur but need not be reflected in this tentative definition. For example, (1) if the payoffs are cardinally and commensurably interpreted, so that we can deal with collective totals, does the *collective maximum* necessarily occur when all choose dominated strategies? Or (2) is the situation for *any subset* among the n individuals invariably MPD when the choices of the remainder are fixed? Or (3) if

column, and Column's payoffs in one row, are equal. The many-person counterparts to these distinctions will show up later.

subsets are formed and treated as coalitions that make bloc choices, does the relation among the coalitions meet the definition of prisoner's dilemma given above? And (4) do we include the limiting cases in which, if n − 1 individuals all choose the same, the nth individual is indifferent to his own choice? These questions are better dealt with as part of the agenda of analysis than as definitional criteria. Since one of the conclusions of the analysis that follows is that the prisoner's dilemma situation is not as distinctive when n is large as when it equals 2, not much is at stake in this initial definition.

A DISTINGUISHING PARAMETER

Taking the four numbered statements as a plausible extension of the prisoner's dilemma idea, and as what I shall mean by MPD when I use the term in this paper, we have at first glance an important parameter, k. It represents the minimum size of any coalition that can gain by making the dominated choice. If k is equal to n, the only worthwhile coalition—the only enforceable contract that is profitable for all who sign—is the coalition of the whole. Where k is less than n, it is the minimum number that, though resentful of the free riders, can be profitable for those who join (though more profitable for those who stay out).

On a horizontal axis measured from 0 to n, two payoff curves are drawn. (We switch, for convenience, to a population of n + 1, so that n will stand for the number of "others" there are for any individual.) One curve corresponds to the dominant choice; its left end is called 0 and it rises to the right, perhaps leveling off but not declining. Below it, we draw the curve for the dominated choice. It begins below 0, rises monotonically, perhaps leveling off, and crosses the axis at some point denoted by k. We use L (Left) to stand for the dominant strategy, R (Right) for the dominated. The number choosing Right on the diagram is denoted by the distance of any point rightward from the left extremity. At a horizontal value of n/3, the two payoff curves show the value to an individual of choosing L or R when a third of the others choose R and two-thirds, L.

ILLUSTRATIVE CURVES

Figure 2 shows several pairs of curves that meet the definition. (The dotted lines will be introduced in a moment.) The only constraint on these curves, under our definition, is that the four extremities of the two curves be in the vertical order shown, that the curves be monotonic, and that the

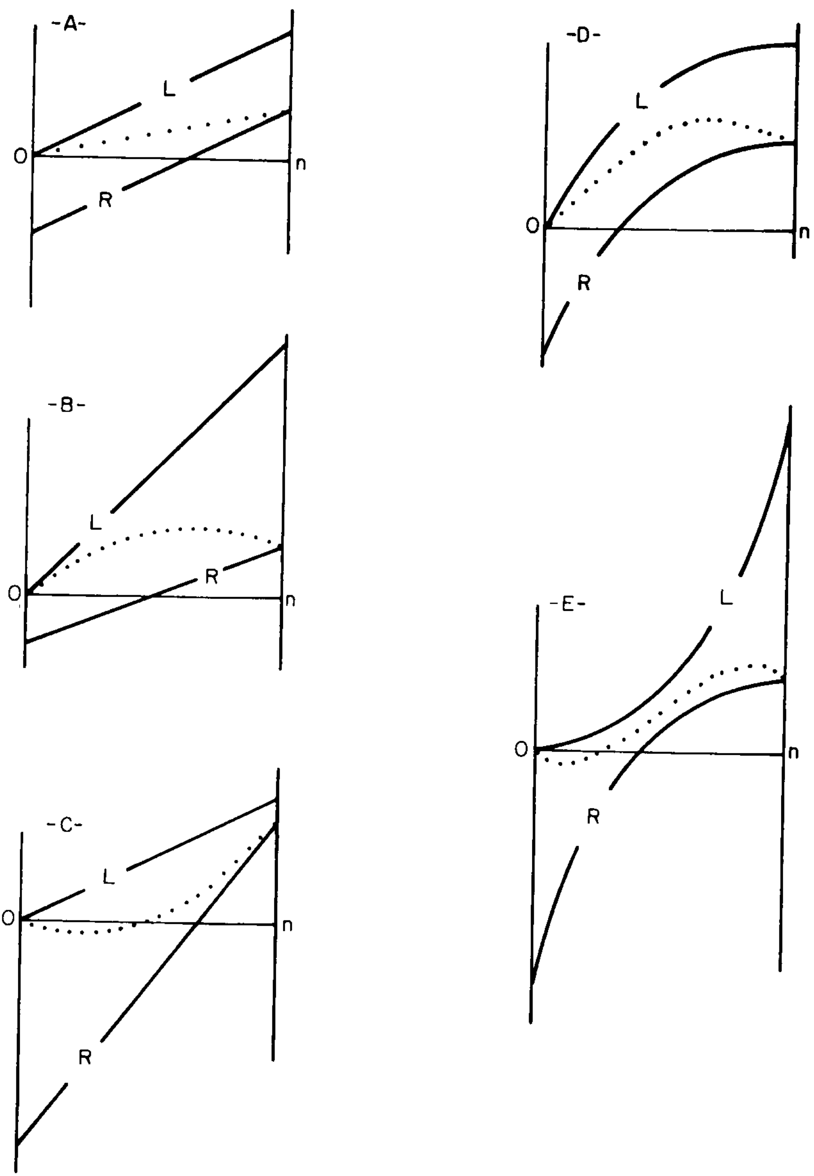

Figure 2.

curves not cross. The *not crossing* represents dominance of choice. (The "internality" is uniform.) *Monotonicity* for both curves in the *same* direction denotes uniformly positive externalities for a Right choice (or uniformly negative for a Left choice). That the *Left* curve is higher while both rise to the *right* reflects the opposition of internality and externality. Finally, the Right curve is higher on the right than the Left curve on the left, reflecting the inefficiency of the uniformly dominant choice. Later, we shall experiment with curves that cross, curves of opposite direction, curves that are vertically interchanged, and curves whose end points (and slopes and curvatures) are differently configured.[3]

The "values" accruing to Right and Left choices for different individuals may or may not be susceptible to some common measure. Reactions to smells, noises, and other irritants cannot be summed over the population. Even if there is a common measure—frequency of illness, time lost in waiting in line, busy signals on the telephone—an indiscriminate summation may produce a total of little interest. (Depending on who has it, a case of rubella can be a welcome relief, a nearly unnoticed nuisance, or a horror; hardly anybody cares directly about the total.) But we have already limited ourselves to choices and payoffs that are identical for all; and although this limitation is meant as a starting point of analysis, not as the whole field to be studied, we may as well enjoy the fact that there are such cases, and that, in such cases, there often is some measurable total that is of interest. Even without supposing that my time is as valuable as yours, it can make sense to inquire about the total amount of lost time between us. And often a simple total can be taken to represent an appropriately weighted sum, if there is no expected correlation between the weights one would attach to different individuals and their likely choices of Left and Right.

The dotted lines in Figure 2 show the total values (or average values) corresponding to the numbers choosing Right and Left. At the left end of the scale, everybody is choosing Left, and the total (or average) coincides with the Left curve. On the righthand side, it coincides with the Right curve. Midway between left and right sides, it is midway vertically between the curves, and at the one-third and two-thirds marks it is located at one-third of the vertical distance, or two-thirds of the vertical distance, from L to R.

3. With merely a binary choice, and an unnamed one at that, there is no way to distinguish "positive" from "negative" externalities. We can equally well say that R is an action with positive externalities and that L is an action with negative externalities. To establish a base of reference, we should have to take either L or R as the status quo.

It is good exercise to match pictures like those in Figure 2 with actual situations. Configuration D, for example, has somewhat the shape of an inefficient rationing scheme, perhaps a road-rationing scheme to reduce congestion. Most of the externalities have been achieved when something over half the population participates. (The collective maximum occurs with about three-quarters' participation.) If the scheme is to drive three days a week and take the train the other two, it looks as though a superior scheme would be to drive four days instead; the two-fifths' reduction is too much, the second fifth generates negative net returns.

Configuration B suggests two things. First, the more people join the cooperative coalition, the greater the advantage in staying out: the differential between L and R increases as the number choosing R increases. (In configuration C, the differential diminishes, and the inducement or penalty required to keep people in the coalition or to induce them to join gets smaller.) Second, the collective maximum in Configuration B occurs with some choosing Left rather than Right; not so in Configuration C, in which the dominated Right choice enjoys the externality more than the Left choice.

The variety, though not endless, is pretty great. Case B, for example, can be drawn so that the collective maximum occurs either to the right or to the left of k.

THE SIGNIFICANT PARAMETERS

It was remarked that, in the *description* of a uniform MPD, a crucial parameter is k, the minimum size of a viable coalition. "Viable" means here that, on an either-or basis, assuming that nobody else cooperates, some group of cooperators can benefit from choosing the Right strategy if their number is up to k. This is the minimum-sized coalition that makes sense all by itself. Evidently it takes more than one parameter to describe one of these situations: Figure 2 suggests how much these situations can differ even if k is held constant. But, staying with k for the moment, we might ask whether we should not focus on k/n, or for that matter, n–k.

If n is given, they all come to the same thing. But n can vary from situation to situation, or it may be a variable in a given situation. (It may even be a function of the values of L and R: if L is to fish without limit, and R is to abide by the rules, the number of people who fish at all may depend on the yields within and outside the rationing scheme.) So the question whether k, k/n, or n–k is the controlling parameter is not a matter of definition. It depends on what the situation is.

If k is the number of whaling vessels that abide by an international ration on the capture of whales, the crucial thing will probably not be the absolute value of k but of n–k. If enough people whale indiscriminately, there is no number of restrained whalers who will be better off by restraining themselves. If there is an infinitely elastic supply of cars for the turnpike, no matter how many among us restrict our driving, we will not reduce congestion. And so forth.

On the other hand, if the whalers want a lighthouse and the problem is to cover its cost, we need only a coalition big enough to spread the cost thin enough to make the lighthouse jointly beneficial to those among us who pay our shares. If the value of the lighthouse to each of us is independent of how many benefit, k among us can break even by sharing the cost no matter how many free riders enjoy the light we finance for them.

These are fairly extreme cases. In one, k is independent of n, and, in the other, n–k is what matters. Special cases could be even more extreme. If the danger of collision increases with n, the light will be more valuable with larger n, and k could actually diminish. On the other hand, if more than 40 vessels clog the harbor, and among 100 shipowners some fraction agrees to operate only one-third of the time, 90 participants operating 30 vessels at a time can hold the total down to 40, making it all worthwhile; but among 120 owners, all would have to participate or the number would go above 40 and spoil the result.

So the derivative of k with respect to n can be negative or greater than 1. But, ordinarily, it might have a value in the range from 0 to 1. And if it is proportions that matter–the fraction of vessels carrying some emergency equipment, perhaps–the derivative will approximate the fraction k/n.

So we have a second characteristic of the uniform MPD: the way that k varies with n.

A third characteristic is what happens to the differential payoff as between Left and Right. Does the incentive to choose Left–to stay out of the coalition–increase or decrease with the size of the coalition? For a given n, the value of staying outside the rationing scheme may increase with the number of cooperators: the more the rest of you restrict your whaling, the more whales I catch by staying outside the scheme if entry is limited and if I am already in the business. Alternatively, if joining the coalition merely means paying my pro rata share of the lighthouse, it becomes cheaper to join if more have joined already.

We can measure this by the proportionate change in the payoff difference—in the vertical distance between our two curves—with the number who choose Right. In Figure 2, some of the curves opened toward the right, showing an increasing differential, and some tapered, with diminishing differential.

There is a fourth important parameter if we treat these payoffs as additive numbers, as we might if they have a "productive" interpretation. It is the number choosing Right that maximizes the total payoff or the total output. If the rationing scheme is too strict and the number of whalers is fixed, whalers may collectively get more whales or make more profit if some of them choose Left, catching all the whales they can catch.

The optimum number of individuals to be vaccinated against smallpox will likely be lower than the entire population; the risk of infection is proportionate to the number vaccinated, while the epidemiological benefits taper off before 100%. This is analogous to the two-person case in which both are better off if coordinated mixed strategies (or alternating asymmetrical choices) can be agreed on than if both choose dominated strategies.

In some cases, collective maximization ought to occur when all choose Right if the terms of the coalition have been properly set. It would be silly to have a limit of one deer per season if the rangers then had to go out and hunt down the excess deer. It makes more sense to set the limit so that deer hunters are best off when all abide by the law rather than relying on some free riders to cull the herd. But sometimes the thing cannot be arranged; it may be hard to devise a scheme that allows everybody one and one-third deer per season.

A conflict of interest intervenes if all the benefits of incompleteness accrue to the free riders who choose Left. Consider vaccination: if people can be vaccinated once only, and nobody can be nine-tenths vaccinated, there has to be a system to determine who gets vaccinated if the optimal number is 90% of the population. (Actually, people can be "fractionally" vaccinated, through longer intervals between revaccinations with some attendant lapse of immunity.) With turnpikes and deer hunters, one can search for a quantitative readjustment that makes maximum membership and optimum benefits coincide, even if people have to be allowed four deer every three years to take care of the fractions.[4]

4. According to *Changing Times* (1972) there has not been a confirmed case of smallpox in the United States since 1949, and it is rapidly disappearing in the rest of the world. "Paradoxically, complications from the vaccine cause six to eight American deaths a year, and nearly one in every thousand vaccinations produces mild allergic reactions, such as rash." The Public Health Service no longer requires travelers

There can be a somewhat greater conflict if the collective maximum occurs to the left of k. Unless the distributive problem can be solved, the achievement of a collective maximum then entails net losses, not merely lesser gains, for those who choose Right. If choosing right is voluntary, all-or-none, and noncompensable, any "viable" coalition has to be inefficiently large.

A final point worth noticing is that a coalition–even, or especially, an involuntary coercive coalition–can change payoffs by its mere existence. In a recent article on high school proms, the author described the reaction, when she tried to make tuxedos optional, of "the boys who wouldn't, on their own, go out and rent a tux, but who like the idea of being forced to wear one . . . For many this would be the only time they'd have an excuse to dress up." Remember Bobby Hull's diagnosis of the aversion to helmets: vanity. A voluntary helmet may be seen as cowardly, but nobody thinks a baseball player timid when he dons the batting helmet without which the league will not let him bat. Motorcycle helmets are not only worn regularly, but probably worn more gladly, in states that require them. Whenever ascribed motives matter, the way a choice is organized or constrained will itself be a part of the "outcome" and affect the payoffs. I shall continue to assume, in this paper, that payoffs depend only on the *choices* made and not on the *way* the choices are brought about, but the reader is now alerted to alternative possibilities.

Coalitions

I have used "coalition" to mean those who are induced to subscribe to the dominated choice.[5] They may do it through enforceable contract, by someone's coercing them, or by a golden rule.

---

entering the United States to show vaccination certificates, nor does it recommend routine vaccination of American youngsters. Because immunity wanes, many adults who were once vaccinated may be unprotected now.

Suppose the Public Health Service announced that, considering together the disease and its contagion and the hazards of vaccination, optimally, the U.S. population should be two-thirds vaccinated. What do you elect for your children? (Suppose it simultaneously mentions that, if two-thirds of the population are vaccinated, it is better to be unvaccinated.)

5. In some tautological sense, the choice was not "dominated" if–all things considered–people actually chose it. But the "all things considered" then includes some things of a different character from the things that were represented in the payoff curves.

But the word coalition often has a tighter institutional definition. It is a subset of the population that has enough structure to arrive at a collective decision for its members, or for some among them, or for all of them with some probability, in this particular binary choice. They can be members of a union or a trade association or a faculty or a gun club or a veterans' organization, who elect to act as a unit in a political campaign, in abiding by some rule, in making a contribution, or in joining some larger confederation. And this can take either of two forms, disciplining *individual* choices of the members or making a *collective* choice on behalf of them.

This kind of coalition is often important because it already exists. It has a membership, a decision rule, and a way of exacting loyalty or enforcing discipline. But unless it was formed especially for the purpose of the binary choice at hand, it is probably not unique. There may be many. People who sign up for the blood bank are an ad hoc coalition, in the looser sense I used earlier; but an American Legion post can decide to support the blood bank and get its members to participate, and a labor union, a student organization, and the members of a bowling team can do the same. Thus, there can be several coalitions. Even if there is just one preexisting coalition, in one of our binary-choice situations, there are then three kinds of individuals: those who belong to it; those who do not but who participate in the Right decision and thus form a second, informal grouping; and those who choose dominant strategies. (It is possible that those who stay out are conscious of belonging to a noncooperating or dissident group, and constitute a third coalition making a collective choice.) Now we have a new set of questions.

SUCCESSIVE COALITIONS

Suppose that k or more elect the Right choice. Looking now at the remaining individuals, n–k in number, are they still in MPD? Originally they were, when they were part of the larger population. They may or may not be now. If at k on the upper (L) curve, we draw a horizontal line, its righthand extremity may be above or below the right extremity of the lower curve. If it is above it, the situation no longer corresponds to MPD for these n–k remaining individuals. Figure 3 illustrates the two possibilities.

This condition determines whether, once the first coalition is committed, any or all of the remaining population could be *induced* to do likewise. The first coalition could still coerce some or all of the remainder

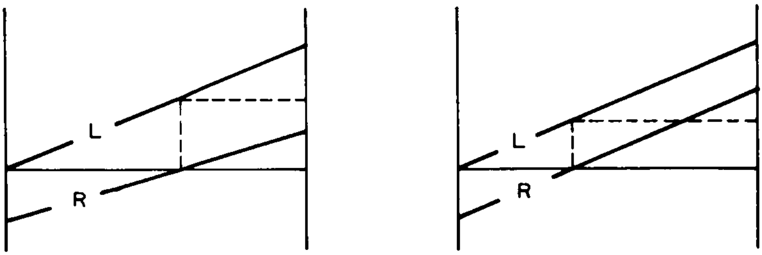

Figure 3.

by threatening to disband and choose Left; but as long as it is committed to choosing Right, the shapes of the curves determine whether or not the outsiders are still in MPD and could benefit from a Right-choosing coalition of their own.

(If they are, and if it takes k′ among them to be viable, we can go on and see whether the n–k–k′ remaining are still capable of another viable coalition, and so forth. We can look at how many coalitions there can be, and whether successive sizes–k, k′ and so on–are increasing or decreasing.)

What is the largest coalition that can choose Right and preserve MPD for the remainder? In a limiting case, the two curves coincide at n, and the situation for any remainder is always MPD; in another limiting case n is infinite, the lower curve asymptotically approaches the upper curve, and the situation again remains MPD for those outside existing coalitions. Otherwise, if n is finite and the two curves do not coincide at their righthand points, there is an upper limit on the number of Right-choosing individuals who can leave a remainder that is itself in MPD.

An interesting consequence is that a Right-choosing coalition can be "too large." It has to be as large as k to be viable; if it exceeds k by too much, it leaves a remainder that has no inducement to coalesce and to join in choosing Right. (We now see why it was not a good idea to include, in the definition of MPD, the condition that, if any subset of the population made the dominated choice, the situation was still MPD for the remainder.)

A TWO-COALITIONS GAME

Next, suppose that there are two coalitions that together exhaust the population. (If they do not, but if the rest of the population is incapable of disciplined organization, the interesting choices relate solely to these

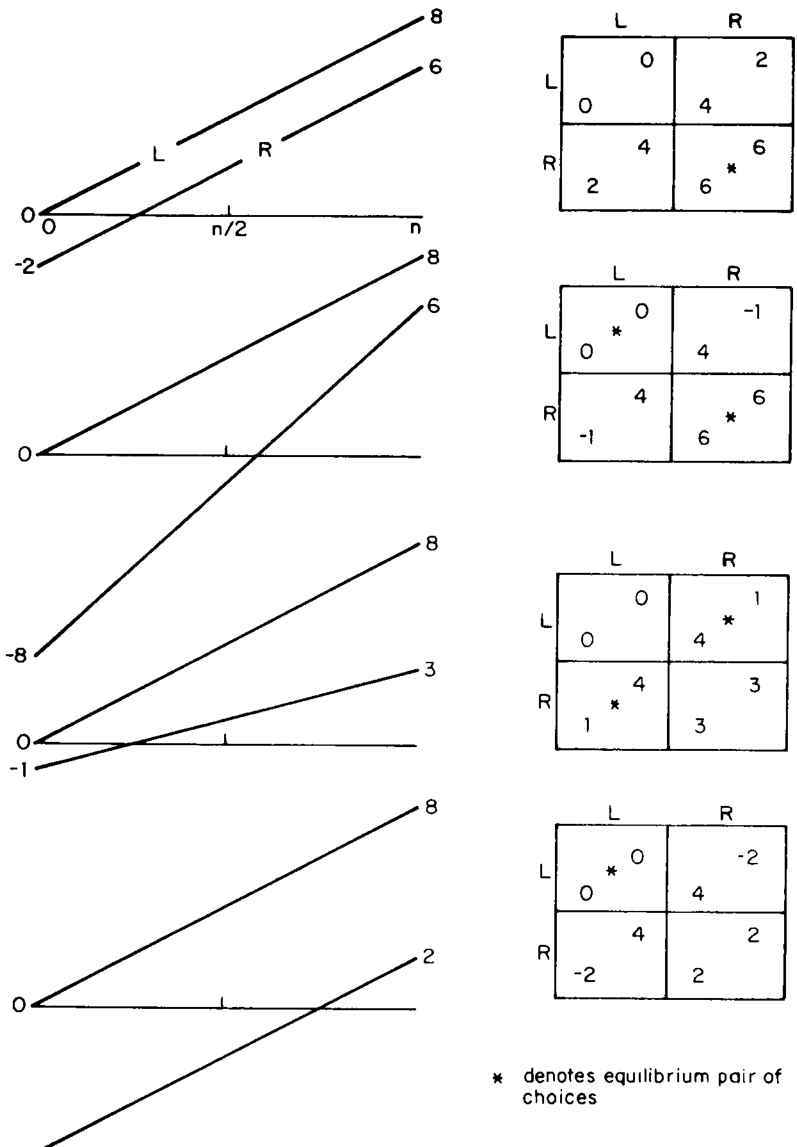
8
6
L
R
0
0
n/2
n
-2
8
6
0
-8
8
3
0
-1
8
0
2
-6
L
R
L
0
0
2
4
R
4
2
6
6
*
L
R
L
0
*
0
-1
4
R
4
-1
6
*
6
L
R
L
0
0
1
*
4
R
4
*
1
3
3
L
R
L
*
0
0
-2
4
R
4
-2
2
2
* denotes equilibrium pair of choices

Figure 4.

two coalitions, and we can move the righthand extremity leftward, reducing n to the sum of the two coalitions. As long as the two coalitions can take for granted that individuals not belonging to either coalition will choose Left, those people can be left out of the analysis and the diagram truncated.)

First, consider two coalitions of equal size. What strategic relation obtains between them? Here are two organizations capable of acting on behalf of their membership or of disciplining their members' choices. The MPD has become a two-organization game. Is this game also prisoner's dilemma? If not, what else can it be?

It turns out that there are four possibilities if each coalition acts as a bloc. One is that each coalition has a dominant interest in choosing *Right.* A second is that each coalition prefers the *same* choice as the other makes, whichever choice that is; there are two equilibria, the Right common choice being jointly preferred. A third possibility is that each prefers to choose *opposite* to the other; the one choosing Left is then the better off. And the fourth is a prisoner's dilemma: *Left* dominates.

Figure 4 shows the four payoff matrices, together with curves that generate them. Thus the uniform MPD can be converted to a symmetrical 2 x 2 game by supposing two coalitions of equal size, each deciding on behalf of its membership. And the ensuing 2 x 2 game may or may not have the payoff structure or prisoner's dilemma, there being three other matrices that can result.[6]

The most curious case is the third. It suggests an asymmetrical outcome: choosing opposite to each other, both are better off than with Left choices, and both are in equilibrium. The collective maximum may not occur with all choosing Right. And it is the only case in which a coalition that can *split* its choice—some choosing Left, some choosing Right—has an incentive to do so. What we then have, rather than just choices of Left and Right, is a "reaction function" relating the *proportions* in which a coalition will allocate Right and Left choices among its members according to the proportions in which the other coalition chooses Right and Left.

With the actual numerical values shown in Figure 4, the payoff-maximizing proportions choosing Right, for each coalition as a function of how the other chooses, are represented by the intersecting curves in Figure

6. If the two coalitions are not equal in size, the number of distinct payoff matrices is not four but nine, five asymmetrical ones being possible. The four preference systems—Right dominant, same, opposite, Left dominant—can occur with a Right dominant preference for the larger coalition and with a Left dominant preference for the smaller (see Schelling, 1972a: 76-78).

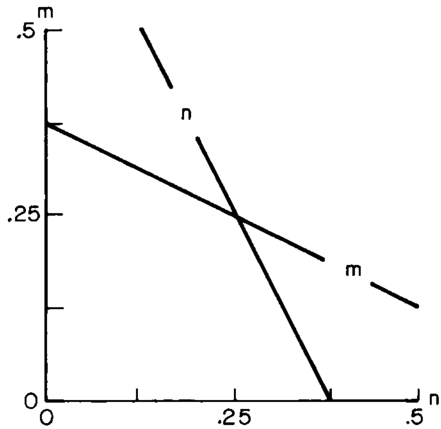

Figure 5.

5. The proportions of the population choosing Right in the two coalitions are m and n. The intersection, at m = n = .25, is an equilibrium point. With m + n = .5 on the horizontal axis in Figure 4, the mean values of the Right and Left choices–the payoffs to the coalitions–are halfway between 1 and 4, or 2.5.

This is evidently an inefficient outcome: Right choices all around yield a payoff of 3. Actually, the collective maximum occurs with seven-eighths choosing Right; with that division, seven-eighths of the population is getting 2.5 and one-eighth is getting 7, for an average of 3.06. (For the mathematics, see Schelling, 1972a: 74.)

Of course, if one coalition reacts in the fashion suggested by the "reaction curve" in Figure 5 and the other knows it, the latter can choose *its* preferred position on the *first's* reaction curve. That is, it can choose its own division between Left and Right that, allowing for the other's reaction, is best. In the case shown, this results in even greater inefficiency: the "anticipating" coalition reduces its own Right vote to induce an increase in the other's; the sum of the changes is negative and the collective total is further reduced.

Some Different Configurations

Thus far, we have examined only a single case, the MPD. We have to look at cases in which the curves cross, with equilibria at their intersection or at their end points and with slopes of the same or opposite direction. We have to look at situations in which people want to do what everybody else does and in which people want to avoid what everybody else does. But rather than switch abruptly, I am going to manipulate our MPD curves, to look at limiting cases and to see what is obtained by shifting or rotating two curves that were initially MPD.

Before doing that, let us remind ourselves of why the prisoner's dilemma gets the attention it does. Its fascination is that it generates an "inefficient equilibrium." There is a single way that everybody can act so that, given what everybody else is doing, everybody is doing what is in his own best interest, yet all chould be better off if all made opposite choices. This calls for some effort at social organization, some way to collectivize the choice or to arrive at an enforceable agreement or otherwise to restructure the incentives so that people will do the opposite of what they would do alone.

For some people, the situation is a "paradox": what is "best" for each person separately is not best for all acting together. Paradox or not, the situation can provoke a search for some kind of organization that can shift incentives, collectivize or surrender choice, or facilitate contingent choices, so that people will stop neglecting the externalities that accompany their choices.

But when the number of people is large, the prisoner's dilemma is not special in that respect. We can draw a number of R-choice and L-choice curves that generate inefficient equilibria and that do not have the shapes, slopes, or end-point configurations of MPD.

Furthermore, the "loose" definition of the MPD allows the possibility that if everybody makes the "right" choice, the result is still not optimal: the collective total would be greatest if a few chose Left. Whatever we call that case–giving it a name of its own or considering it a subdivision of MPD–it is *like* MPD in that there are dominant choices leading to an inefficient outcome, and all could be better off together choosing the opposite. It differs in that everybody would be better off still, if it could be arranged to have something less than everybody make that Right choice, so long as a way could be found to let everybody share in the larger collective total. In the demands it makes on social organization, this is a harder requirement than merely "solving" the problem posed by ordinary MPD and getting everybody to choose Right on condition that everybody else does. In addition to the need to know *how many* should *not* choose Right, there is a need to decide *who* chooses Right and who does not, and perhaps a way to redistribute the results so that, in

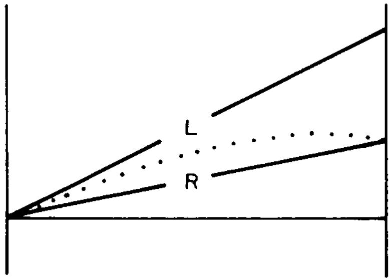

Figure 6.

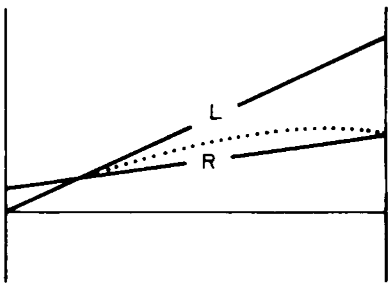

Figure 7.

retrospect or in prospect, the Left-choosers who gain do not detract from the Right-choosers.[7]

It is worth noticing that the number making the Right choice that maximizes the collective total can actually be smaller than the minimum required to form a "viable" coalition: it may be less than k. This entails organizational difficulties; in a refined classification scheme, the situation

7. We can distinguish at least three possibilities here. (1) The choice could be probabilistic: if the weighted-average value is greater with 90% choosing Right than with 100%, people might elect a uniform 10% chance of choosing Left rather than all choosing Right. This would be a "concerted," or "coordinated," or "disciplined" *mixed strategy* in the sense in which that term is used in game theory. (2) If the curves refer to a continual or repeated process, and if the cumulative value for an individual is an average or total computed from those two curves, people can take turns choosing Left one-tenth of the time. (3) If there is an adequate way to transfer value from the Left-choosers to the Right-choosers, the Right-choosers can share in the larger total through compensation. Compensation will be least ambiguous if the L and R curves denote some uniform commodity, activity, or currency that can be directly shared.

might deserve a name of its own. In the absence of compensation, it entails not merely *unequal benefits* from collective action but *actual losses* for some people, for the greater benefit of others, as compared with the equilibrium at Left.

So we should probably identify as the generic problem not the inefficient equilibrium of "prisoner's dilemma" or some further reduced subclass, but *all* the situations in which equilibria achieved by unconcerted or undisciplined action are inefficient—the situations in which everybody could be made better off or the collective total made larger by concerted, disciplined, organized, regulated, or centralized decisions.[8]

There can then be a major division between

(1) the improved set of choices that is self-enforcing once arrived at or once agreed on or once confidently expected—the situation in which people prefer one of two quite different equilibria but may become trapped at the less attractive of the two; and

(2) those situations, including MPD but not only MPD, which require coercion, enforceable contract, centralization of choice, or some way to make everybody's choice conditional on everybody else's.

The MPD then becomes a special, but not very special, subclass of those that require enforcement of a nonequilibrium choice.

INTERSECTING CURVES

To fit MPD into this larger classification, look at the limiting case of the two curves coinciding at the left, Figure 6.[9] Nothing discontinuously different happens here. (The dotted vertical lines denote collective maxima in this and succeeding figures.)

So shift the lower Right curve up a little farther, as in Figure 7. It crosses what used to be the "upper" curve, and Left is no longer dominant. At the left, Right is preferred. If we suppose any kind of damped adjustment, we have a *stable equilibrium* at the intersection.

Because both curves slope up to the right—uniform externality, positive

8. It should be kept in mind that, for people in an MPD or like situation, organizing a disciplined choice is *their* problem, not necessarily ours. "They" can be racketeers enforcing a code of silence, bigots organizing a boycott, conspirators organizing a monopoly, or political opponents forming a caucus against us.

9. For present purposes, do not worry about whether, indifferent between L and R, everybody might happen to choose R rather than L. Everybody may just happen to go to the movies the same night. This detail of analysis is interesting but no more pertinent than if the curves crossed at any other point.

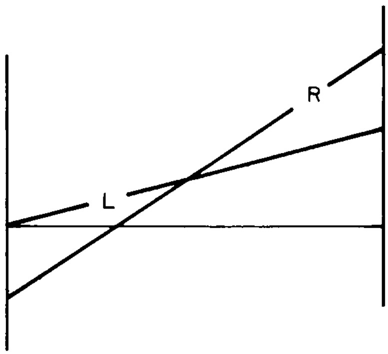

Figure 8.

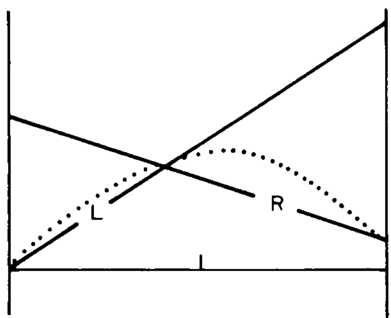

Figure 9.

to the Right–the equilibrium cannot be at a collective maximum. Everybody gains if some choosing Left will choose Right. Those already choosing Right travel upward on *their* curve; those continuing to choose Left travel upward on *their* curve; and all who switch from Left to Right arrive at a higher point on the Right curve than where they were at "equilibrium." (The collective maximum can still occur short of the right extremity.)

Does this differ much from MPD? Both situations contain equilibira that are collectively inferior to *any* greater number choosing Right.

It differs. At the intersection, it takes only a couple of people choosing Right to constitute a "viable coalition," benefiting from their choice of Right. But their action can be offset by the defection of people who were already choosing Right.

What distinguishes MPD is simply that, at the equilibrium, *nobody* is choosing Right; in the intersecting case, with both curves rising to the right, *somebody* is choosing Right. But the difference is not much. In both cases, the equilibrium is inefficient. In both cases, all are better off choosing Right than congregating at the equilibrium. In both cases, the collective maximum can involve fewer than the whole population choosing Right.[10]

While dealing with intersecting lines that slope to the right, we may as well characterize them in the language developed earlier. There is still a *dominant externality;* the internality is no longer dominant but *contingent.*

There is another distinction. The *Left* choice is preferred at the *right* and the *Right* choice at the *left.* Keeping both curves sloping up to the right and intersecting, we could have the two curves interchanged: a Right choice preferred at the right, and a Left choice at the left (see Figure 8).

There we have two equilibria, an all-Right choice and an all-Left. The Right one, enjoying the externality, is preferred. Still, if everybody chooses Left, nobody is motivated to choose otherwise unless enough others do so to get over the hump and beyond the intersection.

So our classification has to consider not only the dominance or contingency of the externality and of the internality, but whether or not the externality favors *more* the choice that *yields* the externality. That is, with a Right choice yielding the positive externality, does it yield a greater externality to a Right choice or to the Left? Which curve is steeper?[11]

10. There is no need for everybody to have a tow cable in his car trunk. It takes two cars to do any good, and two cables are usually no better than one. The "carry" curve should be nearly horizontal; the "don't carry" curve could begin far beneath it, curve over and cross it and become substantially parallel toward the right extremity, at a vertical distance denoting the cost of the cable. The intersection would denote an equilibrium if people could respond to an observed frequency of cables in the car population. Because the "carry" curve is horizontal, the equilibrium is just as good as if everybody bought and carried a cable, and no better; the collectively "best" position would entail a greater frequency of cables, but short of 100%. (And the difference it makes is less than the cost of a cable.) Because of the curvature, a shortfall of cables below the equilibrium value could be severe; an excess above the equilibrium value will benefit some and harm no one. (Most people probably react to a small biased sample of observations; and many may not be mindful that there is such a choice until, in trouble, it is their turn to draw a sample!)

11. In Figure 8, L can stand for carrying a visible weapon, R for going unarmed. I may prefer to be armed if everybody else is but not if the rest are not. (What about nuclear weapons, if the "individuals" are nations?) The *visibility* of weapons can have two effects. If L and R are as in Figure 8, you do not know where you are on your curve–whichever curve it is–if personal weapons are concealed or if nuclear weapons

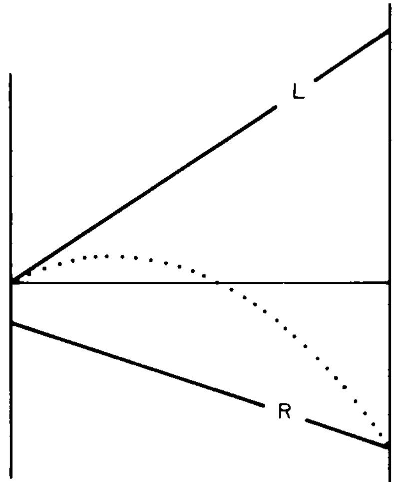

Figure 10.

CONTINGENT EXTERNALITY

Rotate the Right curve clockwise until it slopes downward with an intersection, as in Figure 9. Schematically this is different. *The externality is no longer uniform.* A Right choice benefits those who choose left, a Left choice those who choose right. We have both a *contingent internality* and a *contingent externality.* But we still have an equilibrium. And it is still inefficient (except in a limiting case). (The dotted line again denotes the collective maximum if payoffs are in some commensurable commodity.)

There is a difference. If the collective maximum occurs to the right of the intersection, it is necessarily a maximum in which some—those who choose Right—are *not* as well off as at the equilibrium, unless compensation occurs or choices go in rotation or a lottery determines who chooses Right and who Left. This poses a special organizational problem. But so did MPD when the collective maximum occurred to the left of k. If a system of compensation, of rotation, or of probabilistic determination of who chooses Right or Left is available, the situation is not altogether different from MPD.[12]

---

are clandestine. More likely, visibility will change the payoffs—the risks of being armed depend on whether one is visibly armed—and the curves may have the shape of MPD. (Reliable weapons checks could help, even if the weapons themselves could not be prohibited.)

12. Figure 9 yields some insight into the role of information. For concreteness,

Now, keeping the Right curve sloping downward to the right, but modestly so, displace it downward so that it lies entirely below the Left curve (Figure 10). There is now a *dominant internality* as in MPD. The externality is contingent: a choice of Right benefits those who choose Left while a choice of Left benefits those who choose Right. The situation is unlike MPD because *no* coalition of Right-choosers can be viable (in the absence of compensation). Still, the Left equilibrium can be inefficient. If the Right curve is only slightly below the Left curve at the left extremity, the collective maximum can occur, as it does in Figure 10, with some choosing Right. We still have the organizational problem of inducing the Right choices that maximize the collective outcome.

THE COMMONS

This situation has a familiar interpretation. It is the problem of "the commons." (For the classic lecture, see Hardin, 1968.) There are two common grazing grounds, and everybody is free to graze his cattle on either one. Alternatively, there are two highways, and anybody may drive on either. Anyone who drives on Highway 2 benefits everybody who drives on Highway 1, by reducing congestion there, but adds congestion to Highway 2. Anyone who grazes his cattle on common-pasture 2 adds congestion there, but reduces it on 1, compared with grazing his cattle there. Any problem of congestion with two alternate localities yields the

---

suppose that, during some highway emergency, there are two routes that drivers are not familiar with. If, in their ignorance, they distribute themselves at random between the two routes, with anything like a fifty-fifty division, they will be to the right of the intersection of the two curves in Figure 9. Those who chose R would regret it if they knew; but the outcome is collectively better than an equilibrated division would have been, and, as a "fair bet," all drivers may prefer it to a uniform outcome at the intersection. That being so, the traffic helicopter should keep its mouth shut; it risks diverting just enough traffic to the less congested route to make both routes equally unattractive. (If we had drawn the R curve horizontally, the result would be more striking.) Does the traffic helicopter improve things by telling all those drivers on the congested main routes about the less congested alternate routes?

Next, let R be staying home and L using the car right after a blizzard. The radio announcer gives dire warnings and urges everybody to stay home. Many do, and those who drive are pleasantly surprised by how empty the roads are; if the others had known, they would surely have driven. If they had, they would all be at the lower left extremity of the L curve. An exaggerated warning can inhibit numbers and may lead to a more nearly optimal result than a "true" (i.e., a self-confirming) warning, unless people learn to discount the warning (or subscribe to a service that keeps them currently informed, so that they all go to the intersection of the two curves).

situation represented by two curves that slope in opposite directions. Unless intersecting curves meet special conditions, the collective maximum will not coincide with the equilibrium. And *nonintersecting* curves of opposite shape can yield the same situation!

The fact that the curves do not intersect hardly seems crucial. If the curves did intersect, the problem would be to induce some number greater than the equilibrium number to choose Right, and to share with them the benefits that their so choosing generates for the collective total. But it does not matter much whether the intersection occurs somewhere between the two extremities, at the left extremity, or nowhere. Either way the collective total is maximized with some organized departure from equilibrium and with some choosing in such a way that, without redistribution or sharing, they would suffer net losses.[13]

Dual Equilibria

Turn to the cases of dual equilibria (for straight lines) or multiple equilibria in general.

We have two situations. The curves can have opposite slopes with the *Right* sloping up to the *right* and the *Left* sloping up to the *left,* so that the externality is *contingent* and "self-favoring"—a Right choice favoring a Right choice and a Left choice favoring Left. Or both curves can slope up to the right, the Right curve steeper than the Left. (They can both slope up to the left, of course, but that's the same thing with Right and Left interchanged.) In a classification scheme, these two differ from each other: in one, the externality is dominant; in the other, it is contingent. In social organization, it may not matter whether the curves slope the same or in opposite directions. Either way, there are two equilibria, one at each extremity. The problem of organization is to achieve the superior equilibrium. If both slope in the same direction, there is no ambiguity about which equilibrium is superior; if they have opposite slopes, either may be the superior one.

13. One particular relationship can occur that is worth noticing. With straight lines, it occurs if the two curves are parallel and the right extremity of the lower matches the left extremity of the upper. This is the *zero-sum* situation. The collective total is independent of how many choose Right or Left. It is a limiting case of MPD. If the lower cruve *crosses* the horizontal axis, we have MPD; if it never reaches the axis (and the curves are parallel straight lines), the efficient point is the left extremity. If it reaches the axis just at the right extremity, the collective total, or weighted average, is constant. (Whatever the shape of the upper curve, we can always draw a unique zero-sum lower curve.)

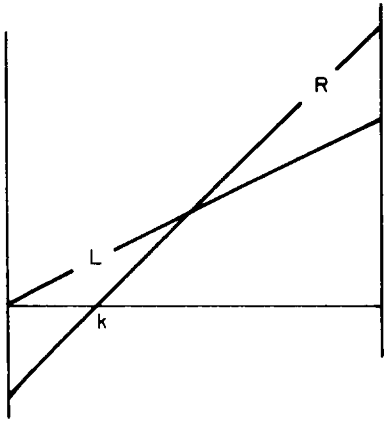

Figure 11.

In any of these cases with two or more equilibria, the problem (if there is a problem) is to get a concerted choice, or switch, of enough people to reach the superior equilibrium. There may be no need for coercion, discipline, or centralized choice; it may be enough merely to get people to make the right choice in the first place. If the choice is once-for-all, it is enough to get everybody to expect everybody else to make the right choice, and this expectation may be achieved merely by communication, since nobody has any reason not to make the right choice once there is concerted recognition.

If an inefficient Left choice has become established, no individual will choose Right unless he expects others to do so; this condition will require some organized switch, as in one-way streets or driving to left or right. People can get trapped at an inefficient equilibrium, everyone waiting for the others to switch, nobody willing to be the first unless he has confidence that enough others will switch to make it worthwhile.

Notice now a difference between the curves' both sloping up to the right and their having slopes of opposite sign. In the former, a coalition can occur that is insufficient to induce the remainder to choose Right, yet is viable. Figure 11 illustrates it. If everybody is choosing Left, there is some number, call it k again, that will be better off choosing Right, even though they are too few to make Right the preferred choice for everybody else. The critical number occurs where the Right curve achieves the elevation of the left extremity of the Left curve, just as in MPD. A Right-choosing coalition is viable if it exceeds this number; if it achieves

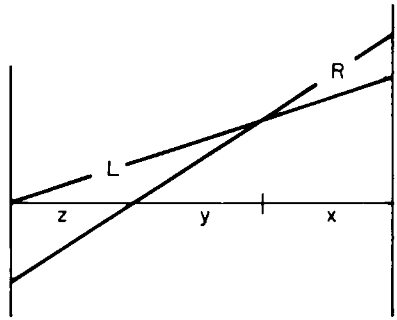

Figure 12.

the larger number corresponding to the intersection, it can induce everybody else to shift. But even it it is too small to accomplish that, the coalition can still benefit. Thus there is an element of MPD even in the situation of two equilibria: there is some coalition that is better off choosing Right, even though the remainder are better off still, and even though any member of the coalition would be better off if he could defect and choose Left. The difference in this case is that there is a still larger coalition that can *induce* everybody else to switch, because it is big enough to make a Right choice the preferred choice. (With MPD a second organized coalition might be so induced, but not the members individually.)

MPD AS A TRUNCATED DUAL EQUILIBRIUM

We can now take a final step in denying MPD any special status, especially any status based on its quantitative structure. The difference between MPD and the dual equilibria need be no more than a difference in size of population. In Figure 12, with a population of x, there are two equilibria. If k is independent of the population—if the curves are anchored on the left—reduce the population to y and MPD results. Reduce it to z and MPD disappears. The MPD is merely a "truncated dual equilibrium," without enough people to carry themselves over the hump. (And the dual equilibrium is merely an "extended MPD," with enough people added to make the coalition self-sustaining.)

This does not mean that every MPD can acquire a second (efficient) equilibrium by enlargement of the population. As remarked earlier, the way k varies with n will be crucial; marginal externalities need not be constant; and parallel or divergent straight lines would not cross to the right anyway. But *any* dual equilibrium that is anchored on the left—one that is a matter of numbers, not of proportions—will truncate to MPD.

Curvatures

There is no end to the shapes we could give our Left-Right curves. But also there is no guarantee that a pair of real choices exists somewhere that corresponds to some pair of curves that we might adopt on heuristic or architectural grounds. Straight lines are somewhat noncommittal and can often serve as proxies for whole genera of monotonic curves. But they are also somewhat prejudicial in their simplicity: they are poor at representing asymptotic behavior; they can intersect only once; and they never reach maxima or minima. A few examples with curvature may dispel the presumption that externalities ought to display constant marginal effect.

COMPATIBILITY

One interesting class may be U-shaped for both curves, like the three variants in Figure 13. The basic relation is one of "compatibility." Uniform choices for *all others* are better for anyone than any mixture. whichever way the one person makes his own choice.

At the top of Figure 13, a Right choice is favored if enough choose Right and a Left choice if enough choose Left. There are two equilibria. One is superior, but either is far better than a wide range of intermediate distributions. A possible interpretation is daylight saving. Let it be summer and let R represent daylight saving. The best is with everybody on daylight saving. Things are not bad if everybody is on standard time. Things are bad if people are divided in the way they keep office hours, schedule deliveries, programs, and dinner engagements. Furthermore, unlike driving on the right or using metric screw threads, the worst thing for an individual is not to be out of step with everybody else; it is to have everybody else not in step with each other. Even if I am on daylight saving, I can better navigate my daily life with everybody else on standard time than if half the world joins me in daylight saving and I never know which half. A traveler who crosses time zones may keep his wristwatch on "home time" and get along all right unless he is with other travelers of whom some do the same.

The middle case is similar overall. But this time everybody somewhat prefers to be in the minority while mainly preferring uniformity for

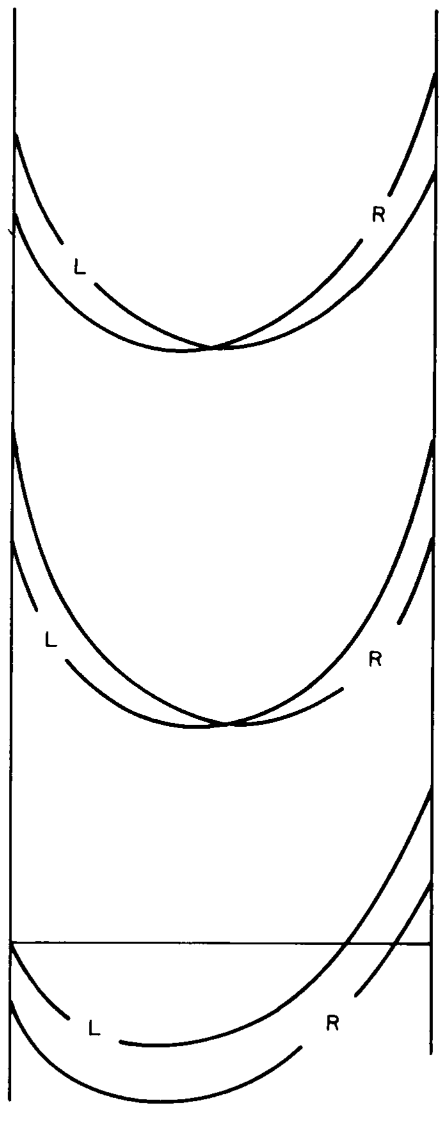

Figure 13.

everybody else. Possibly, to find a parallel with daylight saving, this could be a choice of Monday or Friday as the third day of the weekend when the four-day work week becomes common. To avoid crowds, one may prefer to have Friday off if everybody else drives out of town or goes to the golf links on Monday. (Or, if it is storekeeprs, everybody prefers to be open for business the day his competitors are closed.) At the same time, in getting up a golf game or going to the beach with friends, or just knowing what stores are open and who is keeping office hours, there is advantage in the rest of the world's uniformity; and, on balance, it is better to be in line with everybody else if one cannot enjoy exclusivity. I leave it to the reader to find a more plausible interpretation. In any event, if an equilibrium can be reached, it is an unsatisfactory equilibrium. The temptation to be different stirs things up to everybody's disadvantage, and the advantage in being different dissolves only when there is too little homogeneity to make it worthwhile.

The case at the bottom shows a dominant internality and a single equilibrium, comparatively satisfactory but not completely so. (It could have been drawn with the Left extremity higher than the Right and an efficient outcome. To illustrate a problem, I have drawn it contrary.) I will take a flyer: Left is the decimal system, Right the duodecimal. Either works fine, but if half of us are on one and half on the other, the result is confusion. Furthermore, it is just hard enough to convert to a duodecimal system that, though on behalf of posterity I wish everybody else would change, in my lifetime I would rather stick to my own system, even if it means I am out of step. Another example would be the choice by a group of ethnically similar immigrants to continue using their native tongue or to adopt the language of the host country. As in MPD, I may be willing to adopt the duodecimal system as part of a bargain I strike with everybody else. And, indeed, if we compare end points and ignore the middle range this is MPD, isn't it? We can even identify the parameter, k, denoting the minimum size of viable coalition to switch to the duodecimal system or to the host-country language.

COMPLEMENTARITY

Now invert the curves, as in Figure 14. Here again, there are at least three species. This time, instead of compatibility, we have *complementarity.* Things are better if people distribute themselves between the choices. But though everyone prefers that the universe be mixed in *its* choice, he himself may prefer to be in the majority, may prefer to be in

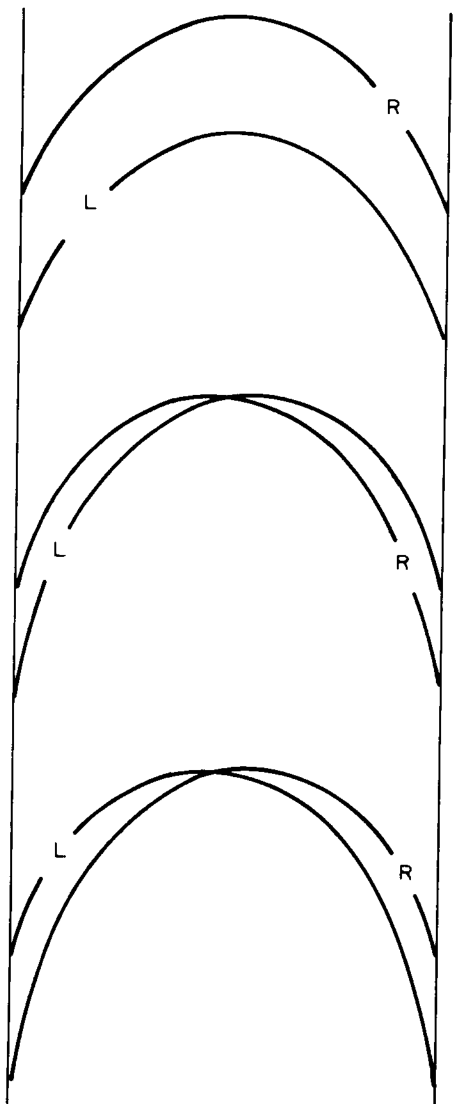

Figure 14.

the minority, or may have a dominant preference no matter how the others distribute themselves.

An obvious binary division with complementarity is sex. Let us conjecture, along lines of biomedical hints that have recently been publicized, that it becomes possible to choose in advance the sex of one's child. (The choice is not binary, since most parents have more than one child and can choose among a few integer mixtures for each family size. But this whole analysis is suggestive and exploratory; so pretend that a family commits itself to boys or to girls.)

It is easy to suppose that most prefer the population to be mixed, and probably close to fifty-fifty. But a parent couple could plausibly have any one of three preferences.

First, there might be a uniform dominant preference, everybody wanting a girl or everybody wanting a boy independently of the sex ratio in the population, while badly wanting that population ratio close to fifty-fifty. Second, everybody might prefer to have a child of the scarcer sex: for dating, marriage, and remarriage, a child of the scarcer sex might be advantaged. Third, the dominant sex might have a majority advantage outweighing "scarcity value," and parents might deplore a preponderance of males or females while electing a child of the preponderant sex.

In one case, there is a happy equilibrium. In one case, there are two unhappy equilibria. And in one case, there is a single unhappy one.

In the unhappy case at the top, we can identify k, the minimum coalition that gains from enforceable contract. (A coalition larger than half the population has to allocate Right and Left choices among its members.)

This is evidently not MPD by the earlier definition; and we cannot make it so by truncating the diagram, because k is, in this case, a constant fraction of the population. Yet if MPD includes cases in which a coalition, beyond some size, maximizes the collective total among its members by allocating some choices Left, the shapes of the curves to the right of that collective-maximum point are inconsequential unless they generate a new collective maximum. They will not if they diverge much; and they surely will not if they slope downward. For some purposes, then, the upper diagram in Figure 14 shares the interesting properties of MPD. (Coalition policy, though different in detail, is similarly interesting in the bottom diagram.)

The real problem, if technology should offer the choice and thus create the problem, is attenuated by the nonbinary character of the choice for couples that end up with more than one child. But even the artificial binary illustration is a vivid reminder that a good organizational remedy

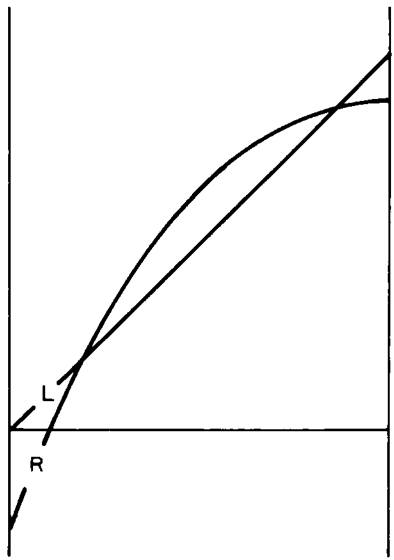

Figure 15.

for severely nonoptimal individual choices is simply *not to have the choice*—to be victims (beneficiaries) of randomization—and thus to need no organization!

SUFFICIENCY

Turn now to Figure 15. A Right curve cuts a Left straight line twice. Everybody prefers that everybody else choose Right, and over an intermediate range people are induced to choose Right. An example might be the use of insecticides locally: you benefit from the use of insecticides by others; the value of your own insecticides is dissipated unless some neighbors use insecticides, too; with moderate usage by others, it becomes cost-effective to apply your own; and, finally, if nearly everybody uses insecticides, there are not enough bugs to warrant spending your own money.

Communication systems sometimes have that property. If hardly anyone has citizen's-band radio, there is nobody to talk to; the externality benefits more the people who have sets than the people who do not, though the latter get some benefits from the communication system; if enough people have sets, others are induced to procure them as a nearly universal means of communication; finally, if everybody else has a set, you can save yourself the expense by dropping in on a friend and using his equipment or handing an emergency message to any passerby, who will transmit it for you.

A more familiar example is the committee meeting. Everybody suffers if nobody goes; it is not worth going unless there is likely to be a quorum; over some numerical range, one's presence makes enough difference to make attendance worthwhile; and if the meeting is large enough, there is no need to give up the afternoon just to attend.

With these payoff curves, there are two equilibria, one at the upper-right intersection and one at the left extremity. If we relabel the curves–and change the interpretation–the equilibria are at the lower-left intersection and the right extremity.

Graduated Preferences

We have assumed indentical payoffs for all. If we relax that assumption, we are in trouble unless we preserve some regularity. If everyone among the n individuals has his own pair of arbitrarily shaped curves, we shall be hard put to identify the incentives of any subset because their preferences will depend on just which people they are.

IDENTICAL EXTERNALITIES

One possibility is to suppose the externalities the same for all but the internalities different. All can then have identical Left curves. Their Right curves will be similar, but displaced vertically from each other by the differences in their internalities. Parallel straight lines are the neatest illustration. We can draw Right curves for the twentieth, fortieth, sixtieth, eightieth and one-hundredth percentiles among the population, with the individuals ranked in order of increasing internalities. For an "MPD-like" situation, we have in Figure 16 a common Left curve with some Right curves beneath it, here drawn parallel to the Left curve as well as parallel to each other. (In Figure 18, the Right curves are parallel to each other but not to the Left curve.) The dashed line connects points on the five Right curves where the number choosing Right, as measured from left to right, matches those percentiles. This curve is labeled MIRV, the "Marginal Individual Right Value." (Economists will recognize it as similar to a "value-of-marginal-product" curve in contrast to a curve denoting "mar-

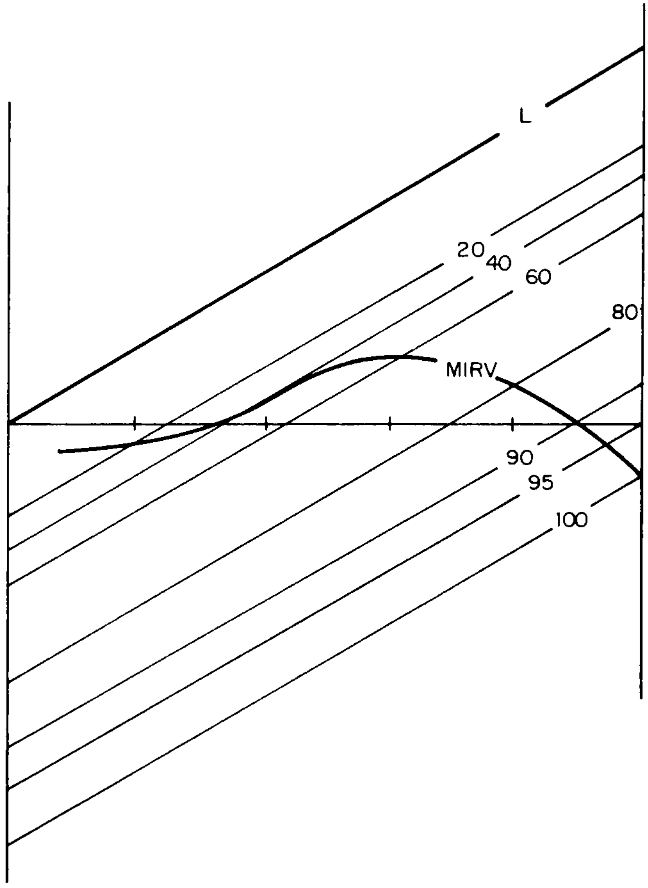

Figure 16.

ginal-revenue-product.") We can interpret this MIRV curve to denote the value of a Right choice for the marginal individual, when individuals are ordered from left to right in terms of increasing internalities.

The right end-point of this curved line corresponds to the right-extremity value for the individual who has the largest internality. If the dashed MIRV line rises to the right, crosses the axis, and stays above it, we have something very much like MPD. (Though all the Right curves slope upward, even if they all cross the axis out to the eightieth or ninetieth percentile or so, the MIRV line need not rise at any point above the axis drawn from the left extremity of the Left curve.)

The particular MIRV curve shown in Figure 16 is a kind of "partial MPD." There can be a viable coalition of any size in the range where the MIRV line is above the axis. About thirty percent of the population, if it is the thirty percent with the smallest (negative) internalities, can form a viable coalition. As much as ninety percent of the population, in Figure 16, could constitute a viable coalition. But there is a residual ten percent among those with the largest internalities that cannot be interested.[14]

IDENTICAL INTERNALITIES

An alternative possibility is that the internalities are the same for all, but the externalities differ. (There is some ambiguity here: we can have the Right curves coincide at the left extremity, at the right extremity, or anywhere between, fanning out from that point with their different slopes.) An example is in Figure 17, where the internalities are uniform for a universal choice of Left, but the externalities differ. Here, up to about the median individual, the externalities sufficiently favor a Right choice that if the entire population had the same Right curve, the upper-right extremity would be a stable equilibrium. And for some number between the twentieth and fortieth percentiles, the individuals who enjoy the strongest externalities from the Right choices of others constitute a stable equilibrium set. If as many as twenty percent choose Right, some number less than forty percent but greater than twenty percent find Right the preferred choice; if they so choose, some larger number find Right the preferred choice, and as more choose Right, more find it the best choice, up to the fortieth percentile. At that point, if a few more choose Right, they would be individuals whose Right curves, *for that number of individuals choosing Right,* were below the Left curve.

Out to something over eighty percent of the population, the externalities are great enough to have the MPD configuration. And, in Figure 17, by the time we reach the one-hundreth percentile, these last few

14. If the payoffs are commensurable—if they have an interpretation as some fungible "output" like an agricutural crop—it is a matter of fact, not a choice of presentation only, that L is uniform and the R curves differ. If the payoffs are personal and incommensurable we can, for parallel curves (constant differentials in internality), just as well align the R curves and let the L curves be displaced vertically by the differentials in the internalities. Essentially, we fix the end-points of L or R as "zero point" and let the other curve reflect the difference. The slope and position of the MIRV curve is the same. (For this special case of parallel straight lines, the collective maximum occurs at the percentile whose R curve meets the right end of the axis drawn from the left end of the Left curve.)

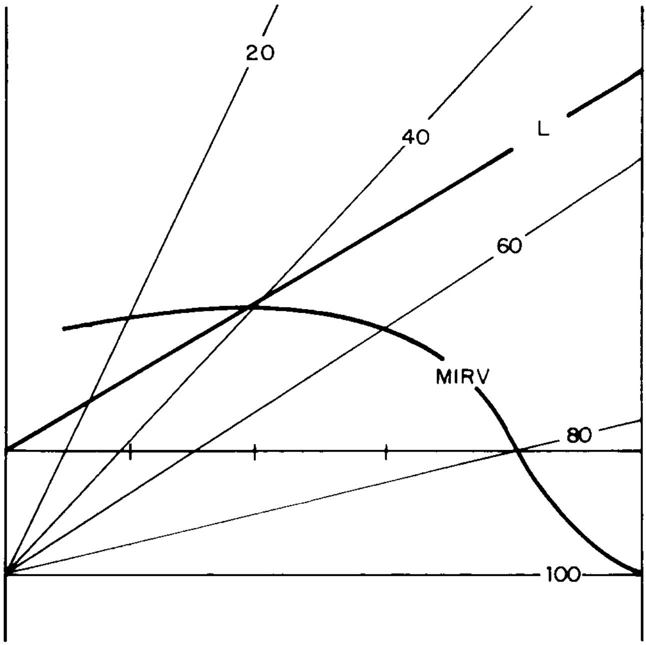

Figure 17.

individuals enjoy no externality whatever; their Right curve is at a fixed elevation all the way.

In addition to a stable *equilibrium* at forty percent choosing Right, we have a viable *coalition* of up to eighty percent. And, with the curves drawn in Figure 17, the collective maximum evidently occurs with all choosing Right, even though the twenty percent of the population least sensitive to the externality suffer net losses from joining unless they can share in the increment to the total that their joining up creates. (That the collective maximum, for commensurable payoffs, occurs at the right extremity is "evident" only because of the steepness of the R curves for the earlier percentiles and is not a necessary result for any such curves.)

For analyzing stable and unstable equilibria, what is crucial is the relation of the MIRV line to the Left curve. Absolute payoffs depend on the vertical distances between those two curves, but the *kinds* of equilibria that occur depend only on whether the MIRV line is above or below, and where it crosses. For this purpose, it is sufficient to know, for each individual, just where his own Right curve is above, and where it is below, his own Left curve. If, as in Figures 16 and 17, everybody's Right curve slopes upward and either cuts his Left curve from below or stays everywhere below his Left curve, we need only to know for each individual whether a crossover occurs and for what aggregate number choosing Right it does so. From this we can derive a cumulative frequency distribution showing, for any percentage of the population that might choose Right, the percentage of the population for which a Right choice would be preferred.

Or, what is the same thing with axes interchanged, we can derive a frequency distribution showing, for any of the least demanding among the population–those whose crossover points occur nearest the left extremity–the minimum number of the population that, choosing Right, would induce this number to choose Right.

The *central* portion of Figure 18 displays exactly that. Any point on the dashed curve indicates, for the number measured horizontally choosing Right, the number measured vertically that would prefer a Right choice. Alternatively, for any point on that dashed curve, the number measured vertically would prefer a Right choice if and only if the number choosing Right were at least as great as the horizontal value. We can call the dashed curve the IRC curve, the "Induced Right Choice."

In the range over which the dashed IRC curve is above the 45-degree line, the percentage preferring a Right choice is greater than the number making the Right choice, and more would be induced to choose Right, up to the point where the IRC curve cuts the 45-degree line to go beneath it. To the right of that intersection near the upper right, say at the ninetieth percentile, there are not ninety percent among the population who would prefer a Right choice if ninety percent were choosing Right; if that many were choosing Right, some among them would prefer to switch to a choice of Left. As they do, the number choosing Right decreases and the number preferring Right decreases, and the switching process continues down to that intersection.

The lower-left region, where the IRC curve lies below the 45-degree line, is a region in which the number preferring Right is everywhere less than such a number choosing Right; some will switch to a Left choice if they are free to do so, and in doing so induce still others to do so, and unless the IRC curve crosses the 45-degree line and lies above it down in the far lower-left corner, zero will be the equilibrium number choosing Right.

This IRC curve, representing the cumulative distribution of crossover points, contains less information than the curve directly above it in the

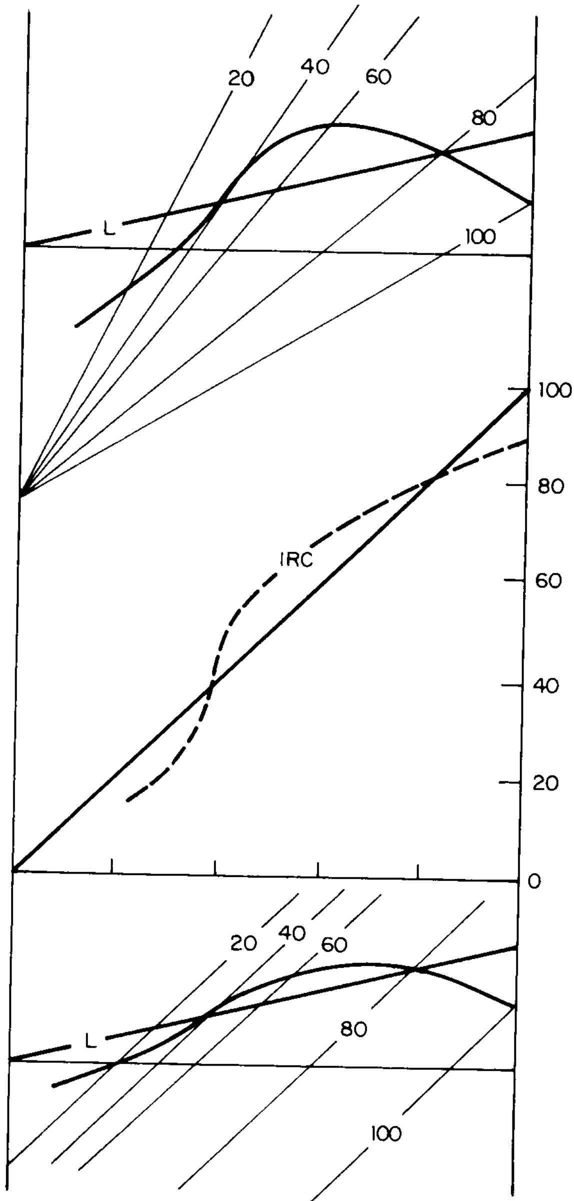

Figure 18.

diagram, from which it has been derived. For each individual, it gives us an algebraic sign, not a number. For the population as a whole, it can display potential equilibria, but not collective totals. But although we have lost information–*because* we have lost information–that curve *exists*. For any (least demanding) fraction of the population measured on the horizontal axis, there is some fraction of the population (possibly none, possibly all) for which the Right curve is above the Left curve. So there exists a unique, single-valued IRC function of the kind plotted in the middle of Figure 18.

If for every individual there is some minimum number that has to choose Right to induce him to do likewise, but no maximum–if his Right curve crosses the Left curve once, from below–the IRC curve in Figure 18 will rise monotonically to the right. It will furthermore denote, at any point, a *particular* group of individuals, a group that includes all individuals represented by points to the left. Nobody drops out as the fraction rises. People are uniquely ordered according to the value on the horizontal axis at which their own preferences become Right. If, instead, everybody's Right curve cuts his Left curve from above, the IRC curve in Figure 18 will be monotonic downward and will represent a depletion of a fixed population as we move to the right. That is, everybody is in rank order, and the fraction denoted by the height of the curve at any point includes all those included at points farther to the right.

Finally, if the Right curves of some individuals cut their Left curves more than once, or if for some there is a single crossover point to the left of which the Right curve is higher while for others it is to the right that the Right curve is higher, the IRC curve need not be monotonic, and, whether it is or not, it will represent a subset of the population that shows "turnover" as we traverse the diagram from left to right. That is, some who "join" the Right-preferred group at a certain point on the horizontal axis disaffiliate at some larger fraction, some who "leave" may rejoin, some are in up to a point, others are out up to a point. The number measured vertically at any point on the IRC curve represents a particular subset of the population, but, in this complicated case, it need not contain all subsets represented to the right.

Because we have lost information in producing this cumulative frequency distribution–the IRC curve–we can reconstruct several alternative arrays of Right curves that could have produced it. In Figure 18, the top part of the diagram shows, for five quintiles, a fixed internality and graduated externalities. The bottom part of the diagram shows a fixed externality with graduated internality. Individuals represented in the lower-left reaches of the frequency distribution can be people either very *sensitive to the externality* or comparatively *insensitive to the internality;* those represented in the upper-right reaches of the diagram can be those least sensitive to the externality or having the larger internality.[15]

More Than Two Choices

We have considered only two choices. To what extent does this analysis generalize to three or more choices?

Some symmetrical cases generalize easily. When two straight lines have *opposite* slopes, whichever way they slope ("self-favoring" externalities or "other-favoring"), the analysis fits three or more choices perfectly well. Generally speaking, if Left and Right curves are similar when referred to their own axes–Left curve plotted against Left choices, Right curve against Right choices–there is nothing especially binary about the analysis.

Consider oppositely sloping straight lines. There are two possibilities: R slopes up to the right and L downward, or R slopes down to the right and L upward. If an action yields negative externalities toward those who choose *that* action, we have the ordinary case of congested highways and the number of highways can be two, three, or a hundred. If choosing an action benefits those who choose the *same* action–people on the metric system benefit from its use by others–the analysis applies equally to two, three, four, or any number of metrics, languages, keyboards, or signaling systems.

The analysis is peculiarly binary when a given choice has positive or negative externalities for everybody, whichever way they choose. (MPD is binary, not symmetrical.) The curves are not reflections of each other. We may be able to use a somewhat similar analysis for a threefold or fourfold choice; but we cannot simply generalize from the binary analysis.

This asymmetrical case is rich in possibilities. As a bare suggestion of the variety obtainable when the action yields uniform externalities–not positive or negative externalities solely to those choosing the same action–consider a threefold choice among Left, Right(1), and Right(2). Let the two Right choices *produce* the same externalities additively but not *benefit* equally from the externalities they produce. The externality reflects the *sum* of R(1) and R(2). Consider it positive.

15. The use of a cumulative distribution of crossover points, like that in the center part of Figure 18, for two groups comprising a fixed population is illustrated in Schelling (1972b). Similar curves relating to variable populations, for two interacting groups, are extensively used in Schelling (1971).

We plot on our left-right scale the sum of R(1) and R(2). We draw three curves—the payoffs to a choice of L, of R(1), and of R(2). This is a special case, and it may be hard to think of an interpretation, but it does yield interesting possibilities.

Look at the top diagram in Figure 19. In the absence of R(1), we would have two curves with two equilibria, the inefficient one on the left. R(1) gets us over the hump. To the left, R(1) dominates L; nobody will choose R(2) unless nearly everybody is choosing one or the other variant of R, but the dominance of R(1) assures that enough will choose R(1) for R(2) to take over, and the right-hand equilibrium results. Thus R(1), never itself an equilibrium outcome or part of one—always a bridesmaid, never a bride—mediates between the other two choices. It can pull the population from a Left extreme equilibrium to the collective maximum at R(2) for everybody.

In the center diagram, it does not quite perform that whole function, not dominating at the left extremity. But it does permit a small coalition to get away from L—here drawn horizontal, for variety—causing others to choose R(1) until R(2) becomes the preferred choice. R(1) mediates over an important range, though not solving the whole problem as in the top diagram.

In those two cases, R(1) yields to R(2) but, in R(2)'s absence, would offer superior equilibria to what L alone offers at the left. In the bottom frame, it plays a more paradoxical role. Alone with L, R(1) offers a highly stable *inferior* equilibrium at the right. No choice of R(1), by any number, benefits those so choosing or those who stay with L. (Indeed, reading from right to left, R(1) is almost the upper curve of an MPD combination with L, offering a second equilibrium at L that is stable over only a small range.) Thus R(1) is an "option" that the population is better off without, in the presence of L alone. But see what it adds to the situation when L and R(2) are the choices. The Left equilibrium is much inferior to the R(2) equilibrium, but stable over a wide range. R(1), which alone with L could only worsen things, nearly dominates L and makes even a small concerted (or unconcerted) choice of *either* R(1) or R(2) sufficient to bring about the R(2) equilibrium at upper right. The dynamics are the same as in the center diagram, but in the bottom case an otherwise wholly unattractive option serves as a self-effacing usher for R(2).

I leave it to the reader to invent interpretations of R(1) and R(2). This trinary choice offers a richer menu, and not merely a generalization of results from the binary situation.

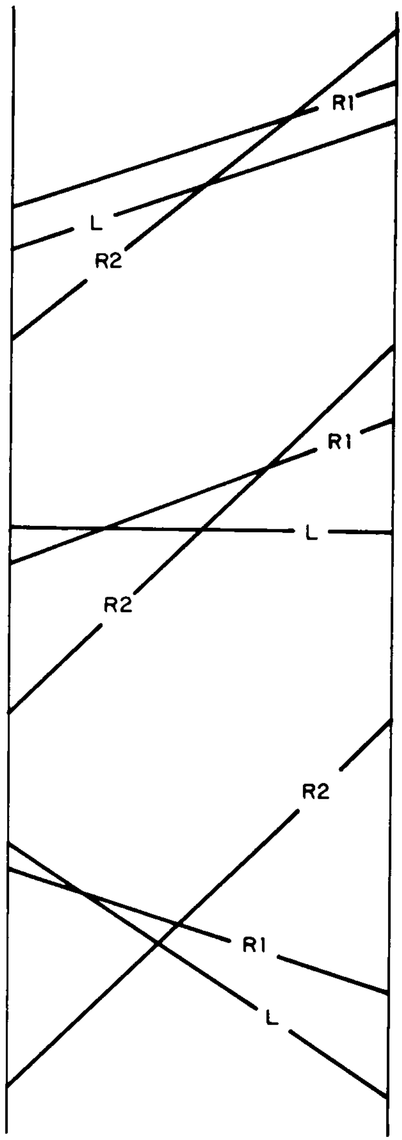

Figure 19.

A Schematic Summary

It is tempting to work out an exhaustive schematic classification for the various possible binary-choice payoff configurations. But the possibilities, though not endless, are many. The curves, even if monotonic, can be concave or convex, S-shaped, flanged, or tapered; and, of course, they need not be monotonic. The shapes that are worth distinguishing depend on what we want to single out for analysis—the number of equilibria, the efficiency of equilibria, the role of information or misinformation, the sizes of potential coalitions, the importance of discipline or enforceable contract, the importance of population size and other things. And still we are dealing almost exclusively with uniform payoffs throughout the population (or the very special case of regularly graduated payoff differences). No logical classification scheme is likely, therefore, to serve everybody's purpose.

One way to generate a classification is to do what we did with "prisoner's dilemma" for all symmetrical 2 x 2 matrices. There are 12 different payoff rankings (not counting ties) that yield symmetrical matrices, and we can interpolate straight-line binary-choice curves for each of them. That is, there are 12 different ways that the end points of a pair of straight lines can be ordinarily ranked if there are no ties. But Figure 4 showed that, for some purposes, subcases are worth considering. And among the 12 straight-line pairs suggested by the 12 symmetrical 2 x 2 matrices, some of the differences are of hardly any interest. For what it is worth, the 12 cases are sketched in Figure 20.

In the upper half of that figure, the *internality* is uniform; in the lower half, it is contingent. In the left half, the *externality* is uniform; in the right half, it is contingent. The small circles mark the points that are potential "equilibrium points" in a very simple sense: at such a point, an individual, given the choices of all others, cannot improve on the choice he is making.

The differences among A, B, and E are not consequential, nor are those between I and J or between K and L. The difference is merely in the relative ranking of two end points that are neither equilibria nor preferred outcomes, and whose comparative positions don't matter. Because ties are omitted and only "strongly ordered" payoffs represented, the zero-sum configuration is missing. (It is intermediate between C and D.) Also arbitrarily omitted for the same reason are curves hinged at left or right extremities and curves that are horizontal lines.

Figure 20 is included primarily to save the reader the trouble of producing it for himself, and for reference in the next section. It is merely

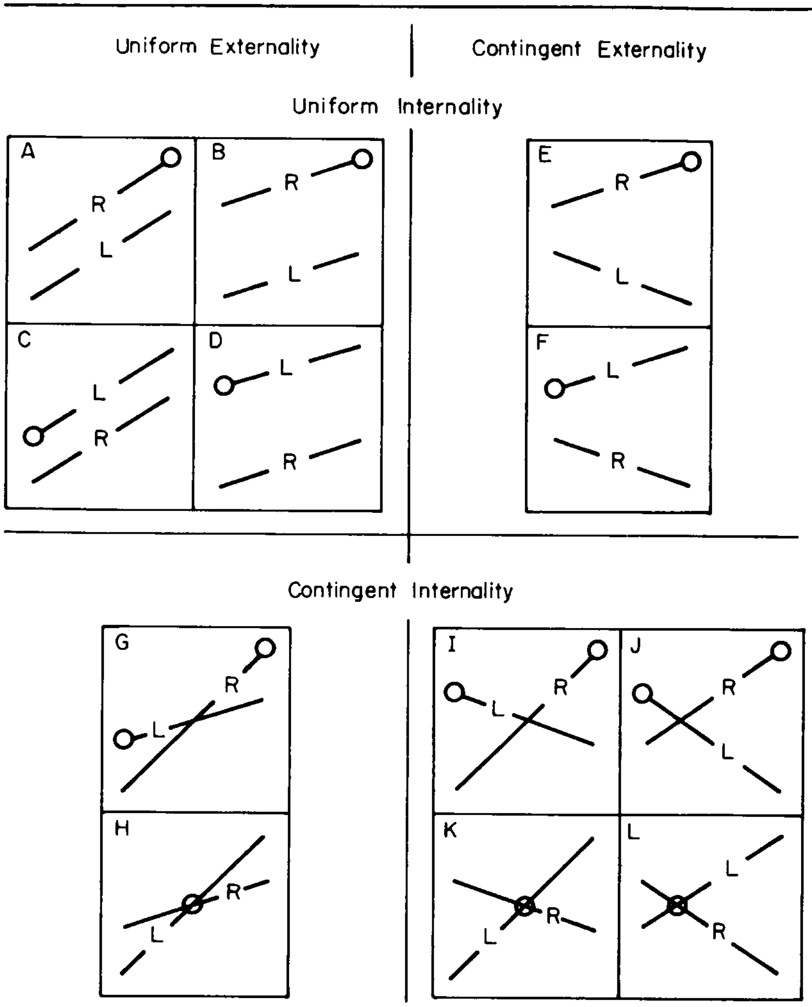

Figure 20.

an answer to the question, what shapes can one get by interpolating linearly among the payoffs of those 12 symmetrical matrices with strongly ordered payoffs (no ties)?

Some of the omitted "limiting cases" may be of greater interest than some cases shown. The problem of "the Commons" will often have a horizontal R curve cutting the L curve or lying just beneath it. (In that case, H, K, and L become identical; C and F become identical if we pivot the R curves about their left end points; and C becomes identical with H, K, and L if the R curve is pivoted around its right end point.) The "special case" of symmetrical I and J configurations (which are then indistinguishable from each other) is really not special, but common: it represents all those situations in which it matters a lot that people follow the same signal, but matters little just which coding they use—e.g., red or green for "go."

Equilibria, Universal Preference, Uniformity, And Collective Maxima

There is a brief, useful classification scheme for straight lines that can be illustrated by Figure 20, especially if we add some of the figures introduced earlier. It distinguishes these situations:

(1) There is a unanimously preferred equilibrium.
  - (a) It is a unique point of equilibrium, as in A, B, and E.
  - (b) It is either of two equivalent equilibrium points, as in I or J, if the upper end points are aligned horizontally.
  - (c) It is one of two equilibrium points that are not equivalent, as in G and, generally, as in I and J.

(2) There is a single equilibrium point and it is "dominated"; i.e., there are other outcomes that would be unanimously preferred to the unique equilibrium point.
  - (a) The collective maximum occurs with the same choice for all, as it may (but need not) occur in C and H.
  - (b) The collective maximum occurs with a mixture of choices and unequal outcomes, as it may in C and H.
    - (i) The collective maximum is unanimously preferred to the equilibrium point, as it is in H and may or may not be in C.
    - (ii) The collective maximum is not unanimously preferred to the equilibrium point, as it may not be in C.[16]

16. Let L rise from 0 to a and R from -1 to $(b-1)$. If X is the fraction choosing Right, $L = aX$ and $R = bX-1$, and the collective total is $X(bX-1)+(1-X)aX$. Since $b > 1$ in the case being considered, the maximum is to the right of $X = 0$; in fact, it is to the right of $X = .5$. It occurs to the left of where the R curve crosses the axis—i.e., with R negative, if $b < 2a/(a+3)$. Note that, for this to occur, $b < 2$ and $a > 3$. The collective maximum occurs with $X < 1$ and $R > 0$—i.e., between where R crosses the axis and the right extremity, if $2a/(a+3) < b < (a+1)/2$. It occurs at the right extremity when $b \geqslant (a+1)/2$—i.e., $a \leqslant 2b-1$. Condition 2–b–ii, with the collective maximum occurring while R is below the axis drawn from the left extremity of the L curve, is much less restrictive if we drop the restriction to straight lines.

(3) The equilibrium point is neither dominated nor unanimously preferred: there are alternative outcomes, involving a mixture of choices, that are better for those making the one choice but not for those making the other.
   (a) The equilibrium point is at the collective maximum, as it will be in D and F if the L curve rises, from left to right, by less than the R curve lies beneath it at the left (and in K and L in the special case of horizontally aligned upper end points—in which case, incidentally, K and L are indistinguishable).
   (b) The equilibrium point is not at the collective maximum, as generally in K and L, and as in D and F when the L curve rises, from left to right, by more than the R curve lies beneath it at the left.

With curvature, of course, there may be two dominated equilibria, as in Figures 14 and 15. In the zero-sum case (a boundary case between C and D), there is no point of collective maximum. With more than two choices, there may be more than two equivalent "universally preferred" outcomes. And so forth.

In every case, the term "equilibrium," or "equilibrium point," should be qualified to read "potential equilibrium." The order and timing of choices and the reversibility of choices; information about others' choices; signaling, bargaining and organizing processes; custom, precedent, and imitation; and many other crucial elements have been left unspecified. So we have no assurance that actual choices would converge stably on what we have identified as "potential equilibrium points."

For that reason, this is not a classification of binary-choice *situations*, which may differ as importantly in those other characteristics as in their payoffs, but refers only to the *shapes* of the binary-choice outcome curves.

REFERENCES

Changing Times (1972) "An end to smallpox vaccination?" Volume 26 (March): 32.

HARDIN, G. (1968) "The tragedy of the commons." Science 162 (December 13): 1243-1248.

HARRIS, R. (1970) "Annals of politics." New Yorker (December 5): 60-161; (December 12): 53-131.

Newsweek (1969) October 6.

SCHELLING, T. (1972a) "Hockey helmets, concealed weapons and daylight saving: a study of binary choices with externalities." Harvard University John F. Kennedy School of Government Public Policy Program Discussion Paper 9.

--- (1972b) "The process of residential segregation: nieghborhood tipping," pp. 157-184 in A. Pascal (ed.) Racial Discrimination in Economic Life. Lexington Mass.: Heath Lexington.

--- (1971) "Dynamic models of segregation." J. of Mathematical Sociology 1: 143-186.