Volume Title: Issues in Defense Economics

Volume Author/Editor: Roland N. McKean, editor

Volume Publisher: NBER

Volume ISBN: 0-87014-490-1

Volume URL: http://www.nber.org/books/mcke67-1

Publication Date: 1967

Chapter Title: The Strategy of Inflicting Costs

Chapter Author: T. C. Schelling

Chapter URL: http://www.nber.org/chapters/c5161

Chapter pages in book: (p. 105 - 127)

The Strategy of Inflicting Costs

T. C. SCHELLING

HARVARD UNIVERSITY

This paper was stimulated by a remark of Frederic M. Scherer on the notion "that a nation should weigh its own costs against the costs of its enemy" in choosing a strategy or in allocating resources to defense. He said the logic of it escaped him. I doubted that it did, and began to develop a complacent defense of the proposition only to discover that indeed the logic—or at least an important part of the logic—had escaped me too; and where first I thought that economists like Scherer and me should neither disagree nor be in doubt about the proper formulation of so simple a principle, I no longer find the principle so simple.

Scherer argued that the proposition is valid only if the amount of resources a rival can allocate to defense is fixed—"an assumption unsupported by recent history," he says—or if an arms race is being waged as a "game of economic ruin." He doubts the advisability of the latter strategy, at least for the United States against the Soviet Union, and he goes on to criticize a good deal of recent defense analysis for its "zero-sum" orientation, its assumption of strictly opposing interests.[1]

The proposition that the enemy's costs in meeting some threat that we pose, or in responding to some measure we take, should be taken into account in deciding whether the measure is worthwhile, is not hard to defend in a zero-sum context; and one can set Scherer straight by saying that the zero-sum context is all anyone had in mind. But though one can defend the letter of the proposition, it is not so easy to defend the spirit. Should we confine ourselves to the zero-sum context, where the answers are comparatively easy, evading the more interesting contexts where the answers may be harder? Should we defend the cost criterion

[1] See his review of E. S. Quade, "Analysis for Military Decisions," *American Economic Review*, December 1965, pp. 1191–1192.

in a way that increases our vulnerability to Scherer's charge that quantitative military analysis too readily assumes a zero-sum relation, even while our discussions of strategy are almost exclusively nonzero-sum?

We cannot justify zero-sum reasoning by assuming that the amount of resources an adversary will put into the contest is fixed. Even with a fixed budget the contest will almost surely be dramatically nonzero-sum. "Deterrence," for example, is meaningless in a zero-sum context. So is "surrender"; so are most limited-war strategies; and so are notions like "accidental war," escalation, preemptive war, and "brinkmanship." And of course so are nearly all alliance relationships, arms-race phenomena, and arms control. The fact that war hurts—that not all the losses of war are recoverable—makes war itself a dramatically nonzero-sum activity whether budgets are rigidly fixed or not.

To get a zero-sum contest one needs a much more restrictive assumption than the one Scherer mentioned. The idea that we have some interest in an adversary's domestic economy and that that interest is not simply opposite to his interest and measurable in the same currency as our military interests, is only one among many reasons—and surely not the most important reason—why military relations are nonzero-sum both in peace and in war. Furthermore, a "game of economic ruin" is not zero-sum if war itself is an alternative to progressive economic ruin, and if war is worse than economic ruin, or the threat of such ruin can be used to extort concessions.

In fact, it is hard to see how military relations can ever even approximate a zero-sum model unless things have reached the point where all but two extremes among the possible outcomes have become—for reasons of diplomacy, personality, technology and geography, or some profound incompatability between the two sides—practically unattainable and irrelevant to decision. There are, of course, tactical arenas within a nonzero-sum relation that can be dealt with, at least in a practical sense, as separable zero-sum local contests, but care is needed even at the tactical level.

Separable Zero-Sum Components

There are indeed zero-sum military problems that can be separately analyzed. Cryptography is an example: one scrambles or enciphers tactical communications to keep secrets. The avoidance of predictable regu-

larities is another: one tries not to be predictably unprepared on Sunday mornings. One randomizes the firing of hardened missiles so that they will not be simultaneously and predictably vulnerable during the moment the lids are off the silos before the missiles are on their way. Search and evasion will be randomized; depth and proximity fuses will be set on zero-sum principles; and scarce ammunition may be allocated among sites on strictly zero-sum criteria. These schedules and allocations relate usually to some fixed set of resources in circumstances where there is no thought of *influencing* enemy intention, even though the wider context may be a nonzero-sum campaign or bargaining process.

There is another way that a zero-sum situation can be embedded in a larger nonzero-sum context and yet be susceptible of isolated zero-sum analysis. This is if the payoffs are some kind of bargaining power, usable in some subsequent process of negotiation or coercion. A zero-sum contest can be fought to determine which side gets some advantage —some ammunition, some intelligence, some technology, or some part of the world that is significant for its later strategic role in some confrontation or bargain. It will be hard to classify such a contest as "zero-sum," though, if the two adversaries have different evaluations of the subsequent bargaining process. It will also be hard to classify it zero-sum if they can anticipate the later bargaining and make later behavior conditional on the conduct of the momentary contest. But if the thing fought over is uniquely recognizable as an advantage to the side that gets it and if any nonzero-sum use of the asset comes only later, the momentary contest can be treated as zero-sum.

There is a special case worth mentioning, though it is difficult to give it a contemporary interpretation. That is the case in which both sides agree to a "trial by ordeal"—to play a game of chess for the treasure, or to hold a tournament for a lady's favor. An agreement may be reached to avoid a potentially destructive nonzero-sum contest by letting some game (in the literal sense) determine the outcome. David and Goliath fought, according to Yadin, to make unnecessary a bloody battle between their armies; both sides agreed (though in the end Goliath's did not keep the agreement) to let the battle of champions determine which army would become slaves to the other.[2] Arbitration has this character; a zero-sum battle of wits determines the outcome,

[2] Yigael Yadin, *The Art of Warfare in Biblical Lands,* New York, 1963, Vol. II, pp. 267–269.

both sides being committed to abide by the result to avoid the potentially more destructive bargaining that would ensue if they were unable to wager their positions on some conventional procedure.

Tactics: Zero- or Nonzero-Sum

It is often alleged that strategy is nonzero-sum while tactics tend to be zero-sum. There is a good deal to the notion; it often means merely that a mission, once ordered accomplished, should be accomplished efficiently. Bombing the civilian population of Japan was not a zero-sum strategy; but once it had been decided to light fires in Tokyo there was every reason to do it with a minimum of risk and effort, and each individual sortie could be scheduled from a zero-sum point of view.

On the other hand, a decision to take prisoners, to bomb civilians in order to destroy a military installation, or to suffer casualties in achieving an objective, are not zero-sum decisions. Furthermore, even though back at headquarters a particular engagement may be viewed as zero-sum, any operation carried out by live human beings is a multiperson, nonzero-sum game. Maybe some kinds of combat have become so mechanized that the individual combatant's valuation of his own life has little relevance to his choice; but the history of warfare tells us that tactical engagements have always been multiperson, nonzero-sum games in which fear disabled more combatants than weapons, and discipline counted more than numbers, weapons, terrain, or technology. General Rommel's orders to his Panzer units to open fire before they were in range, because enemy soldiers ducking their heads were as good as enemy soldiers dead, was a recognition of the nonzero-sum character of local combat.

Costs and "Suboptimization"

Still, the zero-sum contest is an interesting bench mark, a limiting case worth being acquainted with, and often a decent approximation in some localized arena of the larger contest. So it is worth while to get the cost principles straight for that situation, too. The principle says merely that, in taking account of the consequences of some act or decision, any costs that the enemy is obliged to incur have a positive value

to our side, just as the costs we incur in carrying out the act, or in consequence of the decision, must be valued negatively. And a corollary of this principle is that if substantial costs have to be incurred by the enemy, and more modest costs incurred by us, and if as a result of incurring those costs the enemy keeps us from accomplishing anything locally, nevertheless his higher costs are a perfectly valid objective, or motive, for our going ahead with the action. We can even think of the local contest as a battle of costs.

Now, if the whole situation is strictly zero-sum, and if our analysis is intended to comprise the *whole* situation—if we are engaged in what Charles Hitch has called "grand optimization"—there is no need to bring in the concept of "cost," whether the enemy's costs or our own. The best over-all strategy, worked out in all its detail, is just the best strategy, all things considered; and any relevant costs have already been implicitly taken into account. The fact that a particular measure on our part will "cost" an adversary some lives, ammunition, fuel, man-hours or anything else is of no concern if we have made allowance in our over-all model for the withdrawal of lives, fuel, man-hours, and so forth, from their alternative uses. Costs are intended to reflect the payoffs *left out* of our local analysis.

It is when we engage in "suboptimization," and the local contest for which we are suboptimizing uses resources that might have been allocated instead to other arenas, or resources that can be saved and used later in different arenas, that we need to make specific allowance for costs. If some particular measure is under consideration and we have allowed for all of its local tactical consequences but find that it either uses or saves some resources on our side, and that the adversary's accommodation increases or decreases some inputs on his side, the local payoffs do not tell the whole story. The extra resources are withdrawn from other fields of battle, resources saved are available to other fields of battle; it is all the same war, and we have to take account of what happens on those other fields. Costs are simply the measure of consequences, or payoffs, that occur outside the boundaries of our local problem. They are the *external* economies that any partial or local analysis needs to take into account. If there is no economic linkage between our local problem and the rest of the whole problem—if all local resources are specific, if nothing can be added or withdrawn, if nothing can be saved for later use—then the local payoffs tell the whole story,

and we need make no allowance, through "costs," for payoffs beyond our horizon.

Of course, if all resources are so specific that to find out the "worth" of resources added or withdrawn elsewhere we have to examine in detail their impact on those other fields of battle, then we have to solve all the problems at once in order to solve our local problem, and suboptimization is infeasible. Suboptimization only works if there is little or no connection between the local problem and the over-all problem, or if there is some convenient approximate way of making allowance for the impingement of the local problem on the bigger problem. If the enemy has an actual price system that works rationally, or if he manages nevertheless by direct means to allocate his resources efficiently; and if the enemy's economy is characterized by substitution possibilities, alternative uses of outputs, and all the other characteristics we usually impute to a resilient and rational economy; and particularly if the situation allows the enemy time to adapt economically to the measures we confront him with; then we assume that increments of enemy resources can be allowed for as "costs" and "savings" without regard to how the individual items that comprise the resources are allocated elsewhere, or precisely where they are drawn from.

That seems to be all. It could be illustrated by a set of a dozen zero-sum matrices, each dependent on the share of total resources allocated to it, the payoffs to each side being the sum of the payoffs in the individual matrices, with "optimization" achieved only when neither side could improve its position by reallocating resources among the different arenas. A shift in the parameters of one matrix might open opportunities for one side to improve its position by shifting resources into that matrix; if they are shifted from the other eleven matrices in a way that minimizes the reduction of payoffs in those other matrices, there is a calculable "cost" that is of interest to both sides. This "cost" will tend to be the same (for modest increments) whichever the matrix to which resources are being added, and it permits measurement of the impact of local decisions on other fields of battle. If we combine the twelve matrices into a single game we can drop the cost concept altogether; if instead it is convenient to analyze the twelve different subgames separately, the "cost" figures serve an essential purpose and make local optimization consistent with over-all optimization.

The principle does not depend on there being a recognized live, ra-

tional zero-sum adversary. The traditional examples of external economies reflect the same principle. If decisions are decentralized to the managers of twelve fields belonging to the same farm, any managers who put up scarecrows, or plant extra seed, or protect the seeds so that it takes crows longer to get at them, will affect the number of crows on each other's fields. These "adaptive" shifts of the "enemy" have to be taken into account in deciding whether the scarecrow, or the extra seed that draw crows from neighboring fields, is worth the cost.

Costs can of course be used incorrectly, and most of the exposures of the "fallacy" of this cost reasoning relate to some mistaken use. I have been asked whether I would buy a one-hundred-dollar nylon vest to protect myself from a five-cent bullet, as though the cost criterion says I should not—the factor of 2000-to-1 against me making it a bad investment—yet if I treasure my life I shall obviously buy the vest. In this example the cost criterion is merely backwards; the relevant question is whether my adversary should buy a bullet knowing that I can nullify his investment with a bullet-proof vest. He has wasted his money if the vest is cheap, made a splendid investment if my vest is expensive, and if asked what he accomplished by buying his bullet should have the good sense to say that he imposed a cost on me, not that he hoped to kill me and was frustrated.

The "Exchange Rate"

We need of course some kind of "exchange rate" to tell, without extending our analysis into grand optimization, whether a cost imposed on the enemy is worth the cost to us of imposing it. Unless we and our adversary have comparable economies there is no reason why the appropriate exchange rate should equate percentages of gross national product or any other economic magnitudes on the two sides. We can say, as a kind of shorthand, that a cost is worth incurring if it imposes a "greater" cost on the other side; but then we must mean that "costs" have already been measured in the same currency as the payoffs in those other arenas. If we calculate "costs" for the two sides in any other way, such as money costs converted on a foreign currency market, percentages of GNP, or what the enemy's specific inputs would cost at our prices, there is no reason to suppose that a straightforward comparison of "greater" and "smaller" will tell us the story.

The Incidence of Costs on Consumers

Suppose an enemy fuel shortage falls on consumers and the enemy values the health and comfort of his people, deplores the death of infants, suffers in the suffering of his people, and takes those civilian values into account in his own planning, not for the military advantage they give him but just for their own sake. Then the relation between us is nonzero-sum. We may like his babies to be warm and healthy, prefer that they suffer, or not give a thought to it; but we are unlikely to be just as disposed to welcome the suffering of his people as he is to deplore it. In our cost calculations we should take no credit for the prospect that some children will be cold or sick because we damaged enemy fuel supplies.

In fact, the enemy may care less about his own people than we do—if by "enemy" we mean the directorate of the military opposition to us. Suppose the enemy neither likes nor dislikes his subjects—perhaps because he is a military occupant of another country and does not much care whether the citizenry suffers or prospers—and we do not care either, because our sole concern is military advantage. The game is still zero-sum, but if a local fuel shortage falls on the local citizens who will have less heat in their houses, or if a local fuel saving means merely that less is requisitioned locally, there is no "cost" to be allowed for in our problem. The cost is passed along to somebody outside the contest. The costs we want in our analysis are those that reflect the flow of resources among the sectors of the contest, for all the sectors comprising the adversary relation between us; if other sectors are involved and can absorb or yield resources to the sectors engaged in the contest, then either those sectors matter and the situation is nonzero-sum, or they do not matter themselves and distort the price system. We have to correct for them. That makes it harder to ascertain the costs and to use the figures properly, but does not contradict the principle.

The Nonzero-Sum Context

Let me turn now to the nonzero-sum adversary relation that usually characterizes war and military preparations. First let us remind our-

selves of some of the reasons why wars and arms races are nonzero-sum. There is civilian damage—pain, disability, death, privation, lowered living standards, fear and anxiety, both during the war itself and after the war. There are battle casualties; soldiers are people, and when they die there is not only a loss of military assets but also human pain and grief. There is, during military buildup, some influence of that buildup on the likelihood that war will occur; and military preparations affect the character of any war that ensues, including the capabilities of both sides for keeping it limited and for stopping it. Wars and military preparations affect precedents (like the use of gas or nuclear weapons) that affect the conduct of future war. The enemy in a war is often a potential ally or subject. (In colonial wars, the killing of an enemy soldier is the loss of a colonial subject.) War involves hostages; much of the population of western Europe was a potential hostage in German hands during World War II. Adversaries can furthermore gain or lose honor and prestige in a nonzero-sum fashion. War can induce political change; the monarchs of 1914 destroyed most of the system congenial to them. War can weaken both adversaries vis-à-vis other countries. There is a moral cost for some countries in the conduct of war. War is often a multiperson relationship; the generals' plot against Hitler's life and the shift in authority within Japan after Hiroshima remind us that nations are not the same as individuals or teams of like-minded individuals. Finally there is the point mentioned by Scherer, that the military expenditure of potential adversaries can affect the civilian welfare and political conduct of the country, affecting the likelihood of war itself.

We could go on, but there is no need to. Zero-sum adversary relationships are hard to find in any important area of human affairs; and people in need of an example in a hurry sometimes point to war as one of the rare zero-sum social relationships, but this is pure thoughtlessness. About the only thing that could make war appear zero-sum would be a *belief* that war were zero-sum, a belief so obstinately held that the war would indeed be conducted that way. People who put chessmen on the dust jackets of books about strategy are presumably thinking of the intellectual structure of the game, not its payoff structure; and one hopes that it is chess they do not understand, not war, in supposing that a zero-sum parlor game catches the spirit of a nonzero-sum diplomatic phenomenon.

Elasticities of Demand: A Comparative-Advantage Example

Because Scherer and others have already pointed out that costs imposed on the enemy may be passed along to his population, so that we gain less than we hoped or even lose if we prefer his civilians not to bear the cost, let me suppose a fixed enemy military budget and explore some of the ways the cost criterion gets complicated as we acknowledge that our interests and our opponent's are not strictly opposed. Consider that he has a given budget that he can allocate among two, or three, or more kinds of military resources for a war with us. We have to examine how our deployment of forces affects his deployment, and in what ways our choices can influence his choices and how we should evaluate his choices.

Begin with a simple model. He has a budget with which to buy missiles and sites. If he spends it mostly on missiles and concentrates on a small number of sites he gets more missiles than if he went to the added expense of dispersing his missiles among more sites. At the same time, his missiles are more vulnerable if they are concentrated; and for "second-strike" purposes he prefers a dispersed force. His value system, we may suppose, contains two variables: the number of missiles that he can buy (with which he could launch an attack), and the number of missiles he could expect to have (or can threaten to have) left over in case we attack him. We can call these his "first-strike" force and his "second-strike" force. (They are of course the same force, but they are evaluated differently according to who starts the war. A large vulnerable force has a comparative advantage in striking first; a smaller less vulnerable force has a comparative advantage in striking back.)

How he disperses his missile force—how he uses money that might have been spent on missiles to buy more sites—will depend on how expensive the sites are. If sites are cheap he will buy a site for each missile (perhaps more, if we cannot tell empty sites from those that have missiles and must target as many sites as he buys). If sites are expensive he cannot much disperse his force without greatly reducing its size; beyond some point, even for a purely second-strike force, more dispersal becomes uneconomical.

To be explicit, suppose that his costs are a linear function of missiles and sites; missiles are procured at a constant price, sites are procured at

a constant price. We have a given force with which we can attack all of his sites; and, not knowing the particular distribution of missiles among those sites, we would distribute our attacking missiles evenly if we attacked. Suppose our attacking missiles have independent probabilities of kill, so that the survival of a site is merely the survival probability in the face of one attacking missile, compounded by the number of missiles fired at the site. Let X denote the number of missiles he buys, Y the number of sites, and M the size of our attacking force, with P the survival probability of an individual site under attack by a single missile, K his total budget (measured in units equal to the price of a missile), and a the price ratio of sites to missiles. The expected number of missiles surviving our attack would be equal to:

$$XP^{M/Y}$$

Substituting the budget equation, this gives:

$$[K - aY]P^{M/Y}$$

If he wants to maximize the expected value of his second-strike force, he merely maximizes that expression with respect to Y. If he wants to maximize some function of first- and second-strike forces, giving positive value to his first-strike force, we have to specify the function but then we can go ahead and solve his problem.

Now let me pose this question. Suppose we invent a new device, perhaps a new ingredient in the warhead of our attacking missiles, that if unanticipated by the enemy would increase the potency of each of our missiles against any site it is aimed at. It makes all of his sites more vulnerable, and hence his entire missile force. Never mind what it actually is; just imagine that it is available to us at modest cost. Suppose the enemy knows about it and can either see that we have it or assume that we have it. And suppose that this particular capability is one that is readily countered, but at a price. The enemy can defend himself against this measure with complete effectiveness but has to spend money to do so. Do we hope his countermeasure is cheap or expensive?

Now, the distinguishing characteristic of this particular measure is that it attacks *sites,* not missiles. Of course, if it destroys a site it destroys the missiles on the site; that is the whole idea. The point is this: The enemy's countermeasure is associated with his sites, not with his missiles, so that the aggregate cost of defending his force against our

new device depends on how many sites he has to protect, not on how many missiles he has to protect. His countermeasure, we suppose, is adding smoke screens, antiaircraft batteries, police dogs, duplicate generators, or some other active or passive defense to individual sites in such a way that the cost does not depend on how many missiles are on the site.

The point of this distinction, of course, is that it affects his allocation of resources between missiles and sites. That is, it affects the allocation of his budget between first- and second-strike forces. And *we care.* Furthermore, *we care in a way that is not strictly opposed to the way he cares.*

He may, for example, be mainly concerned to deter us with the best second-strike force he can buy and have little interest in either attacking us with his whole force or threatening to attack with his whole force. (He could even prefer to reduce our apprehension, and attach negative value to his first-strike force.) And *we* may prefer that he buy a second-strike rather than a first-strike force, that is, a force whose second-strike capability is relatively larger and whose first-strike capability relatively smaller.

We may not. It depends on whether we want to deter him or to attack him, and it depends on what we think raises or lowers the likelihood of our having to attack him or his having to attack us. However, if we mainly want deterrence, our preferences somewhat coincide with his. He wants a good second-strike force *even* at the expense of his first-strike force; we are willing to see him get a good second-strike force *if* it is at the expense of his first-strike force.

Now, do we want the price of his sites, as distinct from the price of his missiles, to be high or low? If we can raise the price of those sites by threatening them with a measure against which a costly countermeasure is available, do we want to make the threat? If he must assume we have the device, do we want the countermeasure to be expensive or inexpensive? And would we prefer to threaten him with a measure whose defense is associated with his missiles rather than his sites, so that they raise his cost per missile rather than his cost per site?

We can solve the problem explicitly and find out—at least we can if we specify his preference function as among first- and second-strike forces, and our preference function as between his first- and second-strike forces. But offhand about the only conclusion we can jump to

is that we would rather raise the price of his missiles than the price of his sites (if the aggregate costs in both cases are of about the same magnitude) because we anticipate that he will use more of the cheaper of the two inputs, and we know which one it is that we want him to use. Intuition, analogy, or quick mental arithmetic will not give us that answer about site costs. If some neutral country sells the site-defense countermeasure and we can influence its price policy, we do not yet know whether we want it cheap or expensive.

The key to our decision is what, in the familiar language of economics, we might call the enemy's "elasticity of demand" for missile sites. Whatever he spends on sites is unavailable to be spent on missiles. If we want to minimize his first-strike force, we want to minimize the money he spends on missiles, therefore to maximize the money he spends on sites. Will he spend more money on sites, or less, if the price of sites goes up? Is his demand for sites inelastic or elastic? We have to work the problem to find out.

In the specific form in which I set up the problem above, if the enemy's exclusive interest is a second-strike force, it turns out that his demand for sites is slightly inelastic.[3] He spends somewhat more on sites if the price of sites goes up, less on missiles. To put it differently, the "income effect" of the price rise outweighs the "substitution effect" and, though missiles have become cheaper relative to sites, he now buys fewer missiles.

In this particular problem the answer happens to be independent of the initial costs of sites and missiles and independent of the budget; but this is a special case and we would have to give a more qualified answer if we worked with a more complicated model—if, for example, missiles or sites were bought along rising or falling supply curves, if sites supported or interfered with each other, or if missiles did, or if the enemy's preference function were slightly different. It is easy to make a minor change in this model and reach the contrary conclusion—that we do not want sites to become more expensive to the enemy, we want the counter-

[3] $X e^{(M \log P)/Y}$ is maximized at $Y^2 + MY \log P - (KM \log P)/a = 0$. Differentiating this equation yields:

$$-\frac{a/dy}{y/da} = \frac{KM \log P}{aY(2Y + M \log P)} < 1$$

More directly, one can maximize $XP^{aM/(K-X)}$ and differentiate the resulting equation for dX/da. That derivative is negative.

measure sold cheaply; and we should avoid confronting the enemy with our own measure in the first place if we can, to keep him from finding missiles more attractive, relative to sites, than they used to be.[4]

This kind of consideration cannot arise in the zero-sum confrontation. It furthermore can arise, as it did here, without regard to whether we care whether the enemy's people are better housed and fed. It can arise, as it did here, without our putting a positive value on his second-strike force. It requires only that *our* preferences not be *strictly opposed* to those of the enemy, or, in economic terms, that our indifference curves as between his first-strike and his second-strike capabilities not coincide with his in reverse.

A Jointly Valued Capability

Consider another problem, where the answer should no longer surprise us after working through the first problem. Let there be some military capability that the enemy wants and that we want him to have. This is something we want him to buy not just because it uses resources that he would otherwise spend on things that we deplore even more; this is something it is positively in our interest that he have. We want him to have it even if it is free of charge, and we might even provide it if he could not buy it himself. It might be a device to avoid nuclear accidents, to minimize the false alarms that might cause him to launch war itself, to reduce nuclear fall-out, to facilitate armistice, to permit the recovery of dead or wounded from the battlefield, to care for prisoners of war, to avoid the military destruction of fish and wildlife, or anything else. The point is, he wants it, and we want him to have it; we have no military interest in denying it to him, and if it gives him any sort of local advantage in a zero-sum contest somewhere, that disadvantage to us is outweighed by its beneficial effects over-all.

Suppose, now, there is some measure available to us with which we can destroy it. Or suppose we can affect the price he pays for it, causing it to be more expensive or less expensive. What should we do?

There is a lot to be said for being nice about it, if only to dramatize

[4] "Elasticity of demand" is strictly applicable only if the price of sites is fixed to the buyer, who chooses the quantity that maximizes whatever it is that he wants to maximize. From here on we shall use "elasticity" in a looser sense just to suggest the idea of a "revenue effect" and what happens to the budget available for other commodities.

the common interest, to set a precedent for preserving assets of value to both sides, to induce him to avoid taking reciprocal measures that would hurt us as well as himself, and for other essentially diplomatic or bargaining reasons. Leaving those aside—just considering our influence on his allocation of resources—we have two interests here. First, we want him to possess that capability; even if we threaten to destroy it, we want him to take steps to defend it. Second, we want to divert resources from other capabilities that we do not like into this capability that we do like, not just because we like this capability but because we like it to draw resources from the rest of his budget.

Again it is something like elasticity of demand that is relevant. If we know that he will defend that capability at all costs, because it is important to him as well as to us, then we can safely threaten it, forcing him to defend it, knowing that the net effect will be fewer resources available to the part of his force that we do not like. We threaten—to pick an extreme example—to spoof his warning system so that he would launch a successful first-strike against us—a first-strike that, though successful, he would rather not launch and would launch only in the mistaken belief that he was under attack. He anticipates our spoofing and spends money to tranquilize his warning system, to make sure that we cannot provoke him into an attack that he would prefer not to launch, to make sure that we do not bring down on ourselves a war that we deplore as much as he. We watch him cut back his ground forces, or his first-strike capability against our strategic forces, or his second-strike capability against our civilian assets, or some other program that opposes our interests, in order to maintain the capability that, though we like it too, he likes very much.[5]

But we must be careful. If we put the price so high that he prefers to do without it—if we threaten it in a way that makes its defense too costly for him—we divert no resources from other parts of his budget; we merely prevent his having something that we both badly want him to have. Or if the countermeasure is something that he can buy in varying quantities or degrees, raising the price or supply curve may cause him to buy less of it. If his "demand" is elastic he ends up with less of what we want him to have, more of what we do not want him to have, and we lose on both counts; if his "demand" is inelastic he ends up with

[5] In view of Professor Eisner's later comment, I acknowledge that this is an extreme example, chosen just to dramatize the point.

somewhat less of what we want him to have, less also of what we do not want him to have, and we must weigh the one against the other.

The point of this example is twofold. First: to emphasize that there *are* military capabilities that we want the enemy to have—not just that we dislike less than we dislike other capabilities, considering his marginal rate of substitution, but that we value positively, which we want him to have, and would even buy for him if he did not procure it for himself. Second: that just as we might want to *lower* the price of a force that we prefer him *not* to have, to induce him to draw resources from other parts of his budget that we like even less, we could want to *raise* the price of things we prefer that he have, for the same reason. Here it is a case of wanting the income effect on the rest of his budget to outweigh the substitution effect for this particular item.

A Diversionary Threat

Next consider the enemy's response to a new capability on our side to inflict damage on his civilians and their property. And suppose, just to clarify the principle, that what we threaten to destroy is of little or no military value. This could be because he cannot mobilize these civilian resources before or during war, or because they have low military value (a strong comparative advantage toward civilian value). Pain and life itself would have this quality, as would potential members of the labor force too young to work until the war is over; so would physical assets that would be more important after the war than during the war, and structures too specific in their civilian use to be transferable directly or indirectly into military strength. The idea is simply that, whatever the value of threatening such destruction, we get no military advantage from carrying out the destruction. Nor does the enemy suffer a significant *military* disadvantage. It is pure *civilian* loss, a kind of "side payment" that he makes but we don't get.

Aside from revenge, justice, and other moral and psychological satisfactions—and these could be negative—there appear to be two main purposes served by such a military threat directed against enemy civilian value. Both depend on influencing the opponent's behavior. One is *coercion* (i.e. bargaining advantage) and the other is *diversion* (of resources from other uses). "Coercion" implies that he refrains from or performs some action as part of a bargain, according to which we

carry out the destruction or not according to whether he refrains or not. "Diversion" implies that he anticipates our attack on his civilian value and does something to mitigate the effect. We may or may not want him to divert resources. If he does, it reduces our remaining coercive influence. But if we want him to divert resources—if that is the purpose—the result is brought about by coercion, by the threat that we will destroy civilian value unless he diverts resources to protect them. To the extent that he can attenuate our threat by shifting resources, he merely holds us to our bargain by carrying out his part.

Now, how do we feel about the costs to the enemy of countering our threat in some fashion? This will depend on how he counters it, and at the expense of what other capability. Two important possibilities are that he would defend his civilian assets and that he would procure a force of his own with which to threaten ours. If he defends his people, it could be at the expense of his battlefield forces—that is, the forces he could field in a "zero-sum contest" for territory or booty—or it could be at the expense of the force with which he threatens our civilian assets. (It could also be at the expense of some measures, of the kind suggested earlier, to which we attach positive value.) The diversion might also change the character of the forces he threatens us with, such as their first- or second-strike advantage, and change his expectation of the likelihood of war.

There are two limiting cases to consider: First, that the damage we threaten would *certainly* deter the enemy in all circumstances, so that we need no longer worry about how he allocates his military resources (and, if he could acknowledge the adequacy of our deterrence, he might as well disband his force); second, that though the damage we could do would severely reduce the net gain to the enemy from any war he might contemplate, it would certainly *not* deter him if on other grounds he had decided to have his war.

If he is certainly deterred, we have no interest in diversion. Our only interest is that diversion be incapable of spoiling our deterrence. We want the costs of an adequate defense to be prohibitively high.

If deterrence is clearly inadequate, our interest will be diversion. We want him to withdraw as much as possible from his other capabilities as the price of defending himself against our threat—at least, we do unless he withdraws resources from some capabilities that we prefer him to have. Assuming that he withdraws his resources from the battle-

field, or from a first-strike strategic force or from a second-strike force that does not shift in character toward a first-strike force under the pressure of a smaller budget, we want the price of defense to be *low* enough to be attractive, *high* enough to be costly. Again, "elasticity of demand" is the analogy, although we are likely to be less interested in the "price" of a "unit of defense" than in the whole curve relating damage to investment in defense. If defense is too costly—if it is too ineffective compared with its cost—the enemy will not invest in it. We get nothing for our trouble. In fact, anticipating that he will not defend his civilian assets, we should not procure the threatening force in the first place. (Paradoxically, the enemy may prefer that his civilian assets be indefensible in the face of a threat not large enough to deter him. The indefensibility of his cities makes our threat purposeless, and we may not bother to buy it.)

He may instead react by developing a reciprocal civil-damage force of his own. If he does, we want it to be costly. We may not like this response, even if it is costly, and prefer not to provoke it. From his point of view, if he can afford either to defend his cities *or* to threaten ours but cannot afford both, defending and threatening may be equally attractive. From our point of view they are *not* equally attractive.

These considerations are important. Burton Klein pointed out in his study of German war economics that one of the main consequences of the strategic bombing of Germany in World War II was the sizeable diversion of German resources into air defense—up to one third of all German military production went into active defense against air attack at the climax of the war.[6] If the purpose was merely to defend military strength, it was part of a zero-sum contest not related to what I have been discussing. But if the object was to preserve Germany and Germans, not just war potential, it was the diversionary principle expressed above. There is presently a corresponding question whether the United States should want the Soviet Union to invest, or not to invest, in ballistic-missile defenses.

Between the extremes of sure deterrence and no deterrence, the deterrent and diversionary motives can both be present. We might, for example, prefer a large Soviet investment in city defenses that, though inadequate to spoil deterrence altogether, would appear to the Soviet

[6] Burton H. Klein, *Germany's Economic Preparations for War*, Cambridge, Mass., 1959, pp. 232–233.

leaders a useful hedge against war yet would take resources away from their offensive forces or their battlefield forces. One can hope that the threat provokes enough expenditure on the defense of civilian assets to constitute a substantial diversion without the defense being so effective that it greatly reduces deterrence. How to weigh the two considerations depends on what it takes to reduce deterrence. One probably has to pay some price in deterrence in order to achieve the diversion, because the enemy will not waste money on a futile defense; still, an opponent could find it worthwhile to spend money to reduce damage in a possible war, without the extent of such reduction necessarily affecting his willingness to launch war or to provoke it. If we manage to draw the enemy's "supply curve" of defense—the relation of damage-reduction to money spent on it—we want to be careful how we draw it; anybody who then shifts it upward may do us as much harm as if he shifted it downward.

Adaptation and Costs

In zero-sum situations, the enemy's adaptation to a new initiative on our part always reduces the value of that initiative. We see an opportunity or a new capability, take advantage of it, and get a certain score at his expense if he does not adapt; his adaptation takes away some of our advantage. The net advantage to us depends on how costly his adaptation is; but in zero-sum situations, we always prefer that he be unadapted—that he not know of our new initiative, find adaptation too costly, or be institutionally unable to adapt. In nonzero-sum situations we can *prefer* that he adapt; and he can prefer to adapt, knowing that we prefer him to.

In fact, we can undertake moves that have no value for us unless he does adapt. The diversionary principle just discussed is one of these. Facing an enemy threat to our civilian values, we might be attracted to a completely defensive posture that virtually nullifies his threat. We might alternatively be attracted to a completely offensive posture because it *obliges* him to dismantle his offensive posture and devote resources to defense.

It is possible that our preferred posture in the face of an enemy offense is pure defense, but we like even better a pure enemy defense that is induced by a purely offensive posture on our side, and choose the offensive posture accordingly. If he does not adapt as we hope, we end up with offensive postures on both sides; that could be our last choice.

In this case, unless he adapts as we expect, improving our position as he rectifies his own, our move is a bad one, for us and perhaps for him.

Adaptive Cost-Avoidance: Comparative Advantage Again

Consider one more case, just to sharpen up the contrast between zero-sum and nonzero-sum calculations. The enemy is building a force, or about to buy a force, of one hundred "semihardened" missiles at a cost of $1 apiece. Our attack force can inflict 50 per cent attrition on his one hundred missiles if we attack. He could alternatively buy hardened missiles, fully invulnerable, for $2.20 apiece. His goal is a force out of which fifty missiles could survive attack; and at these prices he gets it for $100 with the more vulnerable missiles while it would cost $110 if he bought fifty of the more expensive invulnerable variety. Suddenly we discover a device that increases the vulnerability of the semihardened type of missile; with this device, even though it costs us something and reduces the size of our attack force, the enemy's one hundred-missile force would be 75 per cent vulnerable and only twenty-five missiles would survive. What do we get for our money?

If he fails to adapt—if the device is our secret—we get a reduction by half of his surviving force in the event we attack. If he adapts by increasing his force of vulnerable missiles, we oblige him to buy twice as many (to reach an expectation that fifty will survive) and raise his costs by $100, the resources being detracted from somewhere. But he has another option: he can switch to the different type of missile and get a potential force of fifty surviving our attack for a cost of $110. For the sake of the example, suppose that his procurement is not too far along to permit the switch; perhaps the one type of missile is converted into the other by additional expenditure on the missile's environment of $1.20 per missile.

He switches, gets an estimated fifty surviving attack at a cost of $2.20 each, spending a total of $110, or $10 more than he originally planned. What did we get for our money—the money we spent on the device that caused him to switch his missile-procurement plan?

Mainly we got a qualitative change in his force—that is, a quantitative trade along two dimensions—together with a modest drain on his military resources elsewhere. He would have had a second-strike force of fifty and a first-strike force of one hundred; now, faced by our new

device, he procures instead a second-strike force of fifty and a first-strike force of fifty. Besides costing him $10 we have cut his first-strike force by fifty.

We like it; and we like it quite out of proportion to the $10 cost we inflicted on him.

In fact, our device may cost us so much that inflicting an extra cost of $10 would be a poor bargain; but by cutting his first-strike force in half without firing a shot we may achieve something well worth the cost.

His cost increase can be a gross understatement of what we gained by imposing on him the *incentive* to incur the cost. We tilted his trade-off line and were the beneficiaries of his shift in procurement. (In the same way we might have sold him, or given him, something that would have reduced the cost of the more expensive missile, and been the chief beneficiary of his savings. In fact, we could have subsidized, with cash, his procurement of the more expensive missile, and come out ahead.) Our gain is of the same algebraic sign as his cost, but bears no particular proportion to it. We could have had the same result if the hardened missile cost $2.01 or $3.99, except for the net cost to his other programs. This net cost is really separable from the reduction in first-strike force and has to be evaluated altogether separately. At $3.99 we are more likely to enjoy the result (unless the resources come out of that fail-safe system we were concerned about earlier), at $2.01 we are less enthusiastic; but the drain on his resources is separable from the shift in first-second-strike composition.

Notice this, too. Had he been able to buy an even "softer" missile that was already 75 per cent vulnerable and unaffected by our new offensive device, but that cost, say, $0.55 and was not competitive with the semihardened, our new capability would make the semihardened a poor buy and his response might be to procure two hundred of the soft ones for $110. This we do not like; his first-strike force is doubled at a net cost of $10 and, even if he does not much care, we do. If both the "hard" and "soft" alternatives are available, minor shifts in the ratio of cost to survivability can make, from our point of view, major changes in his force: at $0.53 per soft missile, with hard ones at $2.20, he buys two hundred soft ones for $106 while at $0.57 per soft missile he buys fifty hard ones for $110. To him, if second-strike ability is what he wants, the decision is marginal; to us it is systemic.

I am not now trying to illustrate, once again, that we may like or dislike what our initiative made him do, but that *minor* variations in his relative costs can make *major* differences in our evaluation of his response. A consequence of this is that no single measure of his cost —whether discounted, augmented, or changed in sign from the original estimate to make it a better measure of what we gain—can be used as a "correction" of his costs to reach their net worth to us. It is not just that the gain or loss to us is poorly measured by the cost inflicted on the opponent and needs to be corrected; it is that the gain or loss to us depends on the *particular* way that he reacts in a particular case, on the particular incidence of the new costs within his budget, on the particular price comparisons that determine his response. Two measures that identically raise *his* costs can induce changes in *our* payoffs that bear no relation to each other. No "exchange-rate" correction will take care of this. In the zero-sum case his marginal decisions are marginal to us, and we can use a single measure of "costs" to avoid tracing the ramifications of his response. In the nonzero-sum case there is no short cut; the ramifications have to be traced.

Conclusion About Enemy Costs

What conclusion do we reach about the role of inflicted enemy costs in evaluating our own alternatives?

The main conclusion is that costs, in a nonzero-sum context, do *not* reflect the changes in payoffs outside the arena of our suboptimization. The reason is simple: the costs we inflict, or threaten to inflict, on the enemy can *at most* reflect the changes in *his* payoffs outside the local arena. They do not reflect the relevant changes in *our* payoffs because there is not the convenient perfect negative correlation between his losses and our gains once we drop the zero-sum restriction, that is, once we extend our analysis into the (almost universal) nonzero-sum contexts.

In the zero-sum case his costs reflect the gains to us that our local analysis would otherwise have left out of account. In the nonzero-sum case neither the magnitude of his costs, nor even the algebraic sign, reliably suggests the pertinent changes in our payoffs. Our gains may be greater than his costs suggest, less than they suggest, or opposite in sign

to what they suggest, and will differ according to the particular choice we pose for him.

This strikes a blow at "suboptimization," of course. What makes partial analysis valid is often the dependability of enemy costs as a short cut to all the ramifications outside the local contest. When costs are recognized as undependable—when they are no index to magnitude or even sign, neither an upper nor a lower limit to the effect on us of the adversary's adaptation, to what happens outside our local field of vision—we have to extend our field of vision.

And as mentioned above, our attitude toward his adapting to our own behavior is different; instead of reducing our advantage it can raise it or offset it in a manner not communicated by his costs. As a corollary, our attitude toward secrecy and information can be dramatically different from what it is in the zero-sum case, as Martin McGuire has recently emphasized.[7] Taking the first move and letting the adversary adapt is never the right choice in a zero-sum situation; it often is in the nonzero-sum.

[7] *Secrecy and the Arms Race,* Cambridge, Mass., 1965, especially Chapter 7, pp. 212–232.