

from:

Axelrod

date:

To arrive at the equilibrium position in the absence of threats we might use a function analogous to the threat function — a non-cooperation function.

The threat function states the probability of making a concession of given size as a result of a threat of given size, issued with given <sup>perceived</sup> probability of being carried out.

If we think of bargaining without threats as a process of tatonnement in which each player "threatens" non-cooperation unless he gets a concession relative to the tentative agreement reached, we can say that each has a function connecting the probability of conceding to a given demand ~~or~~ according to the size of "threat" and its credibility. (The size of the "threat" is the amount the player would still stand to gain in the proposed position.)

But of course what determines credibility  
is the other side's potential loss from  
withdrawal of cooperation



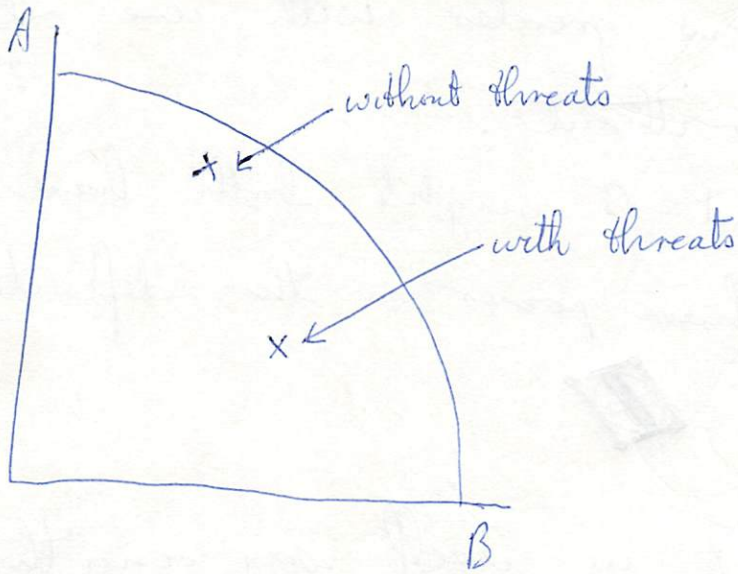


MEMORANDUM

to :

from:

date:



If a player does better when threats are used (B does in the illustration) we can say that he benefits from the use of threats; and if a player does worse we can say that he is damaged by the use of threats (as A is in the illustration).

It may be that neither gains from the use of threats (because of the possibility of threats being carried out and of threats simply producing deadlock in the status quo position); exceptionally both might gain if threats are successful in pushing them towards agreement near the P-optimal frontier in the top right quadrant.



## Possible conception of power:

A has power over B if A's expected utility is greater with use of threats than without.

(Note: A and B might both have power or neither have power on this definition)



NB this doesn't in itself say anything about where the outcome is without the use of threats (where threats = saying you will make the other person worse off than in the non-agreement position; non-co-operation is present as a tactic in the non-threat situation since it means saying that you will see the other person gets the non-agreement position.





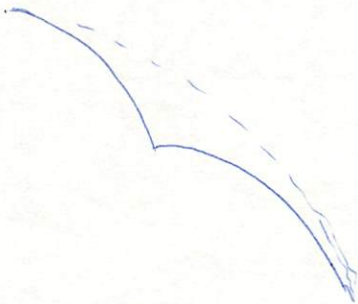
MEMORANDUM

to :

from:

date:

Note that ~~P~~-optimal line can be constructed graphically by taking the envelope containing the most upper-right points when all trades are represented graphically.





COMMENTS ON "POWER IN NON-VOTING GROUPS"

by

Brian Barry

July 1971



Response to invitation to submit advance comments on "power in non-voting groups".

Brian Barry,  
University of Essex.

1. Is there any need for a special concept of "power in non-voting groups"? I think not, but this is because I do not think there is anything special about power in voting groups. Power, as I understand it, is, roughly speaking, the ability to get other people to do what you want. In relation to voting groups this might apply to (a) the power of the group collectively to get others to do what it wanted or (b) the power of a given member to get other members to do what he wanted. In neither case would the fact that it was a "voting group" be of any importance. To use "power" to refer to the a priori or observed frequency of being pivotal seems to me a misuse of the term.

2. What is the point of employing a concept of power in social analysis? We often want to say, in advance of any actual confrontation between potentially conflicting parties that one has a "strong position" and so on. Such statements are connected by an empirical generalization to the prediction that, other things being equal, a party in a strong position will tend to get what it wants if it comes into conflict with a party in a weak position. To have a "strong position" in this sense is to have power and the analysis of power consists in (a) discussing what makes up a "strong position" and (b) discussing how the process of exploiting a "strong position" works.

3. Definition of power. It does not seem to me that the main problem lies in the definition of the concept of power. There is nothing wrong with the often-repeated kind of definition in terms of the ability of an actor to get what he wants over the resistance of some other actor or actors. The problems occur when one tries to specify

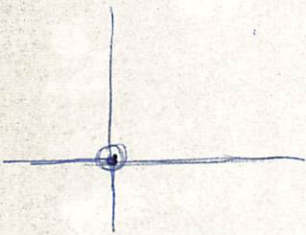
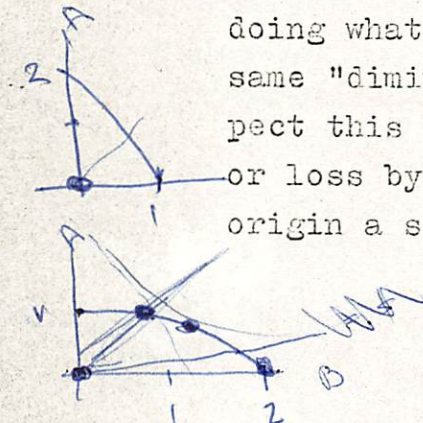


what makes for a greater or less ability and how the ability can be put into effect. Without a full and reasonably rigorous analytic framework discussions of these questions easily founder on unstated assumptions, ambiguities in the specification of hypothetical cases, etc.

4. Object of the paper. At present we have a quite elaborate formal literature on bargaining and on the tactics of conflict and we have a non-formal literature on power. The need is, I suggest, to bring these two approaches together and attempt to treat power in the context of a formal theory of bargaining and conduct generally in conflict situations. This short paper is offered as a modest contribution to that end. I have deliberately restricted myself to setting out the main ideas without elaboration because I think it will probably serve the purposes of the conference best to concentrate attention on the broad outline rather than get bogged down in detail.

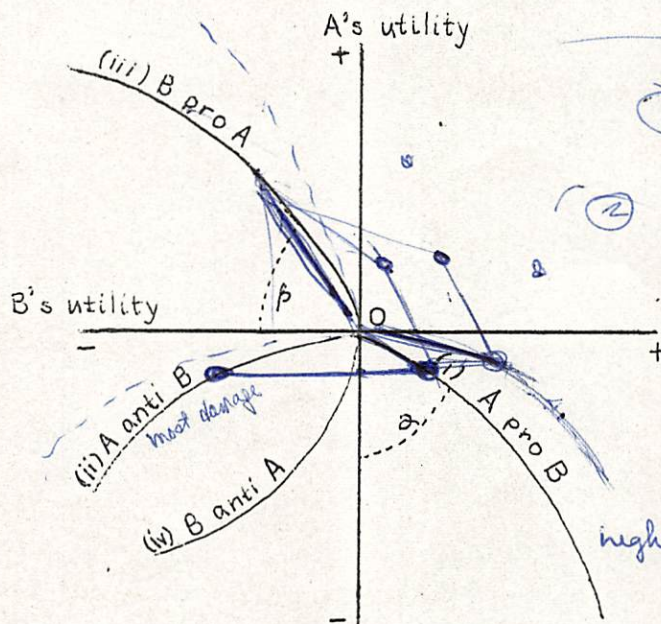
5. Four basic curves. Schelling has pointed out that a payoff matrix allowing each player a quite limited choice of strategies at each point in a sequence of a few moves each can easily generate over  $10^{100}$  cells. Obviously some simpler way of representing possible outcomes is needed. I suggest that, taking a two-person situation as our object of analysis, we can in principle represent the most important facts about their relationship in terms of four curves: (i) for any level of sacrifice of utility by A, the highest amount of utility with which he can provide B; (ii) for any level of sacrifice of utility by A, the largest loss of utility which he can inflict on B; (iii) and (iv) the equivalents of (i) and (ii) substituting B for A and vice versa.

6. Shape of the curves. In general it seems reasonable to suppose that all four curves will show the cost of producing either good or ill for the other actor rising at an increasing rate. The "cost" may in some cases start off as a positive utility - A actually likes doing what makes B better off (or worse off) - but on the same "diminishing returns" assumption we should not expect this to be so for indefinitely large amounts of gain or loss by the other actor. Taking the status quo as the origin a set of four typical curves might look like this:



BARG GAME

FIGURE 1



Normally put A on X axis

- ① what assumptions about bag etc?
- ② eg what does it look like for bag game? (start w dotted line)

highest A can give B for given sac by A

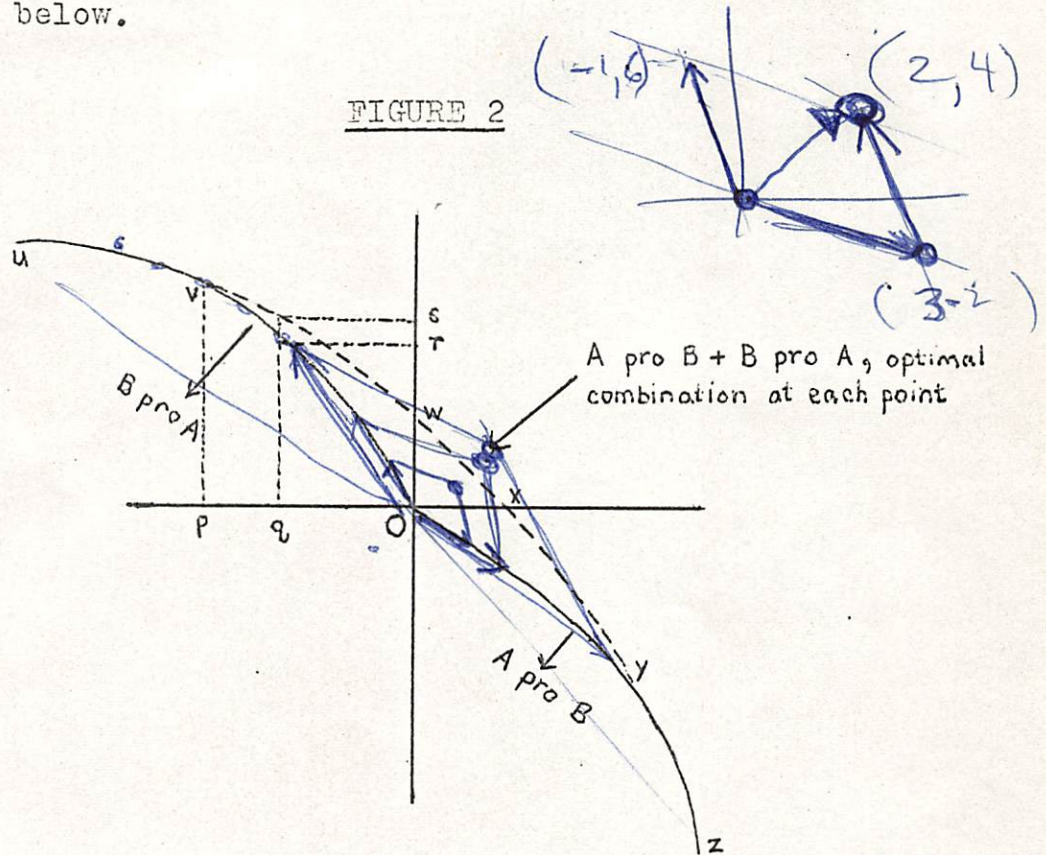
7. Gains from trade. If the curves (i) and (iii) exhibit "diminishing returns" and the angles  $\alpha$  and  $\beta$  sum to more than  $90^\circ$  near the origin, gains from trade are possible. (If the curves are of any other shape things are less simple but there may also, of course, be possible gains from trade over certain ranges.) A situation in which there were potential gains from



trade would be one in which, if A has the opportunity of getting B to move  $x$  units to the left of the origin along curve (iii), then for some  $x$  there would be a position preferred by both A and B in which B moved further along (iii) and A moved some way along (i), i.e., in which instead of a unilateral transfer there was a one-sided exchange. The discussion of gains from trade has been dominated (unfortunately for other purposes) by the special case of noncoercive bargaining where, unless A and B both benefit from a departure from the status quo, the parties stay at O. This has led to a tendency to think in terms of a false dichotomy: gains from trade and noncoercive bargaining on the one hand, unilateral transfer and coercion on the other.

8. The Pareto-optimal frontier. A simpler way of envisaging gains from trade is graphical: gains from trade are possible when the Pareto-optimal frontier is at some point to the upper right hand side of the curves (i) and (iii) which meet at the origin. The Pareto-optimal frontier is, of course, simply the line about which we can say that, given the objective possibilities of the situation, if A and B are on it then A can only get better off if B becomes worse off and vice versa. If there is no possibility of gains from trade, the Pareto-optimal frontier just consists of curves (i) and (iii); since this combined curve runs through O that tells us that the only way A can improve on the status quo is to make B worse off than in the status quo and vice versa. If there are gains from trade possible this means that some points on the combined line are dominated by others which represent the net effects of bilateral transfer. We arrive at the Pareto-optimal line by considering all the conceivable exchanges between A and B running from A giving nothing to B while B gives the maximum possible amount to A through to the opposite extreme. Each of these conceivable exchanges gives rise

to a pair of net utilities for A and B. If we eliminate all those pairs which are dominated by some other pair (e.g., A 6 units B 5 units by A 6 B 6) we are left with the "efficient" points. If there are no possible gains from trade these will simply be unilateral transfer, all the bilateral transfers having been dominated; if there are possible gains from trade we get a curve like that shown below.



9. Exchange and unilateral transfer. (1) Notice that at some point the Pareto-optimal line, even when gains from trade are possible, coincides with the curves for unilateral transfer. (In Figure 2, it extends beyond them along the line vwxy.) This means that, if A is able to extract from B (presumably by coercion) a utility loss greater than  $p_0$  it will be most efficiently accomplished by B doing something that A wants but A doing nothing for B. Thus the fact that there are "gains from trade" as defined does not guarantee that every transaction will be most efficiently accomplished by an exchange.



(2) Perhaps more counter-intuitive is the converse point, already touched on, that, if A is able to get B (again presumably by coercion) to accept a loss of utility (relatively to the status quo) of an amount less than  $p_0$ , he can do better for himself by proposing an exchange than by insisting on a unilateral concession by B. Thus, returning to Figure 2, if A can extract from B a utility loss of  $q_0$  he can get a gain in utility of  $O_r$  by insisting on a straight transfer but, by requiring the optimal exchange with a net loss of utility of  $q_0$  to B he can secure  $O_s$  - a gain of  $r_s$ .

10. Excursus on conflict and conflict of interest:

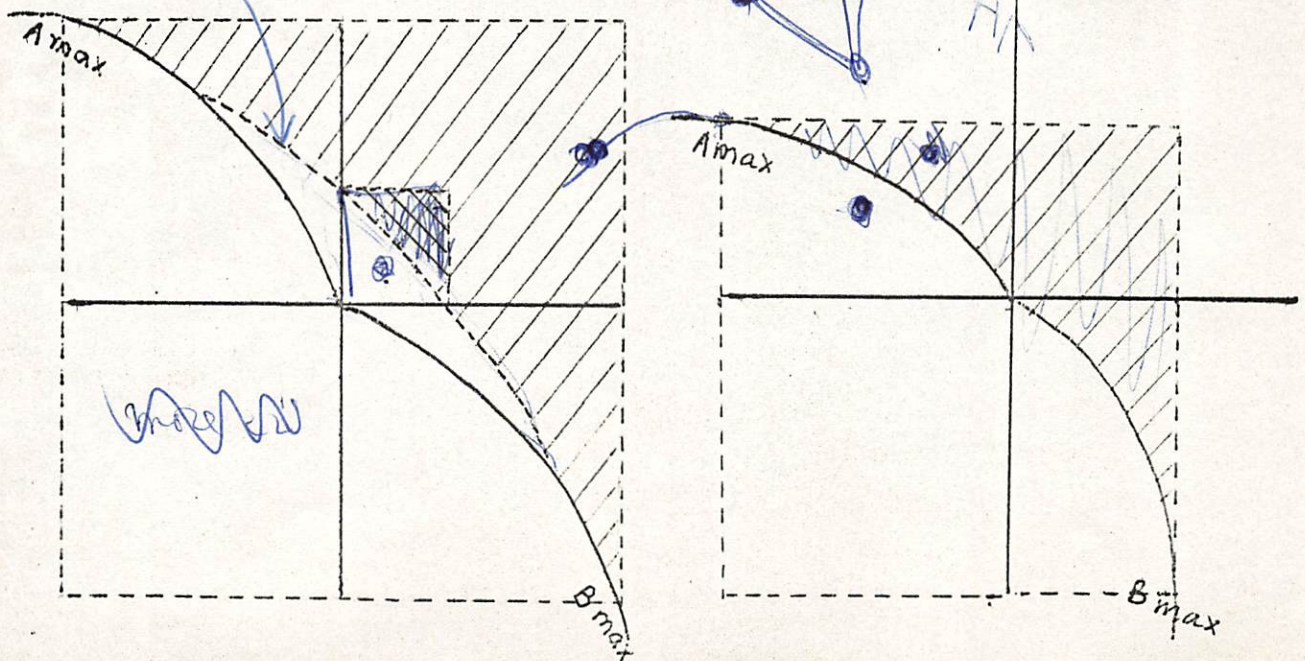
(1) Conflict of interest. The situations we are dealing with exhibit conflict of interest in that the outcome most preferred by A is not that most preferred by B. To get an idea of what is at stake in a confrontation we can look at the points which A and B respectively most prefer, shown as A max and B max in Figure 3. (The significance of the turning points such as A max is that, beyond this point, although there are things that B would dislike doing more than those represented by the point, they would not give A as much satisfaction. Hence A has no incentive to try

*why desired?  
OK*

*as but I think so relevant (7-8)*

FIGURE 3(1)

FIGURE 3(2)





to secure them, and they are not Pareto-optimal since they are dominated for both A and B by A max. N.B. that if what gives A pleasure is B's suffering as such there will not be a turning point and A max will simply represent the maximum suffering possible for B.) Generally speaking we can say that there is "more at stake" between two actors (a) the higher the utility gain to A at A max, (b) the higher the utility loss to B at A max, (c) the higher the utility gain to B at B max and (d) the higher the utility loss to A at B max.

Nash defined  
OK

11. Excursus: (2) Axelrod. Robert Axelrod, in his book Conflict of Interest, defines the concept for a bargaining game in which each player can guarantee himself the status quo as the cross-hatched area in Figure 3(1) or, in relative terms, as the cross-hatched area as a proportion of the rectangle in which it is situated. He also discusses "Prisoner's Dilemma" games but does not discuss the general case where players have means of coercion at their disposal. In the spirit of his treatment of the Prisoner's Dilemma, however, we might take the rectangle whose corners are A max and B max and then look at the shaded area to the upper right of the Pareto-optimal line. It is worth noticing that even in terms of relative conflict of interest on this criterion there may well be more conflict of interest where gains from trade are possible (Figure 3(1)) than when they are not (Figure 3(2)). It should also be borne in mind that until we analyse the threat curves of A and B (curves (ii) and (iv) in terms of our earlier discussion) we are unable to say whether the whole range between A max and B max is really "available" to the actors. All we are saying at the moment is that this is the widest area of disagreement there could be given the actors' preferences and the actual transfer-of-utility relationship between them. Even this, however, should not be understood as entailing that B cannot do worse than at A max or A than at B max: if the actors

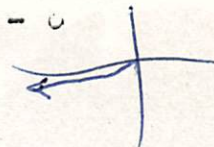
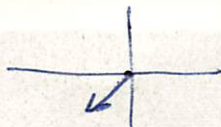
Wrong?  
Harvard  
prop. 100  
origin?

depends on  
threats, coop  
vs noncoop

helping?  
same  
defect

? single  
tech?





apply sanctions to one another they might both finish up to the lower left hand side of the rectangle.

*of interest ab*

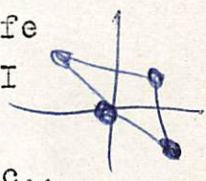
*behavior of interest*

*Don't use my not JCR term*

12. Excursus: (3) Conflict. This kind of measure of conflict is introduced by Axelrod as an empirical predictor of actual conflict. But "conflict" as Axelrod defines it is not what most people would think of as conflictful behavior since it is simply the failure to reach agreement on a mutually advantageous trade. He later extends it to cover the case where both players in a Prisoner's Dilemma game make the unco-operative or "double cross" move but this is still a long way away from the complexities of real-life conflict that we are trying to capture. Most people, I suggest, think of conflict as what is going on at the moment in Vietnam, East Pakistan, Northern Ireland, etc., that is to say the imposition of sanctions by one party on another or (more clearly) the mutual imposition of sanctions by two or more parties. Notice that this means successful coercion does not constitute conflict: if I threaten you with sanctions unless you do x and you comply with my demand this is not conflict. Law-abiding citizens are not in conflict with the police even if it is only the presence of the police that keeps them law-abiding, nor was there conflict in Northern Ireland during the years of unchallenged Protestant ascendancy. There is no advantage in conflating the concepts of coercion and conflict. At most we might want for some purposes to say that obedience based on coercion carries the seeds of possible future conflict.

*letter or both (90D)*

*def cb!*



*good!*

13. Winning and conflicting as outcomes. Power, as I have defined it (along with Max Weber and other worthies) has much the same relationship to winning a particular encounter (getting what you want) as has conflict of interest to actual conflict in Axelrod's conception. Both features of outcomes - who won and were sanctions used? -

*cc:cb = power:win*



are of interest to people for a whole variety of reasons, which is why they are singled out for special attention. And since winning and the use of sanctions are important features of outcomes it is not surprising that we want to identify features of situations which are likely to give rise to these outcomes.

14. Difficulties in power and conflict of interest.

But in both cases there are two kinds of difficulty. First, the things to be explained (or predicted) have a number of different aspects and there is no obviously right way of bringing them together as an "amount" of winning or conflict. (If winning is getting what you want over resistance, how do we weight the two components - is it a bigger win to get a lot of what you want over little resistance or to overcome an immense amount of resistance to get a little of what you want? How do we weight retaliation against primary coercion, and where do attempts to destroy coercive or retaliatory capacity fit in?) Second, once we get away from simple, highly-structured situations of the kind dealt with by Axelrod we have to admit that there are a number of features of a strategic situation which can affect the probability of a given party winning or of the parties becoming involved in conflict and again there is no straightforward way of reducing them to a common measure. We may have to confine ourselves to a series of statements of the form "Other things being equal, the more x the greater probability of A winning (of conflict)" with perhaps some indication of the relative importance of these factors.

15. Implications. This, it should be said, is not an argument either for abandoning concepts like power or conflict of interest (better here: conflict-prone relationships) or for producing more differentiated definitions.



To bring together two points made already in different places: first, winning and conflict are things we are interested in for all kinds of evaluative, practical and other purposes; and any attempt to cut them down by arbitrary restrictions on their scope to make them more manipulable will almost certainly exclude aspects which are important. Second, if we have an adequate analysis of the whole phenomenon of strategic interaction in which both concepts are embedded we shall be able to see our way around both concepts, see where the conceptual choices <sup>lie,</sup> and so on; if we do not, no amount of definitional elaboration will be a substitute.

16. Winning: efficient outcomes. If the object is to say what kinds of position conduce to getting what you want we need to establish more precisely the criterion of getting what you want. For the present purpose I shall not compare different strategic situations but merely ask what in a given situation would count as getting what you want. Taking our two kinds of situation again, we can see that in Figure 4 case (2) does not present problems if the actors finish up on the Pareto-optimal line, since the only move from the status quo (by definition, a draw) is either into the fourth quadrant (A wins, B loses, relatively to the status quo) or the second quadrant (B wins, A loses, relatively to the status quo). In Figure 4(1)

Wrong?  
 A can threaten  
 B with no  
 Pareto opt  
 unless A  
 gets more  
 along p

FIGURE 4(1)

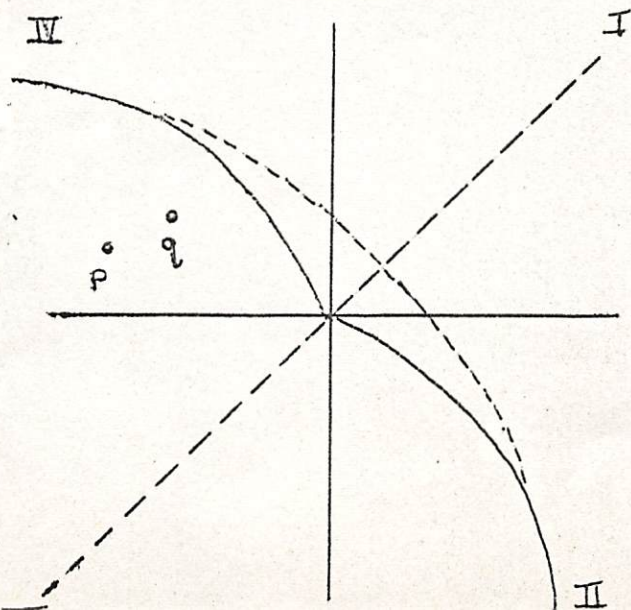
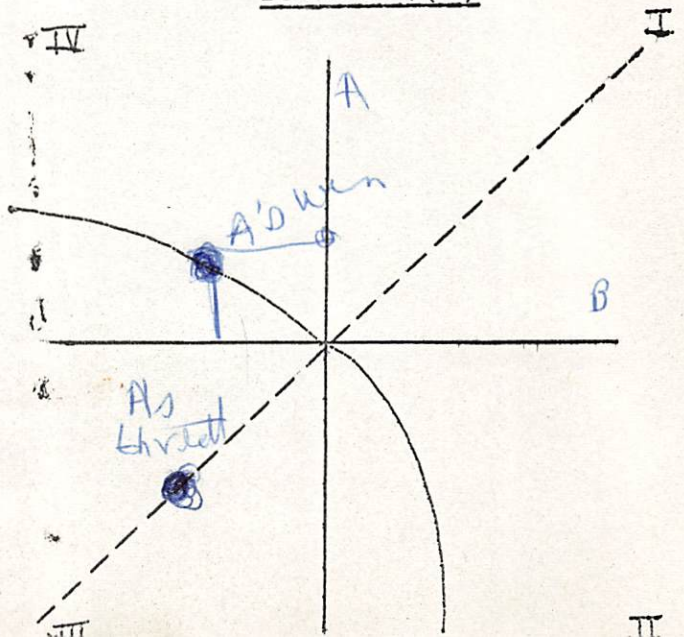


FIGURE 4(2)



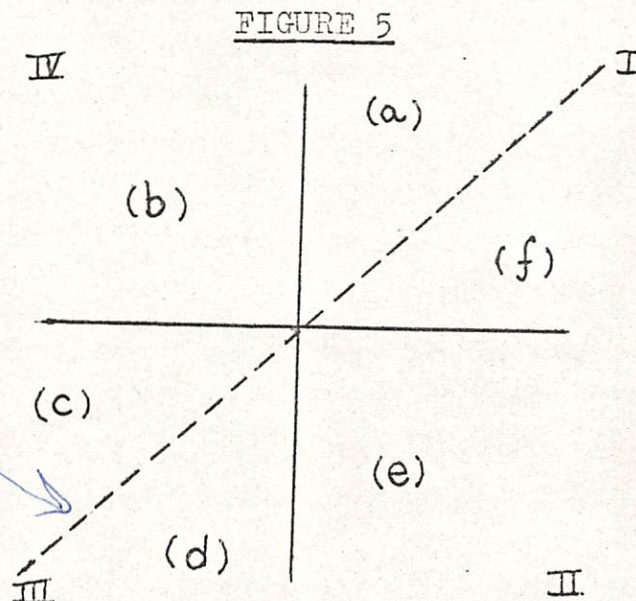


there is again no problem if the result is on the Pareto-optimal line and falls in quadrants II or IV. If, however, it falls in the first quadrant both parties gain relatively to the status quo. Does this mean that both got what they wanted, and we stop there? No: the line between quadrants is only a line and what we really have is a continuum. We often speak of someone "getting the best of the bargain" and by this we mean that, although both parties gained, one gained more than the other. The line of equal gain is, obviously, a line running from the origin and bisecting the first quadrant: to the left of this A gains more than B, to the right of it B gains more than A.

require  
interp  
comp  
int.

said p 12

17. Winning: inefficient outcomes. Any point is one where (a) A gains more than B, (b) A gains and B loses, (c) A loses less than B, (d) B loses less than A, (e) B gains and A loses, or (f) B gains more than A. The areas corresponding to these possibilities are shown in Figure 5.



The dividing line between the first three, in which A comes out ahead of B, and the last three, in which B comes out ahead of A, is, of course, the dotted line shown. Outcomes (c) and (d) are ones in which neither



party gets what it wants and, although we <sup>can</sup> compare losses, no outcome in the third quadrant can be described as a successful attempt to exercise power by either party. The comparison of other non-Pareto-optimal positions or of Pareto-optimal positions with non-Pareto-optimal ones presents difficulties. Consider, for example, points p and q in Figure 4(1): q gives A a slightly better outcome but p is much worse for B. How we relate these depends on our purposes: if we are mainly interested in the particular case we may concentrate on how well A did but if we are looking for an indication of the strategic relationship we may well be more impressed by the loss that A has succeeded in getting B to accept. We should not, in my view, be in too much of a hurry to lose information and should think of an outcome as in general consisting of two elements, A's gain/loss and B's gain/loss.

right } 18. Note on utility. It will be noticed that when the outcome is in the first or third quadrants we have to make an interpersonal comparison of utility. This is, I suggest, unavoidable if we are to capture the notions of "getting the better of the bargain" and of one person losing more than the other. In fact, quite a lot of things can be said without going beyond individual ordinal utilities (i.e., each party being able to say simply whether one outcome is preferable to another or whether two outcomes taken together are better or worse than some other outcome) or individual cardinal utilities (i.e., outcomes taken together with assigned probabilities compared with other outcomes). And it is obviously worth while to see how much of the structure can be reared on these more slender foundations. But as a matter of intellectual strategy it seems to me that the first priority is to be able to cover within an adequate framework all the phenomena we want to be able to talk about.

In many bargaining parleys, reference is made - vaguely, to be sure - to interpersonal comparisons of values, and presumably mathematical abstractions of such situations should incorporate comparisons. (Luce and Raiffa, Games and Decisions, pp. 131-2)



19. Rationale of the equal-advantage equilibrium. If both parties have an accurate perception of each other's utility schedules the position of equal gain may, in the absence of attempts at coercion, have a conscious appeal of symmetry. Luce and Raiffa thus argue for the point at which the equal-advantage line intersects the Pareto-optimal line as a "reasonable" solution (pp. 143-5). If, on the other hand, the parties have a chance to misrepresent their utility schedules, each has an incentive to pretend that it stands to gain less than it does, subject to the danger that if both do this too vigorously they may be unable credibly to accept what is in fact a mutually advantageous bargain. (The recent negotiations between Britain and the E.E.C. for British entry illustrate this, with Britain, especially under the Labour government, making gloomy noises about the costs and the French especially on the other side saying that if any more concessions were given it wouldn't be worth having Britain in. A complication (which greatly weakened the British position) was the need of the British government to avoid being too pessimistic publicly in case this made it too difficult politically to go in on terms the government was prepared to accept.) But even here we can suggest that the equal-advantage position makes some sense by taking a stronger (and more plausible) version of the idea used to support the Zeuthen-Nash bargaining model, which did not in its original form involve interpersonal comparisons of utility. Let us postulate that, in a bargaining situation, the actor with the most to lose, objectively, if the negotiations fail, will be the first to make a concession. It then obviously follows that the actor who is in a position of relative advantage will offer a concession so the long-run tendency must be to finish up on or near the 45° line.

20. Non-co-operation and coercion. We can get to the same conclusion if we imagine that an exchange has been

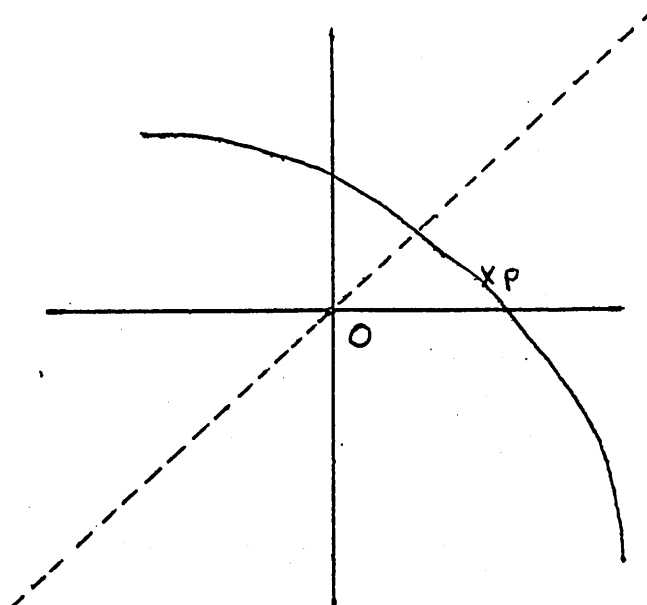
But this  
not require  
comp of  
utilities  
Raiffa  
minimum  
comp of  
\$.

needs more rigor but ok



*just as much*  
 established on a regular basis and there is some question of its ceasing. Blau, following Rupert Emerson, takes this as a paradigmatic situation in which power can be exercised. Like so many theorists influenced by economic models he considers only the curves representing the best A and B can do for one another, and does not consider in any serious way what is surely the essence of power, the capacity to do deliberate harm. In this kind of myopic analysis, we have already seen that mere non-co-operation in striking a bargain is treated as if it were "conflict"; and we can now add that the withdrawal or co-operation is treated as "coercion". Although this is perverse, we can certainly say, with Blau, that if one party gains much more than the other from an exchange then that party is vulnerable to demands. (Blau, of course, obscures things further by treating this kind of relationship as the core of the phenomenon of power.) Thus, in Figure 6, if the parties are at p, A is in a strong position to demand something else as a condition of continuing to exchange on these terms (thus bringing the net utility gain nearer equality) because B stands to lose much more than A from a withdrawal of the exchange. We can show the situation alternatively by putting the established trading position as the status quo and drawing a line to the non-trade

FIGURE 6



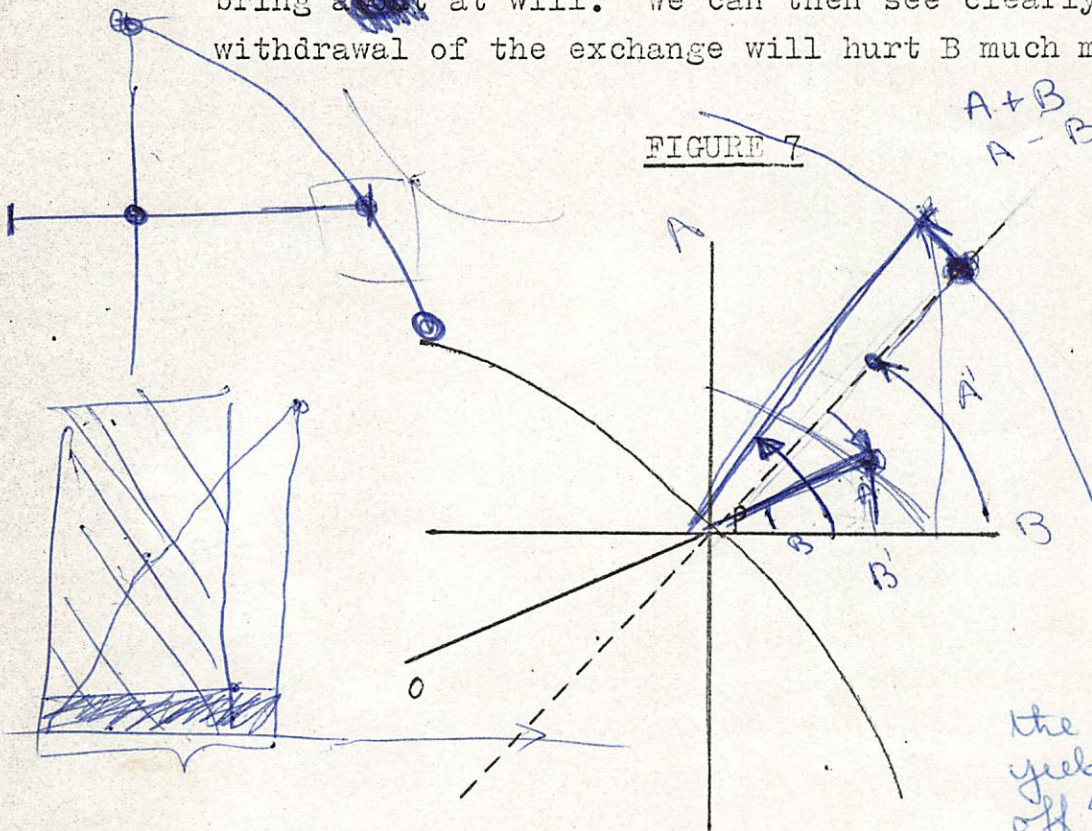


sure no trade

mix all fruit  
~~being~~ no fruit

position to depict a state of affairs either party can bring about at will. We can then see clearly that the withdrawal of the exchange will hurt B much more than A.

FIGURE 7



$A+B$   
 $A-B=0$

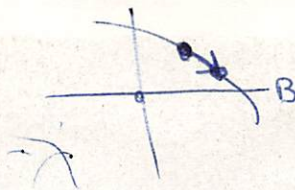
$\frac{A'}{B'} > \frac{A}{B}$   
 $A'B > B'A$

Note then that once power, exercised is no longer useable again (altho the same resources may yield more later). ie once off 45° line, the origin and line presumably more unless a prior way of setting

21. Exchange and power. To say, however, as Blau it as <sup>is true</sup> does, that to be taking part in an exchange on disadvantageous terms constitutes power, seems to me very peculiar. In the absence of any other information we should surely be more sensible to treat the fact that A is taking part in a bad bargain as prima facie evidence that B is exerting power over him. Even if we have the additional information that neither party is in a position to employ coercion against the other, we still have no need to speak of the party with the worst of the bargain having power; all we need say is that he could get a better deal if he tried. By power we should mean the ability to get <sup>at origin</sup> the outcome off the 45° line in one's own favour - not merely the fact that one is off it to one's own disfavour and could get on to it.

But if one can improve oneself he has power. Being unfairly treated my thus give you power  
 Question then is "how def where to put origin"





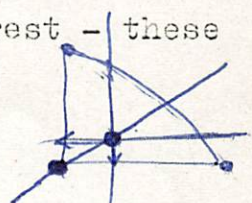
22. Sanctions and power. We have thus, finally, arrived at the point where it can be seen that power depends upon the capacity to apply sanctions. This has, of course, often been asserted as part of the definition of "power" but we have arrived there not by definitional fiat but by an analysis of the prerequisites of being able to get what you want by overcoming resistance and "getting the best of a bargain". Admittedly, one may, by the exercise of guile, by taking advantage of the confusion or incompetence of the other party, or by using the kind of option-foreclosing moves described by Schelling, sometimes come out ahead in a bargain without threatening sanctions, or even occasionally get the other party to agree to something that makes him worse off than he is in the status quo situation. But to be dependably in a position to do this surely requires the availability of sanctions. (Of course, A may be able dependably to get B to do what he (or she?) wants simply in virtue of the fact that B wants to do what A wants him to do. But although this is a way of getting what you want it is not an exercise of power because it does not involve the overcoming of resistance: once A has stated a wish the optima of A and B coincide, the north-east corner of the Pareto-optimal curve is the most northerly and easterly point on it, and there is no conflict of interest - these are just three ways of saying the same thing.)

*Good*

*40*  
*SA*

*3A 60*

*SA 50*  
*SA 50*



23. Threat curves. The analysis of power is thus largely the analysis of the shapes of threat curves and the strategic implications of various combinations of threat curves when occurring with various combinations of transfer curves. The full treatment of these strategic possibilities, and an attempt to relate them in a serious way to empirical phenomena, would require a book, and I in fact intend to write it unless the comments of the other participants at the conference knock irreparable holes in the general notions presented so far. Clearly,

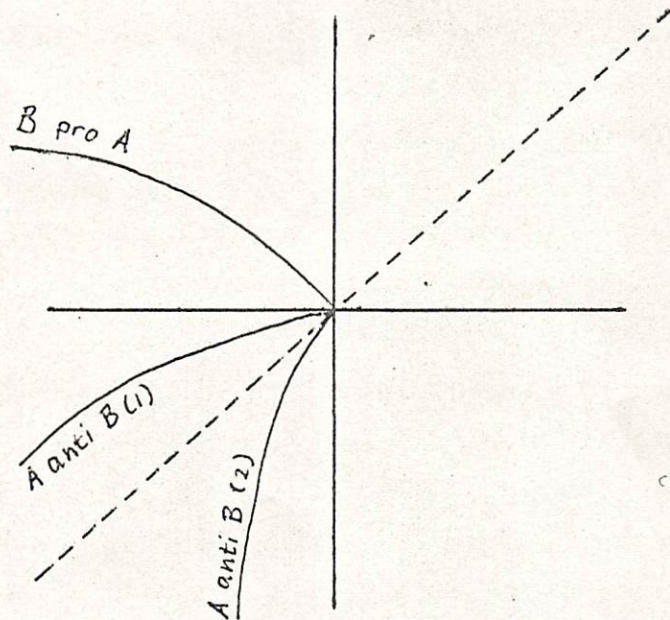
*See empirical lit on barg.*

*(eg Sengal + F.)*



however, we can begin by saying what is a good shape for a threat curve from the standpoint of its owner and what is a bad shape. In Figure 8, of the two threat curves shown for A against B, (1) is favourable to A and (2) is unfavourable to A. By referring to the 45° line it is easy to see that with curve (1) A can do B a lot of damage at a little cost whereas with curve (2) he has to incur a large cost to himself to inflict a little damage on B.

FIGURE 8



*not necessary*

24. Unilateral use of threats. Under the schooling of writers such as Ellsberg and Schelling, we have grown to accept that the essence of a threat is not "This will hurt you more than it hurts me" but "Doing what I want will hurt you less than I intend to make it hurt you not to do what I want". In suggesting that a "strong position" depends on the shape of the threat curve, am I ignoring this? I think not. There are two basic questions to ask about a threat: (1) Is it worth making? and (2) Will it be successful? Whether it is worth making depends on the answer to the second question plus the answers to three other questions: (a) What is the gain if it is successful?

*NOT NEEDED just read "This will hurt me (ie me) a little and you a lot. I can take it. You can't."*

*Only comparison needed is thru behav, ie in war whether the threat is easy to make, you surrender.*



(b) What is the loss from carrying it out? and (c) What is the loss from not carrying it out? Whether the threat is successful depends on (d) the estimate the other actor makes of the probability of the threat's being carried out if he does not comply and (e) how unpleasant the threat would be if carried out compared with the unpleasantness of complying with the demand. Of these five factors, (b) and (d) depend on the shape of the threatener's threat curve; (a) and (e) depend on the shape of the threatener's transfer curve. Holding the shape of the transfer curve constant, the more favourable the shape of the threat curve the more likely it is that making a threat in the hope of securing a concession will be a good bet. Thus, in Figure 8, A looks well placed to make a threat with curve (1) but not with curve (2). Further developments of this line of thought depend on the specification of equations connecting the rates of change of costs, benefits and probabilities. I have done some preliminary work on this but any elaboration of it would require a lot of space.

25. Mutual use of threats. So far I have been writing as if only one party had any threat capacity. There are of course cases of this (e.g., securely-hidden kidnappers) but more usual is the case where both parties have some capacity to threaten even if it is very unequal. What counts as a good shape of threat curve is not affected by whether or not the other party has an available threat, but what happens is liable to be affected. If A threatens B, B may either counter-threaten B or threaten retaliation. These are logically quite distinct: the first is a threat by B to damage A unless A does something advantageous to B; the second is a threat by B to damage A if and only if A carries out his threat following non-compliance by B. The range of outcomes opened up by these is different, though including common items such as the continuation of the status quo and both parties carrying out their threats;

*But just as important  
even as threat (may  
in theory) be symmetric*



and both possibilities have in common that they may in prospect deter A from threatening B at all. It is apparent that any attempt to deal adequately with the strategic situation arising from two parties with available threats would require quite extensive treatment. I shall not therefore pursue the matter any further here.

26. Further developments. Extension to more than two parties is too obvious a need to elaborate on. A development which seems to be essential, though it inevitably introduces additional complexity, is the explicit recognition that the cost of coercion should for many purposes be divided into short-run and long-run costs: coercion in a single case of non-compliance may be quite cheap provided the overheads have been incurred which are required to set up a coercive apparatus. (Banfield, in Political Influence, makes some moves in this direction.) Another complication, which can be introduced quite neatly within the existing framework but requires additions to the basic four-curve diagram, is the recognition of the significance of a "counterforce" strategy in certain situations: that is to say, an actor may find it advantageous to use resources to destroy the coercive capacity of another party rather than seeking to maximize his loss of utility. To end with, I should like to mention the development which seems to me the most interesting and potentially fruitful, though I feel I have so far only begun to find out how to deal with it. This involves getting away from the God's eye view of the curves embodied in the discussion so far and introducing into the heart of the analysis an explicit recognition that what is crucial is the perceptions by each actor of the utility schedules of the other actor. What is the relationship between uncertainty and conflict? Which sorts of misperceptions and in which directions increase or decrease the chances of threatening versus non-threatening, compliance versus non-compliance, carrying out a threat versus doing nothing, threatening retaliation

good

Harsanyi

Me: if ~~no~~  $\int$   
~~from~~ probabilistic  
 beliefs, no  
 problem use  
 mean usually?



