Coalition Formation and Political Outcomes:
A Critical Note

N.R. Miller, Baltimore

As many of the essays in this volume illustrate, two distinct questions have guided much theory and research concerning coalitions.

The first is the question of *coalition formation.* In a strategic environment in which a number of agents can or must act in concert to obtain preferred outcomes, which of the multiplicity of logically possible coalitions will actually form?

The second is the question of *outcome determination.* Which of all the feasible outcomes will actually be realized or chosen through collective action? For example, if the outcomes are alternative allocations of divisible private goods, how will the members of the coalition that actually forms divide the spoils of victory? If the feasible outcomes are alternative levels of non-divisible public goods, what policy will members of the coalition that actually forms enact and implement?

It might seem that an ideal theory, or the most successful empirical research, would provide answers to both questions simultaneously. But the argument of this note is that – if we suppose that the interacting agents are rational in the generally accepted sense of game theory, economic theory, and positive political theory – there is a fundamental tension between these two objectives. The more definitely we can answer the first question, the less definitely we can answer the second, and vice versa. No coherent rational-choice theory can provide interesting and persuasive answers to both questions simultaneously, because – I argue – to the extent that rational collective action predictably determines outcome realization, this determinacy is the consequence of the fact that any one of several competing coalitions might form; conversely, to the extent that the formation of any one coalition is uniquely favored, outcome realization will depend on bargaining within that coalition and will be, to some considerable extent, indeterminate.

I should point out here that I understand rational behavior "in the generally accepted sense of game theory, economic theory, and positive political theory" to imply that bargaining games are indeterminate with respect to outcome realization. Of course, if rationality postulates are strengthened – in a way not so generally accepted – a determinate equilibrium may be implied for bargaining games. Most notably, *Harsanyi* [1977], in a theory which subsumes and generalizes earlier work by *Zeuthen* [1930], *Nash* [1950, 1953], and *Shapley* [1953], has proposed a set of "strong rationality" postulates which give a determinate solution for any bargaining game. Yet

Harsanyi's theory, when generalized to any n-player game, says nothing about "coalition formation" as that term is typically understood (and thus his theory and my thesis really do not contradict one another). Harsanyi argues that, in the absence of some communications bias, every logically possible coalition forms simultaneously, each defending the common partial interests of its members in a bargaining game with the complementary coalition. This "principle of full-coalition formation" quite likely characterizes the broad and long-term operation of the political process in a pluralist society, but pretty clearly it does not characterize political action in the narrower and more short-term arenas typically studied in connection with coalition formation, e.g., nominating conventions, government formation in multi-party parliaments, the formation of electoral alignments, the formation of international alliances, etc.

Despite the fact that many investigators of coalition formation make a bow to game theory and suggest that their hypotheses are derived from – or are at least suggested by – game-theoretical considerations, cooperative game theory is fundamentally a theory of outcome determination, not a theory of coalition formation.

Game theory is the theory of outcome determination when outcomes depend on actions by two or more rational agents. In the original formulation by *von Neumann/Morgenstern* [1953], the feasible outcomes are "payoff vectors," i.e., allocations of "transferable utility," essentially distributions of some divisible private good – in political terms, money, patronage, pork-barrel benefits, or other "spoils." In more recent and probably more politically relevant formulations [e.g., *McKelvey/Ordeshook/Winer*], the relevant overcomes are points in a policy space over which agents have preferences, essentially levels of provision of public goods – in political terms, alternative programs, platforms, or policy packages.

Cooperative game theory is the theory of outcome determination when these rational agents can reliably concert their actions (by restricting their strategy selections) in any agreed upon manner. An *ideally cooperative* strategic environment is one in which agents can concert their actions unimpeded by any kinds of *transaction costs* (resulting from physical dispersion, incomplete knowledge, costly communications, investment in bargaining tactics, etc.). A *coalition* is any set of agents who can concert their actions; thus, under ideally cooperative conditions, any subset of agents is a possible coalition. An *effective coalition* is a set of agents who can realize some common interest (i.e., for one outcome over another) by concerting their actions in a particular way.

Outcome X *dominates* outcome Y if there is some set of agents who all prefer X to Y and who, acting in concert, have the power to make X the actual outcome, i.e., if there is an effective coalition for X over Y. Under fully cooperative conditions, an outcome that is dominated in this sense is clearly "unstable" and we would not expect it to be realized. Thus we are interested in the *core* of the game, i.e., the set of undominated or "stable" outcomes. The core may be large (include many outcomes), which creates a kind of indeterminacy that may be characterized as "social bargaining"; for this it is necessary that domination be incomplete. As domination becomes more complete (because intermediate sized coalitions become more powerful, or because preferences become more consensual), the core shrinks and outcome realization of the cooperative game accordingly becomes more determinate. But the

core can shrink too much, so that it is empty, i.e., so that every outcome is dominated by some other and every outcome is "unstable"; this creates a second kind of indeterminacy that may be characterized as "social instability"; for that it is necessary that domination be cyclical. [Cf. *Miller*.]

The important point to note here is that it is the increasing multiplicity of coalition possibilities (i.e., of effective coalitions) that causes the core to shrink and makes outcome realization more determinate. Likewise, if the core is empty, the multiplicity of coalition possibilities gives stability to certain sets of outcomes (e.g., a "von Neumann-Morgenstern solution," the "competitive solution," etc.). In any case, actual coalition formation is relatively unimportant – indeed, strictly speaking, coalition formation takes on no clear meaning within the context of game theory itself.

In their original treatise, *von Neumann/Morgenstern* [1953, p. 36] aptly characterized the role of multiple coalition possibilities in determining, or at least constraining (their concern was with essential *n*-person zero-sum games, for which the core is always empty), outcome realization.

> Any particular alliance describes only one particular consideration which enters the minds of the participants when they plan their behavior. Even if a particular alliance is ultimately formed, the division of the proceeds between the allies will be decisively influenced by the other alliances which each one might alternatively have entered.... Even if ... one particular alliance is actually formed, the others are present in a "virtual" existence: Although they have not materialized, they have contributed essentially to shaping and determining the actual reality.

Consider the simple example of majority-rule decision making on a single policy dimension with single-peaked preferences. The outcome that entails selection of the policy corresponding to the ideal point of the median voter is the unique element of the core; through many different effective coalitions all other possible outcomes are rendered "unstable." Thus, given fully cooperative conditions, a game theorist would predict that this median outcome would be realized. But he would have nothing to say about which voters might end up voting for this policy. Indeed, being recognized as inevitable, this policy might win on a unanimous vote. In any case, such a unanimous roll call would in no way contradict cooperative game theory, though some analyses of legislative roll calls, relying on *Riker's* [1962] "size principle," might so interpret such behavior.

As a related example, consider a parliament whose members – individuals or parties – are clearly arrayed over an ideological spectrum from left to right. A governing coalition must command majority support. By the same logic as in the direct voting example just above, a cooperative game theorist would predict formation of a governing coalition committed to enacting the platform preferred by the median member of the parliament. But he would have nothing to say about which governing coalition might form. Indeed, the fact that whichever governing coalition does form must agree on this policy depends precisely on the fact that there are so many other coalition possibilities.

In so far as research focuses on coalition formation – and empirical research on governing coalitions in multiparty parliaments typically has this focus (probably because coalition formation is easier to observe and measure than policy or patronage) – this theoretical indeterminacy poses problems. It is appealing, therefore,

to take account of certain "sociological," "extra-rational," or other contextual factors –
such as physical or ideological proximity, differential barriers to communication,
differential patterns of trust and antagonism, differential bargaining costs, etc. – that
plausibly influence coalition formation by making some effective coalitions more
"realizable" and others less so. In general, this means presuming that transaction costs
are greater than zero and unequal among different potential coalition partners. This
presumption may certainly be realistic, but it does involve departure from the
assumptions of ideally cooperative game theory. And it must also be recognized – and
this is the crucial point for the purposes of this note – that in so far as such factors
are deemed to be significant, ruling out the formation of some coalitions or at least
making them less likely, outcome realization becomes less determinate than it other-
wise would be.

Let us consider several examples.

First, consider a simple three-person cooperative game with transferable utility.
The characteristic function is:

$$v(1) = v(2) = v(3) = 0$$
$$v(1,2) = v(1,3) = v(2,3) = v(1,2,3) = 1.$$

The core of this game is empty but it possesses a "main simple solution," according to
which two players divide the unit spoils equally and the third gets nothing. But which
minimal winning coalition will actually form, i.e., which two players will get the equal
positive payoffs, is quite indeterminate. Suppose now that we erect differential barriers
to coalition formation, i.e., differential transaction costs of one sort or other, so as to
rule out, or at least make more costly, the formation of certain coalitions. While this
will make coalition formation more predictable, it will also change the expected
outcome of the game and generally make it less predictable. Suppose, for example, that
an extra-rational antagonism or quarrel exists between 2 and 3, ruling out, or making
very unlikely, formation of the coalition $(2,3)$ as well as the "grand coalition" of all
players. We now can predict coalition formation more reliably, i.e., only coalitions $(1,2)$
and $(1,3)$ may form. But the previously predicted equal division no longer obtains, as
player 1 can exploit the antagonism between 2 and 3 to his advantage. Indeed, if the
coalition $(2,3)$ is ruled out entirely, player 1 can appropriate (virtually) all of the unit
spoils for himself. Now suppose that player 3 is cut off from the others by physical
isolation, communications barriers, or whatever, ruling out, or making far more
difficult, formation of the coalitions $(1,3)$ and $(2,3)$ as well as the "grand coalition."
Coalition formation is now entirely predictable; coalition $(1,2)$ must form. But the
outcome is now entirely unpredictable. The absence of any alternative coalition
possibilities creates in effect an indeterminate bargaining game between players 1 and 2.

The analysis is much the same if we consider instead a three-person majority
game on a two-dimensional policy space, as shown in Figure 1. A game theorist would
predict that one of the three logically possible two-player coalitions would form and
agree to enact the policy Φ_i lying on its contract curve. (The set of policy outcomes Φ_1,
Φ_2, Φ_3 constitute the "competitive solution" to the game, as well as the "main simple
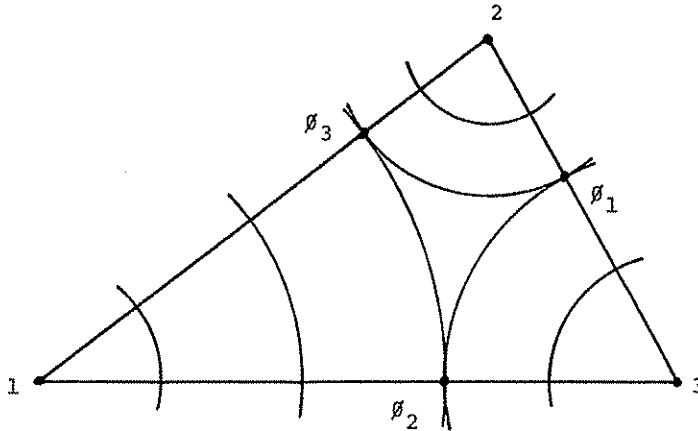solution" and the "bargaining set" – the three solution concepts converging in this

Fig. 1: A Three-Person Policy Game with Euclidean Preferences

simple case [cf. *McKelvey/Ordeshook/Winer*, especially pp. 603, 608].) But once again, if the coalition (2, 3) were ruled out, the policy outcome would approach player 1's ideal point. And if coalitions (1, 3) and (2, 3) are both ruled out, an indeterminate bargaining game on the contract curve of 1 and 2 results.

As a third example, let us briefly consider some empirical research reported by *Leiserson* [1968]. Leiserson uses a variant of the von Neumann-Morgenstern theory of cooperative simple games to interpret the process of coalition formation among factions (players) within the governing conservative party of Japan and to predict the consequent distribution of patronage. He uses the implication of this theory that only minimal winning coalitions will form and accordingly that only distributions supported by such coalitions will be realized. He then narrows down the set of minimal winning coalitions by applying the "bargaining proposition" that the minimal winning coalition with the fewest members will form (because, in effect, it has the lowest trans-action costs). Application of the "bargaining proposition" has two consequences. First, with respect to coalition formation, it renders the initial "minimal winning coalition proposition" redundant, since the winning coalition with fewest members is necessarily minimal [cf. *Taylor*]. Second and much more importantly, with respect to outcome determination (i.e., the distribution of patronage), in so far as the "bargaining proposition" is persuasive and coalition formation becomes more determinate, the patronage distribution associated with the predicted coalition becomes less determinate. That is, in so far as the formation of the predicted coalition is greatly favored, its members are freed from the constraints imposed by other coalition possibilities and can bargain over the distribution of patronage in a way that is more or less indeterminate. (It should be noted, however, that within the tolerance entailed by measurement problems, Leiserson's predictions regarding patronage distributions are quite well supported by the empirical results.)

As a fourth example, consider the familiar "Downsian" theory of two-party competition on a single ideological dimension. We can predict what platform will be

enacted but not what winning coalition will form (i.e., how citizens will vote and which party will win the election). This is because competition induces the two parties into offering identical (median) platforms, leaving all citizens indifferent as to how to vote. In so far as we introduce extra-rational sociological factors (such as ideological motivations on the part of party leaders and/or party loyalty on the part of citizens) in order to make "coalition formation" (between a party and a set of voters) more determinate, the policy outcome becomes less determinate. For example, if many voters are (within some range of tolerance) loyal to one or other party, the "majority party" (in terms of the distribution of party loyalty) can offer a range of different platforms all of which assure it of electoral victory.

As a last example (that stretches the notion of coalition formation somewhat), consider a free-exchange economy. Given a sufficient number of traders, it can be predicted that uniform (competitive) exchange ratios (prices) will emerge and a final allocation of goods can be predicted accordingly. Economists do not, however, attempt to predict what particular "coalitions" will form in equilibrium (i.e., how traders will pair off). Moreover, it is this very indeterminacy with respect to "coalition formation" (all appropriate pairings being essentially equally likely) that creates the uniformity of exchange ratios and determines the final allocation. If on the basis of certain extra-rational sociological considerations (e.g., proximity, friendships) we could predict which "coalitions" would form, the uniformity of exchange ratios and the consequent determinacy of the final allocation would be (in some measure) undetermined, as the exchange economy would be "factored" into a plurality of more or less independent (and indeterminate) bargaining games.

Finally, I should take account of one well-known theory that (typically) predicts, in a determinate way, both coalition formation and outcome realization. This is *Gamson's* [1961] "minimum resource theory" applied to simple games played for divisible spoils. If players possess different "resources" (e.g., political parties controlling parliamentary seats), it is predicted that the "cheapest winning coalition" (e.g., the majority coalition that exceeds 50% of the seats by the smallest amount) will form and divide the spoils of victory (e.g., cabinet portfolios) proportionately to their contributed resources. Thus, if three parties A, B, and C control 40%, 35%, and 25% of the seats respectively, it is predicted that the coalition BC will form and divide the spoils of victory 7/12 to 5/12 respectively. But this allocation is not really a prediction but an assumption; to quote *Gamson* [1961, p. 376]: "Any participant will expect others to demand from a coalition a share of the payoff proportional to the amount of resources which they contribute to a coalition." Given such expectations, A and B both prefer to join with C and C prefers to join with B, so the coalition BC forms. But what this example really shows is that the postulated expectations cannot be rationally sustained. Party A cannot demand a share proportionate to its resources, for the certain result of such a demand is to get nothing at all. Party A must scale down its demands to the point that it is as desirable as a coalition partner as the others. The end result is that any one of the three two-party coalitions may form, dividing the spoils equally between the two partners in exactly the manner of the first example discussed above [cf. *von Neumann/Morgenstern*, 226–227].

References

Gamson, W.A.: A Theory of Coalition Formation. American Sociological Review **26**, 1961, 373–382.

Harsanyi, J.C.: Rational Behavior and Bargaining Equilibrium in Games and Social Situations. Cambridge 1977.

Leiserson, M.: Factions and Coalitions in One-Party Japan: An Interpretation Based on the Theory of Games. American Political Science Review **62**, 1968, 770–787.

McKelvey, R.D., P.C. Ordeshook, and *M.D. Winer*: The Competitive Solution for N-Person Games without Transferable Utility, with an Application to Committee Games. American Political Science Review **72**, 1978, 599–615.

Miller, N.R.: Power in Game Forms. In: Power, Voting, and Voting Power. Ed. by M.J. Holler. Würzburg-Wien 1982.

Nash, J.F.: The Bargaining Problem. Econometrica **18**, 1950, 155–162.

–: Two-Person Cooperative Games. Econometrica **21**, 1953, 128–140.

Neumann, J.v., and *O. Morgenstern*: Theory of Games and Economic Behavior. 3rd ed. Princeton 1953.

Riker, W.H.: The Theory of Political Coalitions. New Haven 1962.

Shapley, L.S.: A Value for n-Person Games. In: Contributions to the Theory of Games, II. Ed. by H.W. Kuhn, and A.W. Tucker. Princeton 1953.

Taylor, M.: On the Theory of Government Coalition Formation. British Journal of Political Science **2**, 1972, 361–373.

Zeuthen, F.: Problems of Monopoly and Economic Warfare. London 1930.