Coalitions

Macartan Humphreys

Department of Political Science, Columbia University, New York, New York 10027;
email: mh2245@columbia.edu

Annu. Rev. Polit. Sci. 2008. 11:351–86

First published online as a Review in Advance on February 15, 2008

The *Annual Review of Political Science* is online at http://polisci.annualreviews.org

This article's doi:
10.1146/annurev.polisci.11.062206.091849

1094-2939/08/0615-0351$20.00

Key Words

commitments, cooperative, noncooperative, contracts, Nash program

Abstract

The game theoretic study of coalitions focuses on settings in which commitment technologies are available to allow groups to coordinate their actions. Analyses of such settings focus on two questions. First, what are the implications of the ability to make commitments and form coalitions for how games are played? Second, given that coalitions can form, which coalitions should we expect to see forming? I examine classic cooperative and new noncooperative game theoretic approaches to answering these questions. Classic approaches have focused especially on the first question and have produced powerful results. However, these approaches suffer from a number of weaknesses. New work attempts to address these shortcomings by modeling coalition formation as an explicitly noncooperative process. This new research reintroduces the problem of coalitional instability characteristic of cooperative approaches, but in a dynamic setting. Although in some settings, classic solutions are recovered, in others this new work shows that outcomes are highly sensitive, not only to bargaining protocols, but also to the forms of commitment that can be externally enforced. This point of variation is largely ignored in empirical research on coalition formation. I close by describing new agendas in coalitional analysis that are being opened up by this new approach.

INTRODUCTION

Coalitions seem to be inescapable in politics. Although much of contemporary political theory holds to the principle that individuals are the fundamental unit of analysis, in actual models the units are often coalitions: trading blocs, nations, states, governments, houses, ethnic groups, parties, classes, households.

In most instances, coalitions are givens; we simply assume that they exist and that they behave as unitary actors, bracketing the questions about how they persist and how actions are coordinated within them.

There are two instances, however, in which coalitional possibilities play a more prominent role in political analyses. The first arises when the possibility of coalitional action poses a threat to analysis that is based on individual maximization. Olson (1965) warns that it is naive to predict actions of individuals based on group interests in settings where agreements cannot be enforced. The converse is also true, however: It is a mistake to predict the actions of groups from the preferences of individuals in settings in which agreements *can* be enforced without duly taking account of that fact. This is what von Neumann & Morgenstern (1944) call "the crucial problem" in modeling games and economic behavior. The basic question is: Given that coalitions can form, what outcomes should we expect?

The second instance arises when groups themselves are an important subject of analysis. The concern here is with the creation and maintenance of group structures—the origins of nations (Alesina & Spolaore 1997), the formation of classes (Roemer 1982) or coalitions between classes (Rogowski 1989, Iversen & Soskice 2006), and the study of endogenous ethnicity (Posner 2005), migration (Tiebout 1956), and government formation (Laver & Shepsle 1996, Snyder et al. 2005). The question here is: Given that coalitions can form, which coalitions should we expect to see forming?

To address these questions we need some notion of what it means to be a coalition. This definitional question is surprisingly thorny and goes to the heart of the project of coalitional analysis. The basic difficulty is that it is insufficient to define a coalition as a group of individuals who act together or act somehow in each other's interests. As Krehbiel (1993) notes in the case of political parties, even if the actions of individuals in a group are correlated, that implies nothing about the agency of *groups*. Krehbiel suggests that evidence for the agency of political parties is found when individuals vote contrary to their preferences but consistent with the group's goals. This approach is compelling, but it leaves any theory of coalitions that purports to be based on methodological individualism and utility maximization in a quandary. On the one hand, if the seemingly distinctive actions of individuals in a coalition can be adequately accounted for simply through a standard examination of individual incentives, then there is no special place for coalitional analysis. On the other hand, if these actions cannot be accounted for by individual incentives, then it would seem that either the assumption of methodological individualism or the utility maximization assumption must be dropped.

Coalitional analysis could indeed be undertaken by dropping either of these assumptions. But for the purposes of this review, I take a middle approach and let a coalition denote a set of two or more agents that have engaged in an externally enforced agreement to undertake a set of actions subject to the terms of their agreement. To emphasize the external enforcement, I use the metaphor of players signing contracts. The insistence on external enforcement is important because "self-enforcing" agreements can be fully analyzed without reference to coalitions. As the enforcement technology is taken to be exogenous, coalitional analysis is in this sense always partial equilibrium analysis. For generality, the definition does not require that actions taken are necessarily in the interest of the group, nor that they are not in the immediate interests of the individuals concerned. The definition excludes some subjects of interest in the analysis of coalitions more broadly defined—such as self-enforcing coordination—but it has the merit of placing the analytic emphasis on the availability of commitment technologies, and, as I discuss below, variation in the types of commitments that are available to players is key to understanding the types of coalitional structures that we expect to observe. Using this definition, the two questions under study can be rephrased as follows: In situations in which individuals *can* make commitments (from some universe of enforceable commitments) to each other, how do we expect them to behave, and who do we expect to coordinate with whom?

The remainder of this review considers game theoretic approaches to answering these two questions. Traditionally, the main attempts to respond to the first question used the tools of cooperative game theory to identify outcomes that are in some sense immune to coalitional deviations. By assuming that gains from cooperation are indeed maximized in such settings, this approach has placed little emphasis on the second question of which groups are in fact likely to form in equilibrium. The next section motivates the problem through the examination of two games, Prisoners' Dilemma and Divide the Dollar. I then describe some of the major solution concepts that derive from a cooperative approach to studying such problems. These solutions, though developed without explicit reference to a particular game form, correspond to solutions that do arise from more fully specified noncooperative analyses.

Although it has significant appeal, cooperative game theory has been criticized by political scientists on a number of grounds. Many of these criticisms miss the mark, perhaps because of a lack of clarity regarding what modeling features are specific to the cooperative approach. But for some applications the cooperative approach does suffer from two shortcomings: its inability to take account of all strategically relevant information and a demonstrable sensitivity of negotiated outcomes to bargaining protocols.

Recent work uses noncooperative coalition theory to respond to some of these shortcomings and to address both questions of substantive interest: what coalitions will form and what strategy profile will ensue. A common framework for doing this is to add a noncooperative pregame in which coalitions are formed and then analyze the base game noncooperatively with coalitions as the unit of analysis. In such cases, players predict the outcomes of the game between coalitions and select coalitional membership on the basis of this. Some of these games use open and closed membership rules; others employ sequential game forms or dynamic processes that allow for continual coalitional evolution. Although it is well known that bargaining outcomes are sensitive to protocols, I highlight the extreme sensitivity of outcomes to details of the legal environment—that is, to implicit assumptions regarding not just how contracts are formed but what types of contracts can be formed. I examine the implications of one such assumption for models of government formation, the assumption that parties can commit to bargaining only with a subset of other parties. This assumption, though rarely if ever defended, is shown to have substantive import.

I close by recommending three fruitful directions for future work: a reassessment of the conditions for coalitional stability; a more explicit focus on the types of commitment devices available to potential coalition partners, and a better linking of these to applications; and an examination of coalitional possibilities in settings in which commitment devices are endogenous.

PD: Prisoners' Dilemma

DD: Divide the Dollar

MOTIVATING CASES

We can motivate the problem of coalition formation through a discussion of the Prisoners' Dilemma (PD) and Divide the Dollar

(DD) games. In each case, I note how the games change once coalitional possibilities are considered. Later, I return to these games when discussing different approaches to modeling coalitional processes.

A Prisoners' Dilemma with Contracts

Assume that two individuals each need to decide whether to provide evidence against the other for some joint crime. Payoffs for the different strategy combinations are as given in **Figure 1**.

This is a standard PD game, and its solution is well known. The inescapable fact is that rational, utility-maximizing players defect (they play *D*), even though they would both be better off if both kept quiet (played *C*). This is a sad result, and the sadness of it has provoked an enormous literature to find conditions under which rational individuals might instead cooperate. One approach is to ask how the prisoners would play if they could sign contracts with each other.

The traditional approach for answering this question is the cooperative approach. Importantly, the cooperative approach "solves" the PD not by bringing a different solution concept to the same game, but by solving a different game: one with contracting possibilities. To emphasize the point, I refer to the PD as the base game and the PD with contracts as the expanded game.

How does play in the expanded game differ from play in the base game? The answer depends on what kind of contracts can be signed. The universe of possible contracts is not well defined, a feature I return to later. I focus here, however, on three prominent types of contract that can be used to form expanded games.

- **Contracts over pure strategies**. One possibility is that contracts are signed with respect to pure strategies: "If both of us have signed this contract, then I agree to play *C*, and if I fail to do so, some terrible thing will happen to me." In the case where contracts can only be signed with respect to pure strategies, there are three efficient contracts, (*C*, *D*), (*D*, *C*), and (*C*, *C*), although one of these, (*C*, *C*), is individually rational in the sense of being an improvement over no contract for both players.
- **Contracts over correlated strategies**. Contracts could also invoke jointly mixed strategies of the following form: "If both of us have signed this contract, then I, Player *A*, will condition my strategy on some publicly observable random event *x*, which occurs with probability *q*. If *x* occurs, I will play *C* (and you will play *D*); otherwise, I will play *D* (and you will play *C*)." This family of contracts is larger and allows players to do better than they would if they only relied on pure strategies. In this case, the individually rational and efficient contracts can be parameterized by $\{q : q \in [\frac{1}{3}, \frac{2}{3}]\}$; and these different contracts correspond exactly to divisions of a dollar.

These two types of contract specify particular strategies to be played by players, but contracts need not be limited to this. A common expansion of the space of contracts is to allow contracts to stipulate side payments to be made between players.

- **Contracts with side payments**. If individuals can commit both to a particular course of action in the game and to bilateral financial transfers, then even if

		B: *C*	*B*: *D*
A	*C*	$\frac{1}{3}$, $\frac{1}{3}$	−1, 2
	D	2, −1	0, 0

Figure 1

A prisoners' dilemma.

contracts over correlated strategies are not feasible, another set of contracts can be implemented that returns the same values. In each of these, one player plays D while the other plays C; the player who plays D then makes a side payment of $1 + \alpha$ to the other, where $\alpha \in [0, 1]$ represents the division of the surplus. In this case (though not always), the set of possible payoffs from the game with transferable utility and pure strategies is equivalent to those from the game with nontransferable utility and correlated strategies.

In the case of the simple contract, there appears to be just one plausible candidate—all cooperate. This could be supported by many bargaining protocols. For example, we could imagine that player 1 makes a take-it-or-leave-it offer to player 2. If players sign, the contract is implemented; if not, then players play the PD (base game) noncooperatively. However, in the latter two cases, with multiple contracting possibilities, the contracting problem seems considerably more difficult, and the solution may plausibly depend on the type of bargaining allowed.

Finding a solution to such a game is of great theoretical interest for the following reason. We see that, for a special case, the PD with contracts reduces to a problem of dividing a dollar. It turns out, however, that this is not unique to this PD or even to two-player PDs more generally; rather, for a very large class of two-player games—excluding zero-sum games—it is possible to normalize utilities such that the set of efficient contracts corresponds precisely to the segment [0, 1]. The introduction of contracts thus vastly reduces the complexity of the universe of games that has to be analyzed. Thus, if we can find one reliable solution to this problem, then we have a solution for a vast array of games that can be characterized by the same simple utility space. That is the tremendous attraction of finding solutions to games with contracts: the promise of generality.

A Divide-the-Dollar Game with Contracts

Contracting becomes considerably more complicated once we move beyond two players. To see why, consider the following normal form (noncooperative) game. There are three players, $N = \{1, 2, 3\}$. The strategy space for each player is given by a two-dimensional unit simplex, Δ, with generic element $s_j \in \Delta$, representing the set of all possible three-way divisions of a dollar. If two or more people agree on an allocation, then each of the three players gets his share in the agreed-on allocation; otherwise, all receive 0.

What are the Nash equilibria of this game? Unlike the PD case, where the Nash equilibrium was unique, here there are many equilibria. Any profile in which all three players play the same strategy is a Nash equilibrium. Any profile in which two players make the same offers is also a Nash equilibrium, provided that the third does not offer more to either of the other two than they offer themselves. These are all efficient equilibria. There is also an inefficient equilibrium, in which each player proposes taking the full dollar for himself.

The Nash equilibrium solution concept clearly has difficulties making fine predictions for this game, since any division and no division can all arise as outcomes. In considering a closely related game, von Neumann & Morgenstern (1944, p. 223) concluded that "there seems to be no escape from the necessity of considering agreements concluded outside the game . . . if we do not allow them then it is hard to see what if anything will govern the conduct of a player." In other words, the authors suggest a form of coalitional analysis in this case not because contracting possibilities are likely to be part of player strategies, but rather in order to select a solution to the base game. In fact, however, it is not clear that allowing agreements solves the problem in this case. If we allow for the possibility that players can write binding contracts over the strategies that they will adopt, the question then becomes: What contracts will be written, with what resulting payoffs?

As with the PD example, we can envision a setting in which, prior to the base game described above, DD players can sign a set of contracts with each other. And as with the PD game, these contracts may specify explicitly or implicitly the actions that will be taken by all the signing partners in the base game and may or may not specify other actions such as the making of side payments. Again, as with the PD case, the set of efficient contracts corresponds exactly to divisions of a pie.[1]

In this case, however, matters are more complicated because players have to select not just the contract but also their contracting partners. The problem is further complicated by the fact that for any initial contract and resulting set of payoffs, some rival contract may be available to a set of players that would make them all better off. This raises new concerns: If a contract is offered, does the offer remain open while a player examines other offers? If a contract is signed, can it be superseded by another contract? With more than three players, the contracting possibilities become still more complex; in this case, the optimal actions of a contracting pair could reasonably depend on whether the other players sign contracts with each other, and what kinds of contracts they sign.

THE COOPERATIVE APPROACH

The classic game theoretic approach to addressing problems in which contracts can be signed uses cooperative game theory. In many accounts, cooperative game theory *is* the study of games in which contracts can be signed. This is misleading because noncooperative approaches can just as easily be used for games with contracts and commitments, as long as the strategies for signing contracts and making commitments are specified. Rather, cooperative game theory is an approach to solving games.

Cooperative Modeling Assumptions

The cooperative approach is characterized by two assumptions: a representation assumption and a solution assumption.

The representation assumption is that, in the presence of contract signing possibilities, the only strategically relevant information from the base game is the set of profiles of utilities that can be attained by coalitions of players, possibly as a function of the full set of coalitions that form. The important difference between different contracts is in the utilities they assign to different players, not in the details of the particular strategies that they specify to achieve these utilities. In practice, the representation assumption takes form in the use of a "characteristic function" (or one of its relatives) to characterize the base game.

The solution assumption is that general protocol-free outcomes should be the object of study. This assumption is logically distinct from the representation assumption. The idea is that the details through which contracts are developed are in practice likely to be unknowable, and predictions should therefore depend on what coalitions are capable of attaining by working together rather than on the particular procedures of bargaining that are used.

Cooperative solutions thus select payoff profiles and do not directly describe either the strategies that may be used to sign contracts or the strategies that are employed in the base game once contracts are (or are not) signed.

I describe the approach in detail for the classic case of "coalitional games with transferable utility" or "cooperative games with side-payments," although more general approaches exist.[2] For a population N of players, the characteristic function v assigns a worth to

[1]Note that the pie being divided in the expanded game is not the same as that being divided in the base game. To see the difference, imagine that players had utility functions of $u_i = ((s)_i)^2$. In this case, a 50/50 division between two players would yield payoffs of 0.25 each, whereas a contract that specified a 50% probability of a 100/0 or a 0/100 division would yield expected utility of 0.5 each.

[2]Two generalizations are worth emphasizing. First, the theory can allow for cases in which coalitions can jointly ensure some set of outcomes that may be valued differently by

each coalition, $v : 2^N \to \mathbb{R}^1$.[3] In practice in most cooperative work, v is a primitive. However, it should be possible to derive v from base games. To do this, we need to make two types of assumptions: effectivity assumptions about how much value a coalition can gain and transferability assumptions about how this worth can be divided among coalition members.

Effectivity assumptions are used to treat the worth of the coalition as the maximum amount that members of a coalition can "guarantee" for themselves (the amount for which the coalition is "effective") (Aumann 1961). As shown in our discussion of the PD, if players can collude, then the set of strategies that are available to them jointly is the corresponding set of correlated strategies. The question then is, given these strategy spaces, what can a coalition guarantee itself? Two notions of coalitional effectivity are prominent in the literature, corresponding to the minimax and maximin derivations of v. Intuitively we think of S and $N - S$ choosing strategies sequentially. S is α-effective for x if it can guarantee itself the payoffs of x even if it goes first. S is β-effective if it can guarantee itself these benefits if it goes last. Other alternatives are possible. For example, $v(S)$ may denote the sum of payoffs to members of S from Nash equilibrium play or from Nash bargaining between S and $N - S$. For Chandler & Tulkens (1995), S is γ-effective for x if x corresponds to the equilibrium payoffs to S in a game between S and the remaining players acting as singletons. The key feature to note here is that in order to *derive* a cooperative game from a base game, we need to already accept some suppositions about how coalitions are likely to react to each other.

TU: transferable utility

The transferability assumption is the assumption that utility can be directly transferred from one player to another. It can be defended without reference to interpersonal utility comparisons if players have sufficient access to a divisible and transferable good, y, that enters each player's utility linearly.

With these two assumptions in hand, a number of minor assumptions allow a useful normalization. First we assume that v is "superadditive," i.e., that $v(A \cup B) \geq v(A) + v(B)$. [In fact, superadditivity is a result of both the minimax (Luce & Raiffa 8.2) and maximin (Aumann 1961, Sec. 9) derivations of v.] The superadditivity assumption is based on the idea that whatever underlying strategies are played by A and B to generate $v(A) + v(B)$ can also be employed by A and B working in concert (although see Hyndman & Ray 2007 for a counterargument). It is considerably weaker than the assumption of "strict superadditivity" (the requirement that this inequality be strict), which is not satisfied, for example, in the DD game we considered above. Second, we assume that v is essential: $v(N) > \sum_{i \in N} v(i)$. Note that under superadditivity, essential games are the interesting ones from the perspective of potential for cooperation. For essential supperaditive games, we can always normalize utilities to define a characteristic function with $v(N) = 1$ and $v(i) = 0$ for all i. This is termed a $(0 - 1)$ normalization of the game v. Henceforth, we assume that all TU games are $(0 - 1)$-normalized.

Consider now the cooperative representation of the PD and DD games with contracting discussed above. Under the assumption that players play Nash in the case of bargaining failure, a characteristic function for the PD is given by $v : 2^{\{A,B\}} \to \mathbb{R}^1$, $v(C) = (|C| \geq 2)$. For the DD game, there are many Nash equilibria in the game between two-person coalitions and the remaining singleton, but they are all payoff-equivalent (for the coalitions): the coalition claims the pie. For singletons, we can consider the game between

different members without the possibility for side-payments among them [nontransferable utility (NTU) games]. See Aumann (1961) for a discussion of the core for NTU games and Maschler & Owen (1989) on the "consistent Shapley value" for NTU games. Second, greater generality can be achieved by replacing the characteristic function with a partition function. A game in "partition function form" (Thrall & Lucas 1963) specifies a payoff for every coalition, S, conditional on a particular partition, P, of N.

[3] Implicitly, single-member coalitions are included in this set.

themselves and some pair acting as a coalition, or else consider the three-player game in which an inefficient Nash equilibrium is played. In this case, each can expect a payoff of 0. This yields the 0–1 characteristic function: $v : 2^{\{1,2,3\}} \to \mathbb{R}^1$, $v(C) = (|C| \geq 2)$.

We now consider some of the prominent cooperative solution concepts for games of this form.

Cooperative Solutions

There exists a very wide class of solutions to games in cooperative form, with different approaches emphasizing different principles of bargaining behavior. Here I describe what are perhaps the three most prominent solutions—the Nash bargaining solution, the core, and the Shapley value—and show how each handles the PD and DD games with contracts.[4]

The Nash bargaining solution. The Nash bargaining solution, used for settings with or without transfers, is designed for what can be called (e.g., Hart & Mas-Colell 1996) "pure bargaining models," in which either a comprehensive agreement is reached or all bets are off. The solution is not designed to handle cases with partial breakdowns, where agreements are in fact made only between subgroups.

The Nash bargaining solution is defined as follows.

Definition 1. The Nash bargaining solution for a $(0 - 1)$ TU game is the payoff vector x that maximizes $\prod_{i \in N} x_i$ subject to $\sum_{i \in N} x_i = 1$.

Nash (1953) justified this solution using a particular noncooperative game in which players made threats in a context with a risk of bargaining failure. However, the real power of the result is that Nash also showed that the bargaining solution is the unique outcome from *any* bargaining procedure, including any noncooperative process, that is efficient, symmetric, and scale-invariant, and whose outcomes are "independent of irrelevant alternatives" (that is, if a bargaining solution picks some point z, and if some unchosen point x is removed from the set of acceptable solutions, then z is still chosen). This axiomatic foundation provides strong support to the bargaining solution. Scale invariance is presupposed in all games employing von Neumann–Morgenstern utility; Pareto efficiency is attained in a wide class of bargaining games; symmetry simply requires that all relevant features of the game are captured by the utility possibility set. The most difficult to defend is in fact independence of irrelevant alternatives (Rubinstein 1982), although even this arises naturally in many bargaining settings.

For the PD example, the Nash bargaining solution predicts an even split of the pie. For the DD problem, the Nash bargaining solution, using only information on coalitions of size 1 and 3, yields division $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. For the PD case, this solution is intuitive, and as we will see, it corresponds to noncooperative solutions to the same game. For the DD example, the solution seems less plausible—why would the players split the pie when any two could do better by excluding the third? The reason is that the Nash bargaining solution is appropriate only if a constrained class of contracts—comprehensive contracts—are allowed. This feature obtains in some multilateral bargaining settings (such as Diermeier & Merlo 2004), but not all.

The core. Perhaps the most prominent solution concept for cooperative games is the "core," defined by Gillies (1959). The basic idea can be traced back to von Neumann & Morgenstern's *Theory of Games*.

Definition 2. The core for a $(0 - 1)$ TU game is the set of payoff vectors $\{x\}$ for which $\sum_{i \in N} x_i = 1$ and $\sum_{i \in C} x_i \geq v(C)$ for all $C \subset N$.

[4]For a sampling of other cooperative solutions, see the top cycle set (Ward 1961), the bargaining set (Aumann & Maschler 1964), the nucleolus (Schmeidler 1969), the uncovered set (Miller 1980), and the stability set (Rubinstein 1980). For more general solution concepts that can be applied under conditions of cooperative or noncooperative behavior, see Greenberg (1980) and Chwe (1994).

Informally, a payoff vector is in the core if no coalition can, by acting alone, achieve an outcome that every member prefers to that vector. There is a clear rationale for the core concept in cases where players can sign contracts. The idea is closely analogous to the idea of Nash equilibrium except that, to be in the core, we require not just that no individual deviations benefit any individual but also that no group deviations benefit any group. In this regard, we may expect it to be more difficult to find points in the core than it is to find Nash equilibria. Unlike in the Nash equilibrium notion, however, we emphasize that the deviations from an outcome x by C in the definition of the core imply changes in coalitional structures and thus, possibly, changes in the strategies employed by all players rather than just those employed by C.[5]

Returning to our motivating examples, we find that whereas the Nash bargaining solution corresponded closely to the noncooperative solutions, the core solution looks very different for both of these examples. In the PD case, the core is set-valued: The entire set of (individually rational) efficient contracts is in the core, reflecting the fact that for each of these, some bargaining procedure could produce them in equilibrium. In all cases, however, the substantive outcome is the same: One player cooperates and the other does not, and a compensating side payment is made. For the DD game, the core is empty. No outcome is predicted at all, since for any candidate solution, an argument can be made for why that outcome would not obtain. The emptiness of the core, as I argue below, reflects a fundamental difficulty of generating a prediction for the substantive outcome of this game when noncomprehensive agreements can be signed.

The Shapley value. The Shapley value tries to relate the size of a share of the pie that accrues to an individual in a cooperative setting to that individual's contribution to the size of the pie. Again there is an underlying bargaining logic. Shapley asks whether any allocation has the following property: Whenever one individual, i, demands more of another, j, on the grounds that if i left the coalition j would lose some benefits, then j can object to i's demand on the grounds that if j left i would lose at least as much as j would if i left. The difficulty is in calculating the marginal contributions of different players, since these may depend on other players. Perhaps a player's contribution is large if some other player is in the group but small otherwise. If so, then perhaps the measurement of that other player's contribution should take account of his impact on the first player's contribution. And so on. The idea of the Shapley value (for cooperative TU games) is that a value can be associated with each individual if we average over the whole set of marginal contributions that she could make to every coalition that does not contain her already. Letting $\mathcal{R}$ denote the set of all permutations of N, with typical element R, and letting $S_i(R)$ denote the set of players that precede player i in the permutation R, the Shapley value is defined as follows.

Definition 3. The Shapley value for a (0 – 1) TU game is the payoff vector x with:

$$x_i(N, v) = \frac{1}{|N|!} \sum_{R \in \mathcal{R}} [v(S_i(R) \cup \{i\}) - v(S_i(R))].$$

With a little calculation, one can check that under the 0 – 1 normalization $x_i \in [0, 1]$ and $\sum x_i = 1$. The Shapley value can be uniquely derived using axioms of symmetry, efficiency, additivity [an axiom on the addition of subgames; if $w = v + u$, $x_i(N, w) = x_i(N, v) + x_i(N, u)$)], and the dummy axiom—if a player's actions never increase the utility possibilities of any set, then that player

[5]The core has been defined here for cooperative games. A closely analogous definition exists for many social choice settings; in this case, the core is a set of outcomes rather than a set of payoff profiles. Consider any preference aggregation rule f that can be represented by a set of "decisive coalitions" $\mathcal{D}(a, b)$, such that b is "socially preferred" to a if and only if there is a coalition, $C \in \mathcal{D}(a, b)$, whose members all strictly prefer b to a. Then a is in the core of f for a given preference profile if there is no b that is socially preferred to a. This is equivalent to the condition that there is no coalition that can and wants to induce a change from a to b.

gains no share of the surplus (Osborne & Rubinstein 1994, 14.4.2).

There are a number of known relations between the Shapley value and the core. One is that if a game is convex (in the sense that the larger the coalitions are, the larger the marginal contributions of one coalition to another are), then the core is nonempty and *x* is in the core (see, e.g., Mas-Colell et al. 1995, 18.AA.1).

In both the PD and the DD games we have been considering, the Shapley value corresponds to the Nash bargaining solution.

Hybrid Solutions

There are also solutions that, though cooperative in spirit, are closely associated with noncooperative solutions. Two of these are the strong equilibrium and coalition-proof Nash equilibrium.

Strong Nash equilibrium. The "strong equilibrium" solution concept, due to Aumann (1959), lies right at the intersection between cooperative and noncooperative approaches. It is defined for explicitly noncooperative games and acts as a refinement of Nash equilibrium that excludes Nash equilibria from which some *joint* deviation is beneficial to some coalition. In terms of the discussion above, the notion drops the representation assumption of cooperative approaches but maintains the solution assumption. We can see immediately from the definition that if an equilibrium is strong, then it is a Nash equilibrium and it is also Pareto-efficient. These properties indicate how strong the equilibrium concept is. In fact it is so strong that it often fails to exist. A strong equilibrium does not exist for either the PD or the DD examples studied above. Although sometimes related to the core, the two notions are distinct, as can be seen from the PD example. The core exists in the PD we studied above because it is assumed that if the two individuals do not form a coalition, then they play the Nash equilibrium. Strong equilibrium, however, assumes that if one player deviates from the coalition, then the other's strategy is unaffected.

Coalition-proof Nash equilibrium. Bernheim et al.'s (1987) notion of coalition-proof Nash equilibrium also drops the representation assumption and substantially weakens the solution assumption. They also drop the assumption invoked in the previous solutions that agreements are binding and examine only the possibility that (unmodeled) communication between players can give rise to self-enforcing agreements. Thus, the coalitions they consider do not in fact satisfy the definition of coalitions we have been using so far. Unlike strong equilibrium, Bernheim et al. consider a group deviation to be admissible if and only if it is not itself subject to deviations by subgroups of the deviating coalition (for a treatment of linkages between the two concepts, see Konishi et al. 1997). All strong equilibria are coalition-proof, but not all coalition-proof equilibria are strong. Like strong equilibrium, coalition-proof equilibrium can be thought of as a refinement of Nash equilibrium; and like strong equilibrium, coalition-proof Nash equilibrium may often fail to exist.

In the PD game we have been examining, all-defect is a coalition-proof equilibrium even though it is not strong. The reason is that although the individual prisoners may well agree not to talk, once they are separated they have incentives to ignore the agreement. This is the first solution we have seen that predicts all-defect as an outcome in the PD—a feature that follows directly from the inadmissibility of externally binding agreements. However, in the DD example, *none* of the original Nash equilibria are coalition-proof. Consider any Nash equilibrium in which player *A* receives a positive amount; then consider a discussion in which players *B* and *C* agree on a deviation that splits *A*'s share positively among them. Then both players *B* and *C* have an incentive to implement this deviation, even in the absence of formal contracts. This result again points to the fragility of the Nash equilibrium solution for distributive games.

Cooperative-Noncooperative Links

A common concern is that cooperative solutions, such as those described above, are not "supported" by a well-defined game that leads to these outcomes through standard processes of individual utility maximization. In response, a large body of research seeks explicitly to characterize the linkages between cooperative and noncooperative concepts. The main aim of the so-called Nash program is to find families of noncooperative games whose outcomes correspond to major cooperative solutions.[6] This extensive literature has found links between noncooperative games and a wide class of cooperative solutions.

Nash bargaining solution. Nash (1953) provided his own contract-signing game that implemented his solution with a refinement of Nash equilibrium. In addition, perhaps the most prominent noncooperative bargaining model, the alternating-offers model due to Ståhl (1972) and Rubinstein (1982), yields the Nash bargaining solution as the payoffs from subgame perfect equilibrium in the limit as δ tends to unity (Binmore et al. 1986, Binmore 1987). Krishna & Serrano (1996) provide a related result for n-person games.

The core. Bergin & Duggan (1999) demonstrate that the core (whether defined with respect to α- or β-effectivity) is subgame perfect Nash implementable (see also Serrano & Vohra 1994 for subgame perfect Nash implementation of the core and Maskin 1999 for Nash implementation). Perry & Reny (1994) model a situation in which bargaining takes place in continuous time; proposers propose divisions subject to ratification by concerned parties. If a contract is accepted, the contracting parties leave the game. The game continues among remaining players. Perry & Reny demonstrate that an allocation can be supported as a stationary subgame perfect equilibrium outcome only if it is in the core. Importantly, in this game, if the core is empty then a stationary subgame perfect equilibrium does not exist. In this case, the emptiness of a core predicts the absence of an efficient stationary subgame perfect equilibrium. Chatterjee et al. (1993) provide a model of n-person coalitional bargaining with sequential offers, transferable utility, and time discounting. Focusing on strictly supperadditive games, they find that efficient stationary equilibrium payoffs converge to a point in the core, as δ converges to 1.

There are also strong results for simple games of the form commonly of interest to political scientists. For one-dimensional spaces and noncollegial rules, Banks & Duggan (2000) show that if players are perfectly patient and have strictly concave utility, then outcomes associated with (no-delay, stationary) equilibria are always in the core (and if players are nearly perfectly patient, outcomes are close to the core). Thus, for example, with simple majority rule, the median (core) is implementable in stationary strategies through a standard bargaining protocol (see also Jackson & Moselle 2002). Indeed, Banks & Duggan (2000) provide conditions under which *every* core point can be selected in equilibrium. In addition, recent work that better incorporates the dynamic nature of policy choice and change demonstrates the stability properties of the core in a noncooperative setting. Banks & Duggan (2006) show for a general bargaining model that if a status quo is in the core then it will stay there; indeed, a "static" stationary equilibrium exists if and only if the status quo is in the core. This result holds with no assumptions on discount rates. Surprisingly, however, the core, though retentive, is not attractive. Banks & Duggan (2006) also show that if a status quo is not in the core in a one-dimensional bargaining game, then no legislator will ever propose it. Nevertheless, consistent with Banks & Duggan (2000), with sufficiently patient legislators, outcomes are arbitrarily close to the core.

These results demonstrate the usefulness of the core when it is nonempty. Can anything

[6]See Bergin & Duggan (1999) for a treatment of different approaches to the Nash program.

be learned about the outcome of noncooperative games when the core is empty? Maskin 2003 shows that, in a class of superadditive games in partition form, core emptiness is a necessary condition for a grand coalition not to form. Below, I describe more results that relate core emptiness to the absence of absorbing states in dynamic processes.

The Shapley value. Bargaining models that produce the Shapley value include those of Harsanyi (1981), Hart & Moore (1990), Winter (1994), and Gul (1989). One simple protocol described by Hart & Mas-Colell (1996) is the following. A proposer is randomly drawn to propose a division of the pie. If the proposal is accepted by all players, it is implemented; if not, then with some probability $q < 1$, the proposer is removed from the game and receives a payoff of 0, and bargaining continues among the remaining players over the value obtainable by them. This protocol has a unique subgame perfect solution, which corresponds to the Shapley value for the TU case (for pure bargaining problems it implements the Nash solution). Even though this procedure implements the Shapley value, it does not readily correspond to a likely contracting procedure in many circumstances. For example, in a majority-rule DD game, as studied above, there is no reason to expect that players should seek unanimity in contracts that they propose. Gul's (1989) analysis is perhaps more convincing in this respect. In Gul's model, there is a pairwise meeting technology. When a pair meets, one player offers to buy out the other, making a transfer in exchange for the individual's resources (for example, her vote). If the offer is rejected, the union dissolves; if accepted, the seller leaves the market and the buyer continues to negotiate with remaining players. Expected payoffs from efficient stationary subgame perfect solutions converge to the Shapley values as players become patient. This provides a strong noncooperative rationale for the Shapley value, although, as noted by Hart & Mas-Colell (1996), the result holds only for efficient solutions, not all solutions.

Other work, not covered here owing to space limitations, has found noncooperative rationales for a range of other cooperative solutions, including the Kalai-Smorodinsky solution (Haake 2000, Moulin 1984, Trockel 1999) and the uncovered set (Coughlan & le Breton 1999). The model studied by Hart & Mas-Colell (1996), described above, implements the Maschler-Owen (1989) solution for nontransferable utility games. The model has been extended by Vidal-Puga (2005) to allow for preexisting coalitional structures.

Criticisms of the Cooperative Approach

The cooperative approach has been criticized in recent writing in political theory and, increasingly, published political theory relies on noncooperative methods. The criticisms of the cooperative approach have, however, been diffuse. A clearer sense of the strengths and weaknesses of the approach is important for understanding when the predictions of cooperative approaches should or should not be invoked. I consider five prominent criticisms to which cooperative theorists can readily respond, followed by two criticisms that I believe are more fundamental.

1. *Cooperative approaches select outcomes based on normative rather than positive criteria*. This critique is more salient for some solutions, such as the Shapley value or the Nash bargaining solution, than for others. The core, for example, has little to offer as a normative solution. In addition, as seen above, a success of the Nash program has been its ability to demonstrate how normatively desirable outcomes can arise from decentralized bargaining.
2. *Cooperative solutions do not exist (enough)*. Diermeier & Krehbiel (2003) suggest that the emptiness of the core for many political applications is a reason not to use it. Their rationale emanates in part from a desire to use a single solution concept for all purposes. One response

is that existence is neither general of, nor unique to, cooperative solutions; many cooperative solutions generally exist (or are nonempty), whereas Nash equilibria can fail to exist for reasons as simple as not having a closed strategy set. A deeper response is that multiple solution concepts are not inconsistent with theoretical coherence, and the emptiness or nonexistence of some can itself contain relevant information.[7] In the case of the core, emptiness reflects a feature of the underlying strategic situation (Hart & Kurz 1983). Thus, in the party platform selection problem, a game has a (noncooperative) Nash equilibrium if and only if there is a nonempty core (Duggan 2005); and, as shown below, in many games with nonempty cores it is easy to construct plausible dynamic noncooperative games that exhibit cycles.

3. *Cooperative solutions do not predict strategies and cannot incorporate institutions.* Diermeier & Krehbiel (2003) criticize the core for saying nothing about strategies. They argue that noncooperative solutions predict behavioral and outcome-related predictions, whereas cooperative solutions only predict outcomes. Garrett & Tsebelis (2001, p. 100) argue that their "most important critique" of power indices follows from the fact that "cooperative game theoretic assumption of enforceable agreements is institution free." Both critiques lose much of their force, however, whenever the characteristic function is derived from a normal form representation. In such cases, although strategies are not typically modeled explicitly, they are implicit, as described above.[8] In the same way, institutional features are not absent; rather, they are folded into the definition of cooperative games.
4. *We do not observe the cycling predicted by cooperative theories*. Paradoxically, although cooperative theory allows for binding commitments in many political settings in which the core is empty, cooperative solutions are associated with instability (see, e.g., Johnson & Libecap 2003, Shepsle 1986). In fact, however, core emptiness is simply the absence of a prediction not the prediction of unstable outcomes. At a deeper level, however, as demonstrated below, in some settings there is a relationship between core emptiness and cycling: When contracts are of short duration and the core is empty, cycles can arise in equilibrium and are predicted by dynamic noncooperative models (see Konishi & Ray 2003, Kalandrakis 2004). I illustrate this point with a discussion of the Laver-Shepsle model of government formation in the "Markov Processes and Spot Contracts" section below.
5. *It is unrealistic to assume that players can arbitrarily make binding commitments to each other.* There are two responses to this criticism. The first is that such binding agreements are not unique to cooperative approaches and are endemic in noncooperative approaches studying related phenomena. In such cases the bindingness is often hidden in the game tree, which simply declares that if such-and-such a set of actions are taken the game ends, or the agreement is implemented, and so on. Even the ability to make a take-it-or-leave-it offer (as for example in the model of Baron & Diermeier 2001) involves an implicit commitment technology; indeed,

[7]The use of a solution concept for a given setting reflects the information available to the theorist. Thus, if all that can be imposed on beliefs is common knowledge of player rationality, then the appropriate solution corresponds to the set of rationalizable solutions, not the set of Nash equilibria.

[8]But, as noted above, cooperative solutions will not distinguish between strategy profiles that produce the same outcomes.

LOSS OF STRATEGICALLY RELEVANT INFORMATION

Consider a game between players A and B in which A has no options and B can choose between L and R. Payoffs are given by $u_A(L) = 0$, $u_A(R) = 1$, $u_B(L) = -1$, $u_B(R) = 0$. Using the standard minimax derivation, the characteristic function is given by $v(A) = v(B) = 0$, $v(A, B) = 1$. The Shapley and Nash solutions suggest an even split; yet clearly if A refused to bargain, the Nash equilibrium would award the entire pie to him. Why then would he bargain? The problem here is subtle and hangs on a point of law: As part of contracting, can B make a unilateral commitment? Can B credibly commit to play L in the event of bargaining breakdown and thereby induce A to bargain? If the answer is no, then the characteristic function loses important information.

in some noncooperative models (such as that of Diermeier & Merlo 2000), parties can make cash payments to parties outside their coalition in exchange for a commitment to support future actions of the government. The second response is that this concern is about scope conditions, not about the basic approach. Cooperative solutions are motivated by the idea that binding commitments can be made between arbitrary groups of players. Yet cooperative solutions are employed in settings in which this assumption cannot be readily defended. If this assumption is not justified, then it is not legitimate to ignore the details of institutional procedures that give rise to social outcomes unless noncooperative foundations that include such features yield the same outcomes (as, for example, was claimed for the case of the uncovered set by McKelvey 1986).

There remain, however, two deep and longstanding concerns with the cooperative approach.

1. *The characteristic function disregards strategically relevant information.* The characteristic function greatly simplifies the representation of a game and in doing so may in some cases lose strategically relevant information. The most obvious problem is that payoffs to coalitions are assumed to be independent of the ways that remaining individuals form into coalitions. Thus, positive and negative externalities across coalitions are excluded. This problem can be resolved by replacing the characteristic function with a partition function, described below.

Other problems persist, however, and arise even in the context of two-player games. An illustration—based on one provided by McKinsey (1952) and referenced by Luce & Raiffa (1957)—of the inconsistencies that may arise from the minimax derivation of a characteristic function is given in the sidebar.

A reasonable response to such problems is that Nash equilibria payoffs should be used to assign worths instead of minimax (or maximin). But, as a general principle, the use of Nash equilibrium to predict play following contracting presents difficulties for games with multiple equilibria. In such cases, it may be possible to associate a cooperative solution for each cooperative game that corresponds to a given Nash equilibrium. But even this does not capture some of the subtleties that may arise. For an indication of some of the more complex reasoning that may arise in the case of multiple equilibria, consider a symmetric Battle of the Sexes game, with a $0 - 1$ representation as given in **Figure 2**.

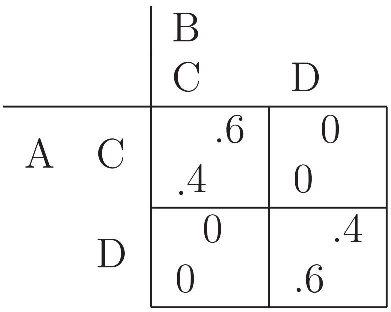


Figure 2

A Battle of the Sexes game. Refusal to sign a contract could signal expectations about the ultimate equilibrium that will be selected.

In this case, the mixed-strategy noncooperative solution is Pareto-inefficient. Bargaining, were it possible, could permit players to reach one of the efficient equilibria or even to randomize between them. In this case, as in the PD that we examined, many cooperative solutions will assign 0.5 each. However, following the logic of forward induction, a reasonable inference following a player's refusal to negotiate is that the player expects an equilibrium in his favor to result in a Nash equilibrium that awards a share greater than 0.5.

2. *Protocols matter*. The solution assumption made by cooperative theorists is that general bargaining results, if they are attainable at all, can be obtained without specifying precise bargaining protocols. Insofar as it implies greater generality, the lack of a detailed description of the game form that implements a solution is a distinct advantage of the cooperative approach, not a weakness. Furthermore, consistent with the Nash program, many articles proposing cooperative theories included noncooperative games that supported these solutions (Nash 1953, Harsanyi 1974, Maskin, 2003). However, the fact that some bargaining protocol implements a given solution does not mean that all protocols do, or that all relevant protocols do. If different plausible contract-signing protocols produce different outcomes (as indeed they do), then cooperative results, despite the advantage of weak assumptions, do not cover the requisite field.

NONCOOPERATIVE COALITION THEORY

New research on coalitions has responded to these concerns by weakening both the representation assumption (characteristic functions are typically replaced with partition functions, or in some cases more general contracts over players' actions) and the solution assumption (explicitly noncooperative approaches are used to examine contract selection). Unlike classic approaches, in most cases, new noncooperative coalition theory focuses not just on outcomes but on studying which coalitional structures (see sidebar) will form in equilibrium.[9]

> ### COALITIONAL STRUCTURES
>
> A coalitional structure is a partitioning of the set of players into a collection of groups, $\Gamma = (G_j)_{j=1}^{m}$, such that everybody is in some group ($\cup_{j=1}^{m} G_j = N$) and no one is in two groups ($G_j \cap G_k = \emptyset$ for $j \neq k$). Typically, we assume that a partitioning is defined uniquely up to the reordering of cells and of individuals within cells (thus {{1,2},{3,4}} is considered the same partition as {{4,3},{1,2}}). Let the set of all such partitions be given by $\mathcal{G}$. For a population of size n, the cardinality of $\mathcal{G}$ is then given by the Bell number (Rota 1964), a number that grows very rapidly in n. The set of partitions for $n = 4$ is 15. For a 100-person senate, the Bell number is on the order of 10^{116} (Weisstein 2005).

As in the discussion of the PD and DD games above, the approach used in much recent work assumes that prior (or in some cases, in parallel) to playing some base game, a coalitional structure forms, which determines in some manner how individuals will play the base game. There exist a range of ways of linking coalitional structures to the ultimate outcomes and payoffs for individuals. The simplest, used for example by Bloch (1996), is to define a utility function for each individual over the universe of partitions, $u_i : \mathcal{G} \to \mathbb{R}^1$ for $i \in N$. Implicitly, once coalitions are formed there is some conjecture about play between coalitions and a corresponding allocation of benefits to individual coalition members. The idea is that, once formed, coalitions play the underlying game with each other, and each coalition maximizes some function reflecting the welfare of its members, subject, perhaps, to some agreed-on rule for internal distribution of benefits. The most common alternative is to associate with each partition a worth for each coalition and allow the distribution of the value and the partition to be determined

[9]For older work emphasizing this objective, see Shenoy (1979), Hart & Kurz (1983) and Aumann & Myerson (1988).

jointly; in this case we have a partition function given by $v_C : \{\Gamma \in \mathcal{G} | C \in \Gamma\} \rightarrow \mathbb{R}^1$ for $C \in 2^N$ and contracts that specify adherence to the group's strategy that returns v_C. More general forms can be considered, however. In principle, it is possible that the coalition formation stage specifies a set of actions that will be taken by coalition members as a function of what other coalitions form, or perhaps as a function of the set of contracts that are signed by other agents.

For concreteness, in the discussion that follows I consider this question for a situation in which the base game is a Baron & Ferejohn–style majority rule bargaining game involving seven players. Coalitional features enter in two ways: through commitments made among players in some, or through the institutions of the base game. We assume that in the base game, each nondummy coalition has an equal probability of being recognized to propose an allocation of the pie. The expected payoffs for each coalition (of patient players), for each possible coalitional structure, are given in **Table 1**, which is based on values provided in **Table 1** of Snyder et al. (2005). My table also reports the expected average payoff for individuals within a coalition.

The question then is, knowing what returns can be achieved by coalitions, what coalitions are likely to form? Or, for the running example we will study, what party structures should we observe in a distributive-politics setting under different commitment mechanisms? The answer depends sharply on what protocols are used and what types of commitment are permitted.

Simultaneous Games

Open membership. In the simplest approach, players simultaneously select a coalition to join. Yi & Shin (1994) propose a model in which all players simultaneously announce a message, and those who announce the same message form a coalition. Utilities are assigned as a function of the ensuing partition, perhaps determined by a subgame played

Table 1 Numeric partitions of seven players

	(1) Numeric partition	(2) MIW[a]	(3) Expected coalition payoff[b]	(4) Expected per capita payoff
a	{1,1,1,1,1,1,1}	{1,1,1,1,1,1,1}	{0.14},{0.14},{0.14},{0.14},{0.14},{0.14},{0.14}	{0.14},{0.14},{0.14},{0.14},{0.14},{0.14},{0.14}
b	{2,1,1,1,1,1}	{2,1,1,1,1,1}	{0.28},{0.14},{0.14},{0.14}, {0.14},{0.14}	{0.14},{0.14},{0.14},{0.14}, {0.14},{0.14}
c	{3,1,1,1,1}	{3,1,1,1,1}	{0.42},{0.14},{0.14},{0.14},{0.14}	{0.14},{0.14},{0.14},{0.14},{0.14}
d	{2,2,1,1,1}	{2,2,1,1,1}	{0.28},{0.28},{0.14},{0.14},{0.14}	{0.14},{0.14},{0.14},{0.14},{0.14}
e	{3,2,1,1}	{2,1,1,1}	{0.4},{0.2},{0.2},{0.2}	{0.13},{0.10},{0.2},{0.2}
f	{4,1,1,1}	{1,0,0,0}	{1},{0},{0},{0}	{0.25},{0},{0},{0}
g	{4,2,1}	{1,0,0}	{1},{0},{0}	{0.25},{0},{0}
h	{2,2,2,1}	{1,1,1,0}	{0.33},{0.33},{0.33},{0}	{0.17},{0.17},{0.17},{0}
i	{3,3,1}	{1,1,1}	{0.33},{0.33},{0.33}	{0.11},{0.11},{0.33}
j	{3,2,2}	{1,1,1}	{0.33},{0.33},{0.33}	{0.11},{0.17},{0.17}
k	{4,3}	{1,0}	{1},{0},{0}	{0.25},{0},{0}
l	{5,2}	{1,0}	{1},{0}	{0.2},{0}
m	{5,1,1}	{1,0,0}	{1},{0},{0}	{0.2},{0},{0}
n	{6,1}	{1,0}	{1},{0}	{0.17},{0}
o	{7}	{1}	{1}	{0.14}

[a]Column 2 provides minimum integer weights (MIW) under simple majority rule.
[b]The expected returns for each coalition (Column 3) are from Snyder et al. (2005).

within the coalition. If mixed strategies are allowed, then equilibrium existence follows from Nash's theorem. Under such a rule, however, there may be a multiplicity of equilibria. Yi (1997) gives a monotonicity condition (ceteris paribus, a member of a smaller coalition benefits from a shift to a larger coalition) that ensures that the grand coalition is the unique equilibrium.

One feature of such open membership rules is that implicitly players cannot exclude others from their coalition. They may choose to leave a coalition they do not like, but they cannot keep other players out (or in). Thus, in the seven-person DD example, we can construct an open-rule game in which players who enter the same coalition receive an equal per capita share of the worth of the coalition in the coalitional structure. This game clearly does not satisfy Yi's monotonicity condition, and a multiplicity of equilibria exist. The set of (Nash) equilibrium coalitional structures is given by the concentrated structure (*o*) and the fragmented structures (*a* and *b*). For every other structure, some player has an incentive to deviate in order to select another group.[10] In this case, all equilibria are payoff-equivalent, because players who receive a subproportionate share of the pie have an incentive to join a group in which they would get at least a proportionate share, and, importantly, cannot be prevented from doing so.

Eguia (2007) studied a model of this form for cases in which coalitions coordinate on voting over binary agendas over which players have 0/1 utility. For more general policy-making settings, Caplin & Nalebuff (1997) studied a related model that takes the following form. There is a fixed number of communities (coalitions), $M = \{1, 2, \ldots m\}$; each community j implements a policy from some space X_j (note that this space may be distinct for each community). Let $\mathcal{G}$ denote a family of partitionings of N into these M communities. A policy vector is a choice of policy for each community and thus is an element of $X \equiv \times_{j\in M} X_j$. Individuals are distributed according to a hyperdiffuse distribution μ over a measure space of types, A, and have preferences over the group to which they belong and the pertinent policy outcomes that are captured by their type. So $u_i : M \times X \times A \to \mathbb{R}$.

Once individuals are grouped into communities, they play a game to select an outcome, and then some political process determines the political outcomes P_j in each community according to $P_j : \mathcal{G} \to X_j$. The within-group political process is not described in any detail in this model, but it is assumed that some political decision is reached in each community and that this decision is continuous in the population of types—i.e., if there is some small change in the population of types, then there will be a correspondingly small change in the policy selected by the community. In this setting, a social equilibrium is a partition Γ with a corresponding policy outcome $(P_j(\Gamma))_{j=1}^{m}$, such that, given the partition, each player prefers to be in the group he is in to being in any other group. A relatively simple proof establishes equilibrium existence in the two-community case with an odd number of dimensions. Gomberg (2004) has also found sufficient conditions for existence for an arbitrary number of dimensions. However, Gomberg (2005) shows that without further constraints on P_j or u_i, an equilibrium can fail to exist in any number of dimensions if there are just three communities.

Closed membership. In an alternative model (game Γ) studied by von Neumann & Morgenstern (1944) and Hart & Kurz (1983), a player's strategy set is the set of coalitions of which he is a member. A coalition is formed if all members of the coalition choose it. Note that embedded in the structure of the game is the requirement that an individual can be part of a given group only if all other members

[10]Define the binary operator $\to$ where $b \to a$ means that some player in coalition structure b has an incentive to make a deviation that induces coalition structure a. We then have the following sets of relations over states:

$$\left.\begin{array}{l} i \to e \to c \to f \qquad \to m \\ d \to h \to j \to \left.\begin{array}{l} g \\ k \end{array}\right\} \to l \end{array}\right\} \to n \to o.$$

consent to his membership. Thus, unlike in the open-membership case, a deviation can be ruled out if the deviation involves a change by a player to a coalition in which he is not welcome.

In the DD game we have been considering, game Γ produces a wider class of equilibria. The same extreme structures are again equilibria, but so are the cases with majority coalitions. In game Γ, a coalitional structure may be sustained as an equilibrium as long as no one person could benefit from destroying the coalition of which he is a member. All but states *e*, *j*, and *i* are equilibria, since in each case, members of the trios would prefer the payoffs they can expect from states *b*, *d*, and *c*, respectively.

In game Δ, a related game also studied by Hart & Kurz (1983), a coalition is formed of all members who announce a given coalition, whether or not all members of that coalition announce it. Thus, if 1 and 2 announce {1,2,3}, while 3 and 4 announce {3,4} the ensuing structure is {{1},{2},{3,4}} under Γ and {{1,2},{3,4}} under Δ. Again, a stable coalitional structure is a coalitional structure that emerges from a Nash equilibrium in the coalition-selection game. The interpretation in this case is that players make an open offer to a set of players with whom they are willing to work, and that offer can be taken up by any subgroup of players to whom the offer has been extended. The key implication with respect to coalitional stability is that, if a player leaves a coalition, that coalition may continue without him. In this case, all but two states are equilibria; unlike in game Γ, structure *j* continues to be an equilibrium because a split by a member of a trio would induce state *h* (not state *d*), which would leave the splitter in the worst possible position.

The merit of each of these coalition-formation games depends in part on the plausibility of the commitment mechanisms they presuppose. The open rule depends on a commitment technology that allows players to commit to hypothetical coalitions without knowing the exact make-up of the coalition; under game Γ the implicit commitment is to a well-defined set, whereas under game Δ the commitment is to a nested set of coalitions. In all cases, the commitment is assumed to be independent of the choices of members in other coalitions. More recent studies have modeled coalition formation as a more dynamic process. This approach provides a more natural framework for studying more complex acceptance and rejection rules and coalitional choices that depend on the coalitional choices of other players.

Sequential Contracting Games

The multiplicity of equilibria in the games described follows in part from the simultaneity of moves. Subsequent work has focused on the more realistic case of sequential contracting.

Chatterjee et al. (1993) provide an early model in which bargaining is built on the characteristic function of a TU cooperative game. A protocol determines an ordering of proposers. Each proposer selects a coalition, *S*, and makes an offer to the members of the coalition; the proposer is committed to that offer until all respondents have accepted or rejected it. If an offer is accepted by all members of *S*, the players in *S* leave the game, and bargaining continues among remaining players. If an offer is rejected by any member of *S*, the rejector becomes the next proposer. Discounting occurs between rounds.

Chatterjee et al. (1993) examine the efficiency of (stationary subgame perfect Nash) equilibrium for all protocols. They find that for equilibrium to be efficient for all protocols, we need the condition that for all *S*, $\frac{v(S)}{|S|} \leq \frac{v(N)}{|N|}$. This condition is satisfied, for example, by the PD game with contracts but not by the DD game with contracts. They find that any efficient stationary equilibrium of a strictly superadditive game is in the core; conversely, if the core is empty, then every stationary equilibrium is inefficient. These results provide another strong link between the core and the efficiency of noncooperative equilibria; however, these games feature no guarantee that efficient equilibria will exist, and indeed in some cases every protocol returns inefficient results even if there is a core.

Subsequent work in this tradition has moved from games built on characteristic functions to games built on partition functions. Two models stand out.

The first, studied by Ray & Vohra (1999), employs a protocol of the following form (see also Bloch 1996 for a closely related model). The game starts with the random recognition of a proposer. The proposer offers to a coalition a contract that specifies an allocation to each coalition member as a function of the ultimate coalitional structure that forms, subject to budget balance. We might think of the contract as specifying the equilibrium strategies that will be played by coalition members for each coalitional structure, together with a redistribution among coalition members of the value attained when those strategies are played. Once such a contract is signed, the signers leave the game, and contracting continues among the remainder until all contracts are signed. If a contract is rejected, the rejector can propose an alternative contract. Once all contracting is complete, players play the base game.

Ray & Vohra (1999) derive their strongest results for symmetric games and show that, for any partition function, there is an associated numeric partition (a partitioning that is defined uniquely up to a relabeling of players) that results from equilibrium proposal making. In addition, they provide conditions under which no-delay equilibria obtain.

For the seven-person DD game we have been examining, the solution under this protocol has the first player offer an equal division of the pie to four players. Each of these accepts and the game ends. Thus, outcome f is selected, although the payoff-equivalent structures g and k could also be sustained.

In the second sequential model, studied by Gomes (2005; see also Gomes 1999), a random player is selected who can propose a contract to some set of players that includes a set of actions and an (unrestricted) set of transfers. Once a contract is proposed, all members of the invoked coalition respond with acceptance or rejection following some fixed protocol. If the signatories of the contract accept the contract, the coalition of these players is formed. Unlike in the Ray & Vohra model, however, these players remain in the game and may, by mutual consent, renegotiate and replace their contract with another contract including more players (conditional on the consent of all players who are having their contracts rewritten). Also in contrast to the Ray & Vohra model, the game is dynamic, with players playing the base game in each period given the contracts in force (I discuss further studies along these lines in the next section). The space of contracts is the set of utilities that are awarded to each player in a coalition, conditional on the coalitional structure and the partition function and subject to budget balance within the coalition. Gomes (2005) shows that if the grand coalition is efficient, it is reached through contracting in finite time. If the grand coalition is inefficient, however, then in games with positive externalities there are bargaining delays, and inefficiencies arise even as player patience approaches unity. In these cases, players delay proposing contracts, hoping to free-ride on the cooperative behavior of others.

Gomes' results are consistent with new findings by Hyndman & Ray (2007) for characteristic function games with (long-term) binding agreements, finite states, and pure strategies. They show under general conditions that, for arbitrary time-dependent strategies, every equilibrium path converges to an efficient outcome. Moreover, under grand-coalition superadditivity (but not for general partition functions), they show that every absorbing payoff limit (steady-state payoff profile) of every Markovian equilibrium is static efficient. These results provide broad support for the Coase theorem in settings in which long-term contracts can be signed subject to renegotiation in cases in which the grand coalition is efficient.

Application of sequential contract formation models to public goods production. To see the logic of the Ray & Vohra solution, consider the following public goods problem (Ray & Vohra 2001). Each $i \in N = \{1, 2, \ldots, n\}$ must choose a level of some public good to produce. Payoffs are given by $u_i = \sum_{j=1}^{n} z_j - \frac{1}{2}z_i^2$. The classic Nash equilibrium solution yields $z_i = 1$ for all i, and so $\sum_{j=1}^{n} z_j = n$ and $u_i = n - \frac{1}{2}$.

What strategies would be selected if coalitions maximized the sum of the welfare of their members? For concreteness, we consider the case with $n = 4$. For any group of size m, it is easy to see that a utility welfare maximizer that controls the strategies of individuals in the group would maximize $m\sum_{j=1}^{m} z_j - \frac{1}{2}\sum_{j=1}^{m} z_j^2$, yielding $z_i = m$ for all i.

Thus, assuming that coalitions maximize the sum of their members' utility but coalitions play noncooperatively with each other, the average payoffs to members of a coalition conditional on a coalitional structure are as given in **Table 2**.

The question then is: What coalitional structure and set of strategies should we expect to observe in a setting in which contracts can be signed? The Nash equilibrium of the game without contracts produces $\{3\frac{1}{2},3\frac{1}{2},3\frac{1}{2},3\frac{1}{2}\}$, but this is not a very compelling solution, since, as in the PD game, any pair could do better by signing a bilateral contract. The grand coalition can achieve payoff vector {8,8,8,8}. However, though efficient, this vector is not in the core. Indeed, the highest payoff that the least well-paid player in the grand coalition can achieve is 8. But any player recieving 8 has an incentive to leave the grand coalition for a payoff of $9\frac{1}{2}$. Further splits or aggregations would not, however, produce net gains for all members concerned. This suggests that, though inefficient, {{1},{2,3,4}} may nonetheless be a plausible equilibrium structure.

This is what Ray & Vohra find. Specifically, they show that for all δ above some threshold, there exists a unique equilibrium coalitional structure. The equilibrium emerges when the first offerer commmits to join no coalition and remains on his own. The second offerer then proposes a coalition of the remaining three players.

In this example, one player free-rides on the remaining three. Ray & Vohra (2001) show the conditions under which such inefficiencies arise. Strikingly, for the game considered here, inefficiency depends sharply on n. In particular, the unique equilibrium is efficient for a sequence of population sizes from the set $n \in \{1,2,3,5,8,13,20,\ldots\}$. In all other cases, including the example given above, the outcome is inefficient. There are, however, upper bounds on the inefficiencies that can result; in particular, Ray & Vohra show that the number of equilibrium coalitions is less than $\log_2 n + 1$. For example, with one million players, no more than 20 coalitions form in equilibrium.

The Ray & Vohra model shows that allowing for side payments does not prevent inefficiencies. However, as the authors note, the inefficient equilibrium appears to depend on the stipulation that players cannot renegotiate. This prohibition rules out the possibility, for example, that a second offerer would propose a contract to all four players, including the player already committed, making an offer of the form {10,10,6,6}. For the public goods game we have been examining, this stipulation appears especially limiting because the "contract" to be renegotiated involves only one

Table 2 Payoffs from different (numeric) coalitional structures in a public goods game

Coalitional structure	Strategies	Payoffs
{{1},{2},{3},{4}}	{1,1,1,1}	$\{3\frac{1}{2},3\frac{1}{2},3\frac{1}{2},3\frac{1}{2}\}$
{{1,2},{3},{4}}	{2,2,1,1}	$\{4,4,5\frac{1}{2},5\frac{1}{2}\}$
{{1,2},{3,4}}	{2,2,2,2}	{6,6,6,6}
{{1},{2,3,4}}	{1,3,3,3}	$\{9\frac{1}{2},5\frac{1}{2},5\frac{1}{2},5\frac{1}{2}\}$
{{1,2,3,4}}	{4,4,4,4}	{8,8,8,8}

player. It is a general commitment that the player cannot renege on, even though doing so would make all better off.[11]

However, as we saw in our discussion of Gomes (2005), in settings in which renegotiation is possible, these inefficiencies do not arise. This is not to say that we can ignore the possibility of the kinds of inefficiencies identified by Ray & Vohra, but rather that the inefficiency prediction turns on details of the contracting, in particular on the possibility for renegotiation.

Application of sequential contract formation models to government formation. Consider now an application of Gomes' results to a government-formation problem in which there is a convex set of policy outcomes, X, and legislators have quasilinear utility of the form $u_i = y_i + v_i(x)$, where y_i is the legislator's holding of a transferable good and $v_i : X \to \mathbb{R}^1$ is strictly concave. In this case, efficiency is achieved by choosing the (unique) policy outcome that maximizes $\sum v_i(x)$ and redistributing income across the grand coalition. Under the commitment mechanism studied by Gomes (2005), the grand coalition should form a government and implement this policy in finite time. This is a classic Coasian solution. However, it runs contrary to the common prediction of minimal winning coalitions (Riker 1962) and to a more recent prediction of undersized coalitions (Diermeier & Merlo 2000).

The difference between the prediction from Gomes' model and the more standard predictions comes down to a subtle difference in commitment technologies. The difference is not whether commitments may be made but *which* commitments can be made. Consider the following example. Assume, consistent with Diermeier & Merlo (2000), that player $i \in N = \{1,2,3\}$ has preferences over points in $\mathbb{R}^3$ given by $v_i(x) = 2x_i - \sum_{k=1}^{3} x_k^2$. The efficient policy is then given by $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. Let the status quo be given by player 1's ideal, ({1,0,0}), with corresponding status quo utilities of (1,−1,−1).

Assume that player 2 is deterministically selected to form a government. What government will form? If the formateur is free to propose a coalition and a set of transfers, as in Gomes' model, we should expect player 2 simply to propose a policy such as $(0, \frac{1}{2}, \frac{1}{2})$ together with a transfer of $\frac{3}{2}$ from 3 to 2 for net utilities of $(-\frac{1}{2}, 2, -1)$. This option is not attractive, however, in the Diermeier & Merlo (2000) model because in that model the decision to enter government and the decision to split the pie are separated. If 2 enters government with 3, then he can expect policy $(0, \frac{1}{2}, \frac{1}{2})$ but no net transfers between the players for utilities of only $(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Instead, Diermeier & Merlo (2000) predict that 2 will elect to form a government on his own, implementing policy (0,1,0) and receiving no transfers for net utilities of (−1,1,−1). Thus, Diermeier & Merlo (2000) predict a minority government, whereas in the Gomes model a majority government is formed. As noted, however, the Gomes model allows for renegotiation subject to the consent of relevant parties, and in this case this leads to still larger coalitions. In our present example, if player 2 were to offer an amended contract, he could propose a government policy of $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ with further net transfers of $(-\frac{5}{6}, \frac{4}{6}, \frac{1}{6})$ and net utilities of $(-\frac{1}{2}, 2\frac{1}{2}, -1)$.[12]

Markov Processes and Spot Contracts

For a given set of membership rules, it is possible to construct dynamic processes in which

[11] See Diamantoudi & Xue (2007) for a setting in which inefficiencies may obtain even with renegotiation of contracts and transfers between players.

[12] This example is illustrative only. I have not taken account here of how expectations of a second-round amendment will affect first-round offers. In fact, if player 2 knew he was a monopoly formateur, he could extract considerably more from player 1 by implementing even less desirable policies in the first round.

EXAMPLE OF A SIMPLE MARKOV PROCESS

Consider the following process of coalition formation in the seven-player DD examined above. At each stage, an individual is selected at random. The individual can then choose either to change groups, conditional on unanimous acceptance by all members of the receiving group (if there are any), or to nominate one player for expulsion from his own group, conditional on the consent of all players other than the victim. If these changes occur, then the coalitional structure changes accordingly; in either case, period payoffs are allocated according to the prevailing structure. Payoffs are given by the average payoff of a player's coalition, and side payments are not allowed. If players are shortsighted and drift occurs with positive probability, that is, weak preference is sufficient to produce a change of groups, then state $h = \{2,2,2,1\}$ is the unique stable state; moreover, it is accessible from every other state. Other processes are directly or indirectly dominated by h, either through expulsion ($\overset{E}{\to}$), strict preferences ($\overset{S}{\to}$), or weak preferences ($\overset{W}{\to}$), as follows.

$$\left.\begin{array}{l} \left.\begin{array}{l} a \overset{W}{\to} b \overset{W}{\to} c \overset{S}{\to} f \overset{W}{\to} g \\ o \overset{E}{\to} n \overset{E}{\to} m \overset{W}{\to} l \overset{E}{\to} g \end{array}\right\} \overset{W}{\to} k \overset{S}{\to} \quad i \overset{S}{\to} e \\ \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad j \overset{S}{\to} e \end{array}\right\} \overset{S}{\to} d \overset{S}{\to} h$$

Note that states involving minimum winning coalitions are not stable, since expulsion or drift may result in coalitional structure $\{4,3\}$. But from this state, a member of the minimum winning coalition has an incentive to defect to create the three-party system $\{3,3,1\}$.

the coalitional partition at each stage is a function of the partition function at the previous stage and the optimizing actions of individuals. In such processes, payoffs can be received in "real time" (Konishi & Ray 2006, Gomes 2005, Hyndman & Ray 2007). In many settings, this structure is more realistic, although it also has the advantage of providing a simple environment in which players can form conjectures about future (equilibrium) play.

The general structure allows for many different ways to model the evolution of group structures. Single-step processes allow for moves between coalitional structures that require change in group membership of one person only. More general processes allow for the wholesale change of coalitional structure from period to period. In such cases, the stability of a state provides a natural equilibrium notion; although different notions of stability obtain for stochastic processes, the simplest is that a state is stable if it is "absorbing," in the sense that once in this state the process will remain in this state. A simple example of such a process for the seven-player DD game is shown in the sidebar.

Laver & Benoit (2003) study a Markov process of coalition formation in which states are partitions and each member of a coalition (party) receives his per capita share of the Shapley value of the partition. As in the example considered above, moves are assumed to occur if they provide immediate benefits to both the mover and the receiving coalition, although in Laver & Benoit's examination, expulsions are not permitted. As in the sidebar's example of a simple Markov process, it is implicitly assumed that players are shortsighted and that once members join coalitions they make two commitments: to vote with the party and to receive a per capita share of the sum of allocations made to their coalition. Laver & Benoit find that although multiple equilibria exist, larger dominant parties are in general more likely than smaller parties to be both attractive to and accepting of new members, suggesting a tendency of evolution toward a single-party monopoly coalitional structure.

Konishi & Ray (2003) examine a more general process in which a (finite) set of states can be interpreted either as partitions or as strategy profiles. In their model, players are farsighted and select moves as a function of the expected value given equilibrium strategies of all other players at each state. Agreements in this model should be thought of as spot contracts—binding, but only for a single period of play. An equilibrium process of coalition formation (EPCF) is one in which, if a move is made between states, this is always to the (long-term expected) benefit of some coalition, and if a strictly profitable move is possible for some coalition from some state, then the state is guaranteed to change and a strictly profitable move will be made with positive probability. Konishi & Ray demonstrate that such EPCFs exist for finite (or countably infinite) state spaces. Moreover, focusing on characteristic functions, they show that for sufficiently patient (or sufficiently impatient) players, if a deterministic EPCF has a unique limit then that limit is a (weak) core point. Further, if an EPCF is absorbing and deterministic then the absorbing states are in the "largest consistent set" (Chwe 1994). The fact that core outcomes emerge in settings in which players are farsighted is especially important because the standard justification for the core, focusing on single deviations, implicitly assumes shortsighted players.

Gomes & Jehiel (2005) study finite state spaces, which can be interpreted as outcomes or as coalitional structures, in a related game. In each period, a player can propose a change of states alongside an (unrestricted) set of transfers to other players (in contrast to the model of Konishi & Ray (2003), where proposals are made by fictitious players and no transfers are allowed). They examine the different implications of long-term and spot contracts. With long-term contracts, they find that efficiency obtains, consistent with the Coase theorem. A striking result is that although core points are consistent with equilibria with myopic players, they are not necessarily equilibria with farsighted players. Gomes & Jehiel provide an extraordinary example in which there are four states, a, b, c, d, with payoff profiles at each state for each of three players given by $v(a) = (60, 40, 0)$, $v(b) = (40, 0, 60)$, $v(c) = (0, 60, 40)$, $v(d) = (64, 64, 64)$. A move from any state to any other state can be made with the consent of a majority. Thus, for each of states a, b, and c, there is a division of the dollar, and so with transferable utilities we may expect the kind of cycling described by Kalandrakis (2004, 2006b). However, unlike in the standard DD game, there is a state d that is a core point in the sense that there is no move away from d that produces immediate benefits to any pair, even if transfers among them are allowed. Gomes & Jehiel (2005) show, however, that with sufficient patience, players can elect to move away from d, incurring short-term losses in the expectation of longer-run gains. The key insight is that if selected, player 1 for example could propose a shift to state a in the expectation that once in state a, players 1 and 2 will both be in strong bargaining positions should they have the opportunity to propose an alternative state to the now suffering player 3. The key weakness of state d is that it does not satisfy the property Gomes & Jehiel term "negative externality–freeness"—that is, from state d it is possible for a coalition to induce another state that hurts a non–coalition partner. Strikingly, Gomes & Jehiel show that if no state exists that is negative externality–free, then the existence of a stable state cannot be guaranteed—there exist some proposal probabilities that induce cycles and, possibly, inefficiencies.

Application of Markov models to the Laver-Shepsle government formation model. As an application of these results to government formation, consider the Laver-Shepsle model of government formation (Laver & Shepsle 1996). In each period, a party is selected to propose a coalition, that is, an allocation of government portfolios to political parties. It is assumed that all parties have some positive probability of being selected. Once a government is proposed, it can be vetoed by any of its members. If it is not vetoed, it is subjected to a majority vote. If the motion carries, the government takes office. If the proposal is either vetoed or defeated on the floor, the status quo government remains in place. In either case, play moves to the next period. Once a government takes office, policy outcomes are selected via a noncooperative game in which each party chooses the optimal policy on the dimensions it controls. Note that the implicit contracting is very limited. When a coalition is voted in, players accept the policy choices of the coalition members; but players are not permitted to sign arbitrary contracts, a feature that can result in the selection of Pareto-suboptimal policies in equilibrium. The equilibrium concept Laver & Shepsle invoke is stationary subgame perfect Nash equilibrium.

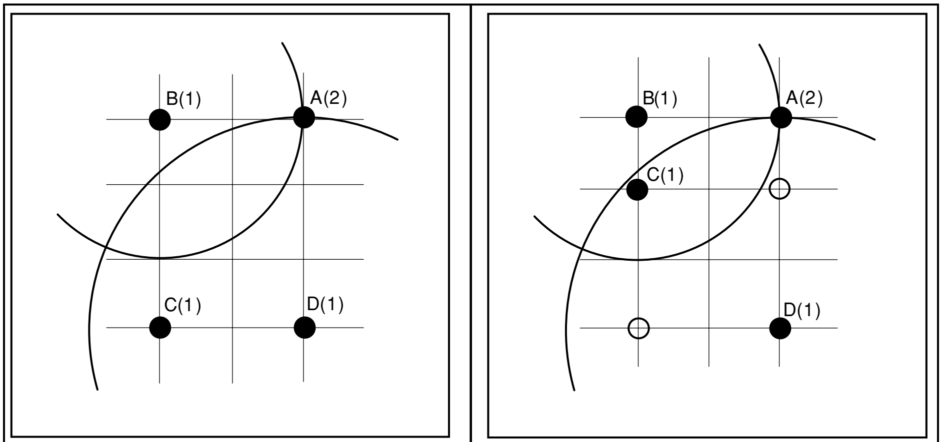
The Laver-Shepsle model of coalition formation is unambiguously noncooperative. However, it has been described in the literature as a non-noncooperative, perhaps cooperative, perhaps structure-induced-equilibrium social choice theoretic model (Diermeier 2006). The reason for this likely relates to the way the model is solved. Although Laver & Shepsle describe the solution concept as being stationary subgame perfect Nash equilibrium, in fact they solve for states, not strategy profiles. Of course, strategies and states are closely linked, and so Laver & Shepsle could respond that "equilibrium states" is shorthand for "states that obtain when players play their subgame perfect Nash equilibrium." There is a second and perhaps more difficult problem, however: Although Laver & Shepsle provide a game form, they do not provide the universe of player preferences that combine with the game form to produce a fully specified game. In particular, they do not provide information regarding how streams of utility are measured. In the absence of this information, Laver & Shepsle leave some parts of the game unsolved (such as which players win a standoff); but more worryingly, the absence of this information makes it difficult to know whether the solutions proposed indeed correspond to subgame perfect equilibria.

In some readings (e.g., Diermeier 2006), the solution provided by Laver & Shepsle is in fact the core and not subgame perfection. In consequence, Diermeier (2006, p. 167) argues that when the core of the (single-period) game is empty, as sometimes it is, "the Laver-Shepsle theory has no empirical content; it does not predict anything." In fact the issue is more subtle. For cases in which the core is empty, Laver & Shepsle indeed do not predict anything. But this is not because of the structure of their model or the solution concept employed; it is because they simply do not fully identify equilibrium strategies for all preference profiles including those with an empty core. Once sufficient utility information is provided, the model can yield an empirical prediction resulting from stationary subgame perfection even in those cases when the core is empty; indeed, as we will see below, in such cases it can predict government instability.

To see this, consider a version of the model in which each player is perfectly impatient, placing value only on the policies produced by governments in the present period, and in which policy is defined over two policy dimensions. Let ij denote a government in which the horizontal dimension is controlled by i and the vertical by j. The state space is the set of governments. We now consider two examples (**Figure 3**). In each case there are four parties, A, B, C, D, with voting weights given by 2,1,1,1, respectively. The first case corresponds to a situation in which there is a very strong party; the second case, based on one provided by Austen-Smith & Banks (1990), corresponds to a case in which the core is empty.[13]

We identify Markov perfect equilibria in which, in each period, a party is selected to propose a government. This proposal is then subjected to a veto and a sequential vote and, if accepted, becomes the next status quo. Stage utilities are then awarded, and play moves to the next period. Using backward induction, in equilibrium proposers propose their most preferred cabinet among those that will not be vetoed or invested; if there is no such cabinet that they prefer to the status quo, they pass. Vetoers veto cabinets if they prefer the status quo to the proposal, and all parties vote in favor of a coalition if and only if they prefer this proposal to the status quo. We assume that the status quo always lies on the lattice. Given this behavior, the equilibrium proposals for the first case are provided in the lower panels of **Figure 3** as a function of the status

[13] For simplicity we examine cases with nongeneric ideals, but the main result does not depend on this.



θ	$\sigma_A(\theta)$	$\sigma_B(\theta)$	$\sigma_C(\theta)$	$\sigma_D(\theta)$
AA, DA, AB,DB	SQ	SQ	SQ	SQ
BB, BA, CA,CB	AA	SQ	SQ	AA
CC, BD, BC, CD	AA	BB	SQ	DD
DD, DC, AC,AD	AA	BB	SQ	SQ

θ	$\sigma_A(\theta)$	$\sigma_B(\theta)$	$\sigma_C(\theta)$	$\sigma_D(\theta)$
AA,DA,DB, AB	SQ	CC	CC	CC
BB, BA, CA,CB	AA	SQ	SQ	AA
CC,BC	BB	BB	SQ	SQ
DD,AD	AA	BB	CC	SQ
AC,DC	AA	AA	SQ	SQ
BD,CD	AA	BB	CC	DD

Figure 3

The upper part of each panel shows a collection of four ideal points and possible cabinets from the Laver-Shepsle model. The voting weight of each group is given in parentheses. Black dots signify ideal points and white dots signify other possible coalition positions. The lower panel shows stationary proposal strategies for each player conditional on a given status quo government.

quo (θ). For the first case, it is easy to check that there is a single absorbing state at *AA*. Indeed, given equal proposal probabilities, the probability that *AA* is reached within 10 rounds is well above 99% independent of the status quo. With a fixed ordering, state *AA* is guaranteed within four rounds. This is an example of a case in which the unique absorbing state is a core point. In addition, the outcome does not depend precisely on features such as recognition probabilities. We note, however, that from Theorem 7 of Gomes & Jehiel (2005), if transfers can be made between players then we are not guaranteed that this state (or any state) is an equilibrium of this game for perfectly patient players for all proposal probabilities. The reason is that players *B*, *C*, and *D* can induce a move away from *AA* that hurts *A*. Such a move is beneficial in the short run for *B* and *C* but harmful for *D*. Player *D* might nevertheless support such a move if he could subsequently extract a payment from *A* to return to *AA*.

In contrast, in the second case, in which the core is absent, an equilibrium cycle can occur with the state passing from *AA* to *CC* to *BB* to *AA* over time. In this case, the asymptotic distribution does depend on recognition probabilities; with equal probabilities, state *AA* occurs with frequency $\frac{1}{4}$ and states *CC* and *BB* each occur with probability $\frac{3}{8}$.

Coalitions and Commitments

We have seen that the types of commitments that are permissible have strong implications

GOVERNMENT FORMATION: EMPIRICAL REGULARITIES AND THEORETICAL PREDICTIONS

Recent work on government formation has been motivated in part by two empirical regularities (Ansolabehere et al. 2005). First, the share of the pie (cabinet positions) assigned to a party in government is proportionate to its seat share (Gamson 1961); second, there is little advantage in being a formateur. Depending on the protocols used, both regularities in the empirical literature on coalition formation can be supported (or rejected) by theoretical models.

As shown by Baron & Ferejohn (1989), when there are at least five (patient) players, any distribution can be sustained as an equilibrium (though not as a stationary equilibrium), no matter who proposes first. In addition, Kalandrakis (2006a) shows that for monotonic voter rules, including rules with vetoers and dictators, any (expected) payoff can result from a stationary equilibrium of distributive games for an appropriate selection of recognition probabilities. For the DD game considered above, any profile of expected payoffs $\{\alpha_i\}_{i\in N}$ with $\alpha_1 \geq \alpha_2 \geq \alpha_3$ can be implemented through an alternating-offers contracting game with patient players and recognition probabilities given by $(\frac{\alpha_1}{\alpha_1+\alpha_2}, \frac{\alpha_2}{\alpha_1+\alpha_2}, 0)$.

In many models that build on the Baron & Ferejohn (1989) approach with random recognition rules, the formateur enjoys a premium (Snyder et al. 2005). The formateur's advantage disappears, however, with patient players under a protocol in which the rejector of a proposal becomes the next offerer (Ray & Vohra 1999). In Morelli's (1999) model of demand competition, the formateur advantage is also absent under some conditions (see also Montero & Vidal-Puga 2007).

for the types of coalitional structures we may expect. **Table 3** highlights the extreme dependence of coalitional predictions on the specific commitment technologies that are deemed feasible. There is no coalition (party) structure that is predicted under all the commitment technologies we have examined, and only two ({3,2,1,1} and {3,3,1}) are not predicted by *some* model of contract writing.

I close this section with a discussion of the implications of one salient feature of commitment technology for models of government formation: the ability to commit to negotiate only with a specified set of other players.

The politics of government formation has been a very active area of research in political science during the past 20 years. Much recent work ignores policy positions and treats the problem of coalition formation as a DD problem (e.g., Snyder et al. 2005).[14]

When formulated as a DD game, theoretical predictions clearly depend on the bargaining protocols used (see sidebar). But the predictions also depend on the manner in which commitment possibilities are modeled. To see how, we reconsider the DD game under two treatments. In the first ("majoritarian") treatment, commitments are invoked only upon the signing of the final deal. A proposal is made by a randomly selected player on the floor of the parliament; if it is accepted by a majority, the division is implemented. If rejected, another player is randomly selected. In the case with three (patient) players, $\frac{1}{3}$ is offered and $\frac{2}{3}$ is retained; each player's ex ante take is $\frac{1}{3}$.

Now consider a second ("coalitional") treatment. Before going onto the floor of parliament, players can elect to form a coalition (or in some treatments, a "proto-coalition"). In forming a coalition, players can commit to only support motions that they can jointly agree on. But they cannot simultaneously commit to a particular motion. That is, they agree to reach a substantive agreement with each other but do not (yet) agree on what that substantive agreement is. In this case, when selecting partners, individuals calculate what they expect to receive from bargaining with each partner. Given the simple symmetry of

[14]Important exceptions stand out. Calvert & Dietz (2005) consider externalities in preferences; Jackson & Moselle (2002) and Morelli (1999) allow both ideological positions and pork allocations. Baron & Diermeier (2001) allow both types of preferences, although, since they assume quasilinearity and efficient bargaining, within-coalition bargaining takes place over divisions of a dollar only. For Strom et al. (1994), ideological differences constrain coalitions; for Laver & Shepsle (1996), preferences are driven by policy concerns only.

Table 3 Numeric partitions that correspond to stable coalitional structures under various assumptions regarding commitment technologies

		(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Rule	**Open rule**	**Game Γ**	**Game Δ**	**Markov example**	**Sequential contracting**	**Repeated spot contracting**
	Commitment mechanism	Individuals coordinate with any player who selects the same coalition label and commit to equal within-group divisions	Individuals coordinate with player-defined sets of players provided these all agree and commit to equal within-group divisions	Individuals coordinate with any subset of a player-defined set of players provided these all agree and commit to equal within-group divisions	Individuals coordinate with coalition members subject to realignments through voluntary joining and acceptance, expulsion, or drift, and commit to equal within-group divisions	Individuals coordinate with coalitions and commit to a given (unrestricted) long-term allocation of payoffs across group members	Individuals coordinate with coalitions for a single decision and commit to a given (unrestricted) allocation of single-period payoffs across group members
a	{1,1,1,1,1,1,1}	✓;	✓	✓			
b	{2,1,1,1,1,1}	✓	✓	✓			
c	{3,1,1,1,1}		✓	✓			
d	{2,2,1,1,1}		✓	✓			
e	{3,2,1,1}						
f	{4,1,1,1}		✓	✓		✓	
g	{4,2,1}		✓	✓		✓	
h	{2,2,2,1}		✓	✓	✓		
i	{3,3,1}						
j	{3,2,2}			✓			
k	{4,3}		✓	✓		✓	
l	{5,2}		✓	✓			
m	{5,1,1}		✓	✓			
n	{6,1}		✓	✓			
o	{7}	✓	✓	✓			
	no absorbing states						✓

the example we are considering, under Rubinstein bargaining they should expect to receive half the pie from either partner. A process is required, however, for determining the set of coalitions. The simplest approach is one in which a formateur is randomly selected and asked to propose a coalition. If his proposal fails, another formateur is drawn.

Although the distinction between majoritarian and coalitional approaches is not standard in the literature, in fact both of these systems have been used by students of coalitional government. The majoritarian approach is used, for example, by Snyder et al. (2005) and Laver & Shepsle (1996), and the coalitional approach is used by Baron & Diermeier (2001), Iversen & Soskice (2006), and Eguia (2007).

Does this modeling choice matter? Underlying the distinction is a presupposition of different commitment technologies. Initially, one might expect it makes little difference whether one engages in multilateral bargaining or selects a coalition and bargains within it. In the first case, one has a one-third chance of getting two thirds of the pie and a one-third chance of getting one third of the pie; in the second case, one has a two-thirds chance of gaining half the pie. Either way, one expects one third of the pie. In fact, however, players may have different preferences over the two methods in the presence of asymmetries that are prominent in models of coalition formation. We consider two such asymmetries.

Size asymmetries. Consider first asymmetries arising from groups of different sizes or weights of the form studied by Snyder et al. (2005). In this model, a bargain is reached on the floor to divide up a pie between the members of some winning majority (a coalition). The question is: Would payoffs differ if allocational decisions were made within rather than across coalitions? The linearity of the results might seem to suggest that, under risk neutrality, benefits would be similar under both methods. This is not the case, and as I demonstrate next, different players can have markedly different preferences over the institutional form in cases like this.

To illustrate, consider the case of bargaining over the division of a pie of size 1 between groups of size {2,1,1,1}. The expected payoffs under a majoritarian model are {.4,.2,.2,.2} (Snyder et al. 2005).

What can players expect in a coalitional version of this game? Consider first bargaining within minimum winning coalitions. In this case, there are two types of coalition: small coalitions containing the strong player and one weak player, and large coalitions containing the three weak players. Under Rubinstein bargaining, players in the small coalitions receive $\frac{1}{2}$, while players in the weak coalitions receive $\frac{1}{3}$. The following then is an equilibrium of the coalition game: Each strong player proposes to bargain with one weak player, each weak player proposes a coalition with the strong player; all players accept all offers to enter minimum winning coalitions.

In the coalitional game, the strong player is always sought after as a coalition partner; his expected share of the pie therefore is $\frac{1}{2}$. A weak player has an expected share of just $\frac{1}{6}$. Hence, because of his attractiveness as a coalition partner, the strong player prefers to bargain within coalitions, but the weak players prefer to bargain across coalitions, on the floor. Since the weak players form a majority here, it is clear that in an up-down vote over the institutional rules, the noncoalitional approach would win.

However, majorities do not always oppose coalitional institutions. Consider the case with groups of size {2,2,1,1,1}. There continue to be strong benefits to the coalitional system for the larger parties. Indeed, if patient, large parties will reject any offers to be in a coalition with small parties, since any such winning coalition contains more than two players; they only enter government together. Their expected benefits are $\frac{1}{2}$ each compared to $\frac{2}{7}$ under Baron-Ferejohn bargaining. In an up-down vote, a majority will favor a coalitional government commitment device.

Constraints over admissible divisions. Consider finally a case in which players are risk-neutral and have uniform voting weights but in which constraints on the set of feasible allocations generate asymmetries between players. This is the setting examined by Iversen & Soskice (2006), who focus on redistribution in the presence of class coalitions subject to a nonregressivity constraint.

Suppose player $j \in \{1,2,3\}$ has endowment $t_j^o = \frac{j-1}{3}$. The net transfers to j, denoted p_j, are limited by three constraints:

1. Budget balance: $\sum_j p_j = 0$
2. Minimum negative payoffs: $p_j \geq -t_j^o$
3. Nonregressivity: $p_1 \geq p_2 \geq p_3$.

Together these constraints determine a subset of the unit simplex that constitutes acceptable distributions. Let an arbitrary point in the unit simplex be denoted by t, where $t \in [0, 1]^3$, $\sum t_j = 1$. The admissible subset of offers is then given by the set shown in **Figure 4**.

The question again is: What will the outcome look like under each institutional setting?

As above, we solve the game among each of the pairs using a Rubinstein-style alternating-offers bargaining model. Where t_{ij} denote the triple of payoffs when i and j bargain bilaterally, under risk neutrality in the limit as players tend to perfect patience, equilibrium outcomes are given by $t_{12} = \{\frac{1}{2}, \frac{1}{2}, 0\}$, $t_{23} = \{0, \frac{1}{3}, \frac{2}{3}\}$, $t_{13} = \{\frac{2}{3}, 0, \frac{1}{3}\}$.

The game thus produces sharp asymmetries in preferences over coalition partners. Given a choice, 1 would prefer to bargain with 3, 3 would prefer to bargain with 2, and 2 would prefer to bargain with 1. Under random appointment of a formateur, the following is an equilibrium: Each player, if selected as formateur, proposes his preferred coalition; all offers of partnering are accepted. Thus, even though each coalition will produce markedly different policies, any coalition can form in equilibrium.

Under random appointment of a formateur, the expected redistribution vector is

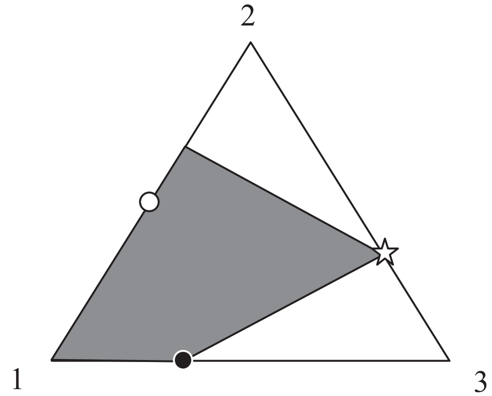


Figure 4

Example of a redistribution game of the form studied by Iversen & Soskice (2006). The set of acceptable redistributions is shaded; the "zero net transfers" case is marked with a star.

$\{\frac{7}{18}, \frac{5}{18}, \frac{6}{18}\}$ and expected net transfers are $\{\frac{7}{18}, -\frac{1}{18}, -\frac{6}{18}\}$.

As in the other cases, we can also study a majoritarian version of this game with equal recognition probabilities. In the present example, there is a particularly simple majoritarian equilibrium for cases in which players are patient. When selected, each player i proposes $\frac{2}{3}$ for himself and $\frac{1}{3}$ for player $(i+2)(\text{mod}3)$; all players accept any offers that accord them at least $\frac{1}{3}$. In this equilibrium, expected benefits are $\frac{1}{3}$ each, and expected net transfers are $\{\frac{1}{3}, 0, -\frac{1}{3}\}$.

Again we find disagreement over the optimal institutional procedures. The wealthy player, 3, is indifferent between systems. Player 1 benefits from the coalitional approach, while 2 prefers taking her chances on the floor. The key difference in outcomes arises when 2 is selected; if selected as a formateur, 2 will form a coalition with 1 and receive $\frac{1}{2}$. If, however, 2 can propose an outcome directly, then she can make better use of her agenda-setting power and propose retaining $\frac{2}{3}$, yielding only $\frac{1}{3}$ to player 1.

We find therefore that in a wide variety of settings, the presence of a coalitional commitment device can have substantial effects on the outcome of bargaining. In itself, this makes commitment devices of this form an interesting object of study—under what circumstances are we more likely to see coalitional institutions used?—but beyond that, it implies that this modeling choice is not innocuous, and the utility of a model should be judged on the reasonableness of this modeling choice. For majoritarian approaches, the question is whether, once formed, members of a coalition have the ability or incentive to push their partners to renegotiate deals that were previously reached in plenary. For coalitional approaches, the question is whether coalition partners have the ability or incentive to seek support for alternative deals from parties outside their coalition.

CONCLUSIONS

Recent research has greatly expanded our understanding of the kinds of coalitional structures that can, or are likely to, emerge in settings in which players can sign binding agreements with each other. This work has moved beyond an unproductive opposition that sometimes obtains between adherents of cooperative and noncooperative approaches. In this new work, coalition formation is modeled as an explicitly noncooperative process that occurs in settings in which binding agreements can be made. The correspondence between Nash equilibrium strategies and core outcomes can then be fruitfully examined.

A number of consistent messages have emerged from this work. First, in many cases, when core outcomes exist, they correspond to predictions that result from noncooperative analysis, both in the more obvious cases where players are myopic and in the more surprising cases where players are farsighted (Banks & Duggan 2000, Konishi & Ray 2003). In settings in which the core is empty, this fact also provides insights into the types of equilibria that may obtain, and indeed, consistent with older insights, the emptiness of the core may give rise in noncooperative settings to cyclical processes induced by noncooperative play. However, there are cases in which the core is not a good prediction, even when it is not empty. For the case in which contracts are temporary and large transfers can be made between players, Gomes & Jehiel (2005) show that although core outcomes may obtain for myopic players, these may not obtain if players are sufficiently farsighted. In such settings, unless a new condition is satisfied—negative externality–freeness—coalitions may choose to leave core states in order to extort other players who will make side payments to achieve subsequent gains.

Second, this research has found that in flexible negotiation environments (with durable but modifiable contracts and transferable utility), the Coase conjecture that efficient outcomes result from contracting typically holds (under subgame perfection) if grand coalitions can implement efficient outcomes. However, in cases in which the grand coalition cannot implement efficient outcomes, or in which contracts cannot be renegotiated, inefficient coalitional structures may result (Hyndman & Ray 2007).

Third, and most important, this new work highlights the dependence of predicted outcomes not simply on the bargaining protocol but, more substantively, on details of the space of admissible contracts—that is, on the types of commitments that can be made.

These new developments point to a number of fruitful avenues for new work on coalitional politics. I emphasize three of these.

The Stability of Coalitional Structures

We have seen that there has been some confusion regarding what lessons to take from the emptiness of the core, particularly in majority-rule settings. Core emptiness does not predict instability. But scholars have historically been conscious that in some sense, an empty core makes the status quo vulnerable; if the majority could simply find the right game form, they could change it. Recent work has clarified that the problem of instability is distinct from the problem of the emptiness of a solution concept. In truly dynamic settings, noncooperative solutions can generate stable behavioral predictions, but in settings where the core is empty, such as the DD game, these do not translate into stable political outcomes (Kalandrakis 2004, 2006b). Instability then seems to be inherent in the problem, not in the solution concept, and it may arise in many settings of importance to political scientists (Gomes & Jehiel 2005) as a function of preferences and protocols rather than simply because of exogenous shocks. There is an open agenda to better understand when such instability is likely to arise in different political and contractual environments.

The Differential Effects of Different Commitment Technologies

We have seen a striking dependence of predicted outcomes on the type of commitments that can be made between players. Whereas much work has focused on the impacts of different bargaining protocols, less attention has been given to a determination of what kinds of commitments may be made. As we saw in the example of government formation, some models invoke an ability to commit only to negotiate with a subset of players while others do not. This ability has material effect, yet the rationale for such a commitment device is rarely defended and the impacts of the assumption are rarely analyzed. Contracting possibilities may vary along a number of dimensions; some of these have been examined, whereas others remain open questions. In applied work, there has been little effort to match the commitment technologies made available in game description to the actual settings under study.

The signatories. In most treatments, contracts can either be signed by no individuals or by arbitrary coalitions of individuals. In principle, however, there may be a rich set of constraints on who can sign contracts, or who can sign contracts with whom. Plausibly, for example, contracts may only be signed between pairs, or between some pairs and not others. In some settings we may consider contracts signed by only one person, but in others these may be void. In other settings we might consider commitments credible only if made by some share of a group, a majority for example, or perhaps the grand coalition. Finally, in most work on coalition formation, contracts are signed uniquely among members of a given coalition; yet more general structures are conceivable in which contracting takes place across networks rather than within coalitions In such cases, individuals may have a series of contracts with different individuals or groups rather than within the same group.

The domain. Variation in the domain of contracts is also relevant, though largely unexplored in the context of coalitional analysis. Considerable work has been done on incomplete contracts arising where transaction costs make the specification of all possible eventualities infeasible. Maskin & Tirole (1999) argue that in such settings, transaction costs need not prevent optimal contracting; rather, contracts can be written in utility space rather than action space. Their results are very general whenever renegotiation is not permitted and are still general even with renegotiation if players are risk-averse. However, there are other reasons why limited domain may be relevant; in particular, there may be legal or other inhibitions on what is contractible. In the model of Iversen & Soskice (2006), contracts are not enforceable if they impose costs on poorer players. In the Laver & Shepsle (1996) model, parties, although they are in office together, cannot make a binding agreement on policy choices. Finally, there may be legal restrictions on the ability to make side payments. The Gomes & Jehiel (2005) model is appropriate for analyzing coalition formation only if side payments can be made for favors rendered.

The costs. All the models examined here assume that contracting is essentially costless. In all cooperative models, the payoffs are determined by outcomes and not by how those outcomes are reached. Some noncooperative approaches include costs in terms of time to contracting but ignore other features. The number of contracting parties is typically not considered a cost to contracting, although it is a plausible concern. Little has been done in this area, although we refer to Gul (1989) for the case of games played on characteristic functions and to Macho-Stadler et al. (2006) for a recent examination of bilateral mergers in a Cournot setting.

The duration. A number of the studies examined above highlight the importance of contract duration for equilibrium outcomes. Hyndman & Ray (2007), for example, usefully distinguish between temporary (or spot) agreements, irreversible agreements (which allow no renegotiation), and permanent agreements (which allow renegotiation among consenting signatories). Many features of duration have not been examined, however, such as contracts with endogenous sunset provisions or contracts whose termination depends on actions taken by only a subset of parties. Further issues arise if the possibility of reneging is considered, a point I return to below. For such cases, further details of contract design may become salient, such as whether contracting provides for joint or several liability.

Publicity. In present studies, contracts are public documents. However, in principle, the publicity or secrecy of contracts could affect play in ways not presently allowed for in contracting games. This matters especially if contracts specify actions to be taken. Such contracts could serve not only to commit players within a coalition to work in each other's interests but also to commit them to a course of action in the game between coalitions. In other settings, the publicity of contracts could reveal previously unavailable information to outside parties. Finally, as noted above (see Criticisms of the Cooperative Approach), even the refusal to sign a contract, if this is known, could provide information regarding intended play in a base game and thus alter play.

The process. As is well known, bargaining protocols can have a large effect on the outcome of negotiations. In the context of coalition formation, Gul (1989) shows that pairwise bargaining can lead to Shapley value payoffs when all (live) pairs are matched with equal probability. However, he also shows that in a related partnership game with different recognition probabilities, different outcomes obtain. Quite different procedures are proposed by Morelli (1999), Snyder et al. (2005), Diermeier & Merlo (2000), and others, all with substantive implications. Determining which approach is "right" is difficult because coalitional bargaining procedures are typically not embodied in constitutions (Diermeier & Merlo 2004). Moreover, even if formal rules exist, these rules may not reflect the relevant commitment devices available to players. The largest party may well be the first to be recognized on the floor, but a smaller party might have been the first to gather a group and reach a deal in the corridors. The challenge for applied work is to identify protocols that accord with the actual procedures used in the cases of interest.

Endogenous Commitment Technologies

Throughout this discussion I have treated commitment technologies as exogenous. In many accounts, these commitment devices are external institutional features; compliance is guaranteed by an outside agent or institution. In other accounts (e.g., Jackson & Moselle 2002), the commitments that hold parties together are seen less as legal devices and more as the result of repeated interactions within a larger, unmodeled game. Yet a clear lesson from the new work on coalition formation (examined in the previous section, Noncooperative Coalition Theory) is that different commitment technologies produce materially different outcomes that are valued differently by different individuals. Understanding this mapping from commitment technologies to welfare can provide information on the origin and institutionalization of those commitment technologies. In turn, a deeper understanding of the origins and the breaking points of commitment technologies can shape our understanding of when and how political coalitions will elect to employ them.

ACKNOWLEDGMENTS

I thank David Epstein, Jeff Lax, Lucy Goodhart, and Cyrus Samii for very helpful comments and suggestions. I am especially indebted to Michael Laver for extremely useful conversations on all parts of this article.

DISCLOSURE STATEMENT

The author is not aware of any biases that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Alesina A, Spolaore E. 1997. On the number and size of nations. *Q. J. Econ.* 112:1027–56

Ansolabehere S, Snyder JM, Strauss AB, Ting MM. 2005. Voting weights and formateur advantages in the formation of coalition governments. *Am. J. Polit. Sci.* 49:550–63

Aumann RJ. 1959. Acceptable points in general cooperative *n*-person games. In *Contributions to the Theory of Games*, *Annals of Mathematical Studies*, No. 40, ed. AW Tucker, RC Luce, IV, pp. 287–324. Princeton, NJ: Princeton Univ. Press

Aumann RJ. 1961. The core of a cooperative game without side payment. *Trans. Am. Math. Soc.* 98:539–52

Aumann RJ, Maschler M. 1964. The bargaining set for cooperative games. In *Advances in Game Theory*, ed. M Dresher, LS Shapley, AW Tucker. Princeton, NJ: Princeton Univ. Press

Aumann RJ, Myerson RB. 1988. Endogenous formation of links between coalitions and players: an application of the Shapley value. In *The Shapley Value*, ed. RJ Aumann, RB Myerson, pp. 207–20. Cambridge, UK: Cambridge Univ. Press

Austen-Smith D, Banks J. 1990. Stable governments and the allocation of policy portfolios. *Am. Polit. Sci. Rev.* 84:891–906

Banks JS, Duggan J. 2000. A bargaining model of collective choice. *Am. Polit. Sci. Rev.* 94:73–88

Banks JS, Duggan J. 2006. A general bargaining model of legislative policy-making. *Q. J. Polit. Sci.* 1:49–85

Baron DP, Diermeier D. 2001. Elections, governments, and parliaments in proportional representation systems. *Q. J. Econ.* 116:933–67

Baron DP, Ferejohn JA. 1989. Bargaining in Legislatures. *Am. Polit. Sci. Rev.* 83:1181–206

Bergin J, Duggan J. 1999. An implementation-theoretic approach to noncooperative foundations. *J. Econ. Theory* 86:50–76

Bernheim BD, Peleg B, Whinston MD. 1987. Coalition-proof Nash equilibria. I. Concepts. *J. Econ. Theory* 42:1–12

Binmore K. 1987. Nash bargaining theory. In *The Economics of Bargaining*, ed. K Binmore, P Dasgupta, pp. 27–46. New York: Blackwell

Binmore K, Rubinstein A, Wolinsky A. 1986. The Nash bargaining solution in economic modelling. *RAND J. Econ.* 17:176–88

Bloch F. 1996. Sequential formation of coalitions in games with externalities and fixed payoff division. *Games Econ. Behav.* 14:90–123

Calvert RL, Dietz N. 2005. Legislative coalitions in a bargaining model with externalities. In *Social Choice and Strategic Decisions*, ed. D Austen-Smith, J Duggan, pp. 227–47. New York: Springer

Caplin A, Nalebuff B. 1997. Competition among institutions. *J. Econ. Theory* 72(2):306–42

Carraro C, Marchiori C, Sgobbi A. 2006. *Advances in negotiation theory: bargaining, coalitions and fairness*. Fond. Eni Enrico Mattei (FEEM) Work. Pap. No.66.05

Chatterjee K, Dutta B, Ray D, Sengupta K. 1993. A noncooperative theory of coalitional bargaining. *Rev. Econ. Stud.* 60:463–77

Coughlan PJ, Le Breton M. 1999. A social choice function implementable via backward induction with values in the ultimate uncovered set. *Rev. Econ. Design* 4:153–60

Chandler P, Tulkens H. 1995. A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. *Int. Tax Pub. Financ.* 2:279–93

Chwe MSY. 1994. Farsighted coalitional stability. *J. Econ. Theory* 63:299–325

Diamantoudi E, Xue L. 2007. Coalitions, agreements and efficiency. *J. Econ. Theory* 136:105–25

Diermeier D, Krehbiel K. 2003. Institutionalism as a methodology. *J. Theor. Polit.* 15:123–44

Diermeier D, Merlo A. 2000. Government turnover in parliamentary democracies. *J. Econ. Theory* 94:46–79

Diermeier D, Merlo A. 2004. An empirical investigation of coalitional bargaining procedures. *J. Pub. Econ.* 88:783–97

Diermeier D. 2006. Coalition government. In *The Oxford Handbook of Political Economy*, ed. B Weingast, D Wittman, pp. 162–79. Oxford, UK: Oxford Univ. Press

Duggan J. 2005. *A survey of equilibrium analysis in spatial models of elections*. Work. pap., Dep. Polit. Sci., Rochester Univ.

Eguia J. 2007. *Endogenous parties in an assembly: the formation of two polarized blocs*. Work. pap., Dep. Polit. Sci., New York Univ.

Gamson WA. 1961. A theory of coalition formation. *Am. Sociol. Rev.* 26:373–82

Garrett G, Tsebelis G. 2001. Even more reasons to resist the temptation of power indices in the EU. *J. Theor. Polit.* 13:99–105

Gillies DB. 1959. Solutions to general nonzero-sum games. In *Contributions to the Theory of Games IV*, ed. RD Luce, AW Tucker. pp. 47–85. Princeton, NJ: Princeton Univ. Press

Gomberg AM. 2004. Sorting equilibrium in a multi-jurisdiction model. *J. Econ. Theory* 116:138–54

Gomberg AM. 2005. An example of nonexistence of three-community equilibrium. *J. Pub. Econ. Theory* 7:285–94

Gomes A. 1999. *A theory of negotiations and formation of coalitions*. Work. pap., Wharton School Bus., Univ. Penn.

Gomes A. 2005. Multilateral contracting with externalities. *Econometrica* 73:1329–50

Gomes A, Jehiel P. 2005. Processes of social and economic interactions: on the persistence of inefficiencies. *J. Polit. Econ.* 113:626–67

Greenberg J. 1990. The theory of social situations. New York: Cambridge Univ. Press

Gul F. 1989. Bargaining foundations of the Shapley value. *Econometrica* 57:81–95

Haake CJ. 2000. *Supporting the Kalai-Smorodinsky bargaining solution*. Work. pap., Inst. Math. Econ., Bielefeld Univ.

Harsanyi JC. 1974. An equilibrium-point interpretation of stable sets and a proposed alternative definition. *Manage. Sci.* 20:1472–95

Harsanyi JC. 1981. The Shapley value and the risk dominance solutions of two bargaining models for characteristic-function games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, ed. RJ Aumann. et al., pp. 43–68. Mannheim: Bibliogr. Inst. Mannheim

Hart O, Moore J. 1990. Property rights and the nature of the firm. *J. Polit. Econ.* 98:1119–57
Hart S, Kurz M. 1983. Endogenous formation of coalitions. *Econometrica* 51:1047–64
Hart S, Mas-Colell A. 1996. Bargaining and value. *Econometrica* 64:357–80
Hyndman K, Ray D. 2007. Coalition formation with binding agreements. *Rev. Econ. Stud.* 74:1125–47
Iversen T, Soskice D. 2006. Electoral institutions and the politics of coalitions: why some democracies distribute more than others. *Am. Polit. Sci. Rev.* 100:154–81
Jackson M, Moselle B. 2002. Coalition and party formation in a legislative voting game. *J. Econ. Theory* 103:49–87
Johnson R, Libecap G. 2003. Transaction costs and coalition stability under majority rule. *Econ. Inquirer* 41:193–207
Kalandrakis T. 2004. A three-player dynamic majoritarian bargaining game. *J. Econ. Theory* 116:294–322
Kalandrakis T. 2006a. Proposal rights and political power. *Am. J. Polit. Sci.* 50:441–48
Kalandrakis T. 2006b. *Majority rule dynamics with endogenous status quo*. Work. pap., Dep. Polit. Sci., Univ. Rochester
Konishi H, Ray D. 2003. Coalition formation as a dynamic process. *J. Econ. Theory* 110:1–41
Konishi H, Le Breton M, Weber S. 1997. Equivalence of strong and coalition-proof Nash equilibria in games without spillovers. *Econ. Theory* 9:97–113
Krehbiel K. 1993. Where's the party? *Br. J. Polit. Sci.* 23:235–66
Krishna V, Serrano R. 1996. Multilateral bargaining. *Rev. Econ. Stud.* 63:61–80
Laver M, Benoit K. 2003. The evolution of party systems between elections. *Am. J. Polit. Sci.* 47:215–33
Laver M, Shepsle KA. 1996. *Making and Breaking Governments: Cabinets and Legislatures in Parliamentary Democracies*. Cambridge, UK: Cambridge Univ. Press
Luce RD, Raiffa H. 1957. *Games and Decisions*. New York: Wiley
Macho-Stadler I, Perez-Castrillo D, Porteiro N. 2006. Sequential formation of coalitions through bilateral agreements in a Cournot setting. *Int. J. Game Theory* 34:207–28
Mas-Colell A, Whinston MD, Green JR. 1995. *Microeconomic Theory*. Oxford, UK: Oxford Univ. Press
Maschler M, Owen G. 1989. The consistent Shapley value for hyperplane games. *Int. J. Game Theory* 18:389–407
Maskin E. 2003. *Bargaining, coalitions, and externalities*. Work. pap., Inst. Advanced Stud., Princeton Univ.
Maskin E. 1999. Nash equilibrium and welfare optimality. *Rev. Econ. Stud.* 66:23–38
Maskin E, Tirole J. 1999. Unforeseen contingencies and incomplete contracts. *Rev. Econ. Stud.* 66:83–114
McKelvey RD. 1986. Covering, dominance, and institution-free properties of social choice. *Am. J. Polit. Sci.* 30:283–314
McKinsey JCC. 1952. Some notions and problems in game theory. *Bull. Am. Math. Soc.* 58: 591–611
Miller NR. 1980. A new solution set for tournaments and majority voting. *Am. J. Polit. Sci.* 24:68–96
Montero M, Vidal-Puga JJ. 2007. Demand commitment in legislative bargaining. *Am. Polit. Sci. Rev.* 101(04):847–50
Morelli M. 1999. Demand competition and policy compromise in legislative bargaining. *Am. Polit. Sci. Rev.* 93:809–20
Moulin H. 1984. Implementing the Kalai-Smorodinsky bargaining solution. *J. Econ. Theory* 33:32–45

Nash JF. 1953. Two-person cooperative games. *Econometrica* 21:128–40
Olson M. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard Univ. Press
Osborne MJ, Rubinstein A. 1994. *A Course in Game Theory*. Cambridge, MA: MIT Press
Perry M, Reny P. 1994. A noncooperative view of coalition formation and the core. *Econometrica* 62:795–817
Posner DN. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge, UK: Cambridge Univ. Press
Ray D, Vohra R. 1999. A theory of endogenous coalition structures. *Games Econ. Behav.* 26:286–336
Ray D, Vohra R. 2001. Coalitional power and public goods. *J. Polit. Econ.* 109:1355–85
Riker WH. 1962. *The theory of political coalitions*. New Haven, CT: Yale Univ. Press
Roemer J. 1982. *A General Theory of Exploitation and Class*. Cambridge: Harvard Univ. Press
Rogowski R. 1989. *Commerce and Coalitions: How Trade Affects Domestic Political Alignments*. Princeton, NJ: Princeton Univ. Press
Rota GC. 1964. The number of partitions of a set. *Am. Math. Mon.* 71:498–504
Rubinstein A. 1980. Stability of decision systems under majority rule. *J. Econ. Theory* 23(2):150–59
Rubinstein A. 1982. Perfect equilibrium in a bargaining model. *Econometrica* 50:97–109
Schmeidler D. 1969. The nucleolus of a characteristic function game. *SIAM J. Appl. Math.* 17:1163–70
Serrano R, Vohra R. 1994. Non-cooperative implementation of the core. *Soc. Choice Welfare* 14:513–25
Shenoy PP. 1979. On coalition formation: a game-theoretical approach. *Int. J. Game Theory* 8:133–64
Shepsle KA. 1986. Institutional equilibrium and equilibrium institutions. In *Political Science: The Science of Politics*, ed. HF Weisberg, pp. 51–82. New York: Agathon
Snyder JM, Ting MM, Ansolabehere S. 2005. Legislative bargaining under weighted voting. *Am. Econ. Rev.* 95:981–1004
Ståhl I. 1972. *Bargaining theory*. Work. pap., Stockholm School Econ.
Strom K, Budge I, Laver MJ. 1994. Constraints on cabinet formation in parliamentary democracies. *Am. J. Polit. Sci.* 38:303–35
Thrall RM, Lucas WF. 1963. N-person games in partition function form. *Nav. Res. Logistic Q.* 10:281–93
Tiebout C. 1956. A pure theory of local expenditures. *J. Polit. Econ.* 64:416–24
Trockel W. 1999. Unique Nash implementation for a class of bargaining solutions. *Int. Game Theory Rev.* 1:267–72
Vidal-Puga J. 2005. A bargaining approach to the Owen value and the Nash solution with coalition structure. *Econ. Theory* 25:679–701
von Neumann J, Morgenstern O. 1944. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton Univ. Press
Ward B. 1961. Majority rule and allocation. *J. Confl. Resolut.* 5(4):379–89
Weisstein E. 2005. Number of decimal digits in BellB[10_n]. **http://www.research.att.com/~njas/sequences/A113015**
Winter E. 1994. The demand commitment bargaining and snowballing cooperation. *Econ. Theory* 4:255–73
Yi SS. 1997. Stable coalition structures with externalities. *Games Econ. Behav.* 20:201–37
Yi SS, Shin H. 1994. *Endogenous formation of coalition in oligopoly*. Unpublished manuscript, Dep. Econ., Dartmouth College

Annual Review of Political Science

Volume 11, 2008

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Political Science* articles may be found at http://polisci.annualreviews.org/