STABILITY AND THE DISTRIBUTION OF POWER

By ROBERT POWELL*

THE relation between the distribution of power and the probability of war has been at the center of a long and important debate in international relations theory. Is an even distribution of power more or less stable than a preponderance of power? The balance-of-power school generally argues that an even distribution of power is more stable, whereas the preponderance-of-power school generally argues the opposite, that a preponderance of power is more stable.[1]

Empirical efforts to resolve this debate have yielded equivocal results. Siverson and Tennefoss, for example, find an even distribution of power to be more stable, whereas Kim and Moul find a preponderance of power to be more stable. Singer, Bremer, and Stuckey find an even distribution to be more stable than a preponderance of power during the nineteenth century and the opposite during the twentieth century. Maoz and Bueno de Mesquita and Lalman, by contrast, find no significant relation between the probability of war and the distribution of power. Finally, Mansfield finds evidence of a nonlinear, quadratic relation in which the probability of war is smallest when there is an even distribution and a preponderance of power.[2]

* I am grateful to Bruce Bueno de Mesquita, James Fearon, Joanne Gowa, Robert Keohane, James Morrow, Paul Papayoanou, Leo Simon, R. Harrison Wagner, and Barry Weingast for their comments and criticisms. This work was assisted by a grant from the National Science Foundation (SES 9121959) and a National Fellowship at the Hoover Institution at Stanford University.

[1] Arguing that a balance is more stable are Innis Claude, *Power and International Relations* (New York: Random House, 1962); Hans Morgenthau, *Politics among Nations*, 4th ed. (New York: Alfred Knopf, 1967); John Mearsheimer, "Back to the Future," *International Security* 15 (Summer 1990); Quincy Wright, *A Study of War* (Chicago: University of Chicago Press, 1965); and Arnold Wolfers, *Discord and Collaboration* (Baltimore: Johns Hopkins University Press, 1962). Arguing that a preponderance is more stable are Geoffrey Blainey, *The Causes of War* (New York: Free Press, 1973); A. F. K. Organski, *World Politics* (New York: Alfred A. Knopf, 1968); and A. F. K. Organski and Jacek Kugler *The War Ledger* (Chicago: University of Chicago Press, 1980). For a review of this debate, see Jack Levy, "The Causes of War: A Review of Theories and Evidence," in Philip Tetlock et al., eds., *Behavior, Society, and Nuclear War*, vol. 1 (New York: Oxford University Press, 1989); for a discussion of some of the important conceptual issues, see R. Harrison Wagner, "Peace, War, and the Balance of Power," *American Political Science Review* 88 (September 1994).

[2] Randolph Siverson and Michael Tennefoss, "Power, Alliance, and the Escalation of International Conflict, 1815–1965," *American Political Science Review* 78 (December 1984); Woosang Kim, "Alliance Transitions and Great Power War," *American Journal of Political Science* 35 (November 1991); idem,

World Politics 48 (January 1996), 239–67

Claims about the relation between the distribution of power and the likelihood of war have generally lacked any formal foundations. Some recent work has attempted to fill this gap, however. Bueno de Mesquita and Lalman and Fearon analyze models of international bargaining. The former find, first, that the probability of war is smallest when the distribution of power is uneven and, second, that "dissatisfaction with the status quo is unrelated to the likelihood of war."[3] Fearon, by constrast, finds that the probability of war is independent of the distribution of power; that is, a change in the distribution of power has no effect on the likelihood of war.[4]

This essay examines the relation between the probability of war and the distribution of power in the context of a game-theoretic model in which two states are bargaining about revising the international status quo. In the model the states make proposals for changing the status quo. The bargaining continues until the states reach a mutually acceptable settlement or until one of them becomes sufficiently pessimistic about the prospects of reaching an agreement that it uses force to try to impose a new international order. The states' equilibrium strategies specify the demands the states will make of each other and the circumstances in which they will resort to force. The equilibrium strategies thus make it possible to calculate the probability that bargaining will break down in war as a function of the distribution of power between the two states. The shape of this function can then be compared to the relation claimed to exist by the balance-of-power and preponderance-of-power schools.

The probability of war in the model turns out to be a function of the disparity between the status quo distribution of benefits and the distribution expected to result from the use of force. The outcome expected from using force is, in turn, directly related to the distribution of power.

"Power Transitions and Great Power War from Westphalia to Waterloo," *World Politics* 45 (October 1992); William Moul, "Balances of Power and the Escalation to War of Serious International Disputes among the European Great Powers, 1815–1939," *American Journal of Political Science* 32 (February 1988); J. David Singer, Stuart Bremer, and John Stuckey, "Capability Distribution, Uncertainty, and Major Power War, 1820–1965," in Bruce Russett, ed., *Peace, War, and Numbers* (Beverly Hills, Calif.: Sage, 1972); Zeev Maoz, "Resolve, Capabilities, and the Outcomes of Interstate Disputes, 1815–1976," *Journal of Conflict Resolution* 27 (June 1983); Bruce Bueno de Mesquita and David Lalman, "Empirical Support for Systemic and Dyadic Explanations of International Conflict," *World Politics* 41 (October 1988); idem, *Reason and War* (New Haven: Yale University Press, 1992); Edward Mansfield, "The Concentration of Capabilities and the Onset of War," *Journal of Conflict Resolution* 36 (March 1992); idem, *Power, Trade, and War* (Princeton: Princeton University Press, 1994).

[3] Bueno de Mesquita and Lalman (fn. 2, 1992), 204–5, 190.

[4] James Fearon, "War, Relative Power, and Private Information" (Paper presented at the annual meeting of the International Studies Association, Atlanta, March 31–April 4, 1992), 20.

Thus, the probability of war is a function of the disparity between the status quo and the distribution of power. When this disparity is small, the division of benefits expected from the use of force is approximately the same as the existing status quo distribution. The gains to using force are therefore too small to outweigh the costs of fighting. Neither state is willing to use force to change the status quo, and the probability of war is zero. When the disparity between the status quo division of benefits and the distribution of power is large, then at least one state is willing to use force to overturn the status quo. Moreover, as the disparity grows, the probability of war generally increases.

These results seem quite intuitive: as the disparity between the distribution of power and the distribution of benefits grows, the probability of war generally increases. Indeed, this formal result echoes Gilpin's explanation of hegemonic wars.[5] Such wars arise, he argues, because of a disparity between the distribution of power and the international status quo distribution of benefits imposed by the hegemon following the previous hegemonic war.[6] Although intuitive, these results contradict the expectations of both the balance-of-power and preponderance-of-power schools. The former expects the probability of war to be smallest when power is evenly distributed, but the probability of war is smallest in the present model when the distribution of power mirrors the status quo distribution of territory. These claims will coincide only in the special case in which the status quo distribution is even. The preponderance-of-power school expects the probability of war to be smallest when the distribution of power is highly skewed and to increase as the distribution of power becomes more even. But the probability of war is largest in the model developed below when the distribution of power is highly skewed.

The formal results derived here also differ from those obtained by Fearon and Bueno de Mesquita and Lalman. Unlike Fearon, the probability of war in the present formulation does vary with changes in the distribution of power.[7] Unlike Bueno de Mesquita and Lalman, who find that "dissatisfaction with the status quo is unrelated to the likelihood of war" and that the probability of war declines as the distribu-

[5] Robert Gilpin, *War and Change in World Politics* (Cambridge: Cambridge University Press, 1981).

[6] Although the potential use of force is not the source of the coercive pressure, some suggest that international regimes are more likely to break down when the underlying distribution of power differs from the distribution of benefits the regime confers. See Robert Keohane and Joseph Nye, *Power and Interdependence* (Boston: Little Brown, 1977), 139; and Stephen Krasner, "Regimes and the Limits of Realism," in Krasner, ed., *International Regimes* (Ithaca, N.Y.: Cornell University Press, 1983), 357–58,

[7] Fearon (fn. 4).

tion of power becomes more skewed, the probability of war is directly related to the level of dissatisfaction with the status quo and generally increases as the distribution of power becomes more skewed.[8]

The remainder of this essay is organized as follows. The next section motivates and defines the model and discusses some qualifications and limitations. The subsequent section describes the game's equilibrium. The final section uses this equilibrium to specify the probability of war as a function of the distribution of power and compares these results with the expectations of the balance-of-power and preponderance-of-power schools and with other formal work. That section also offers a possible explanation for the equivocal empirical results of efforts to measure the relation between stability and the distribution of power.

The Model

To motivate the model, suppose two states are bargaining about revising the international status quo. More concretely, imagine that two states are negotiating about changing the territorial status quo. Each state proposes a division of the territory that the other state can accept or reject. The bargaining continues with offer followed by counteroffer until the states agree on a proposed division or until one of the states becomes sufficiently pessimistic about the prospects of reaching a mutually acceptable settlement that it resorts to force to try to impose a new territorial settlement.[9]

How do changes in the distribution of power between these two states affect the bargaining and, more specifically, the probability that the bargaining will break down in war? Are the negotiations more likely to break down if one state has a preponderance of power or if there is an even distribution of power? It might seem at first that the preponderance-of-power school is clearly correct and that the bargaining is less likely to break down if one of the states has a preponderance of power. That is, suppose a state, say S_1, has just proposed a particular division to the other state S_2. S_2 would be tempted to use force to impose a settlement rather than accept S_1's offer only if the expected payoff to using force were at least as large as the payoff to agreeing to the proposal. But the weaker S_2, the lower its expected payoff to fighting. In-

[8] Bueno de Mesquita and Lalman (fn. 2, 1992), 190, 204–5.

[9] In a more general model the two states might be bargaining about altering the international status quo where the set of feasible outcomes is an n-dimensional policy space. Each point in this space represents a different international order and a different distribution of benefits for the two states. The formal analysis developed below applies equally well to this more general formulation.

deed, if S_2 is sufficiently weak, its payoff to using force will be less than its payoff to agreeing to S_1's offer. In this case, S_2 will acquiesce to S_1's demand and will not be tempted to use force. Nor will S_1 be tempted to use force, since S_2 is willing to agree to S_1's demand. This reasoning, which suggests that a preponderance of power will be more stable, is essentially the reasoning Organski advances: "A preponderance of power . . . increases the chances of peace, for the greatly stronger side need not fight at all to get what it wants, while the weaker side would be plainly foolish to attempt battle for what it wants."[10]

There is, however, a countervailing factor at work. The preceding argument was based on the assumption that S_1's proposal was fixed. But as S_1 becomes stronger and S_2 becomes weaker, S_1 knows that S_2 is more likely to accept any given proposal. Thus, S_1 will demand more of S_2 as S_2 becomes weaker. S_2, in turn, may be more likely to resist these larger demands, and war may become more likely when there is a preponderance of power. If so, then a preponderance of power would be less stable. By contrast, an even distribution of power would restrain all states' demands and thereby enhance stability. This restraining effect of an even distribution of power underlies the balance-of-power school's analysis of the relation between stability and the distribution of power. As Wolfers argues, "Thus, from the point of view of preserving the peace . . . it may be a valid proposition that a balance of power placing restraint on every nation is more advantageous in the long run than the hegemony even of those deemed peace-loving at the time."[11]

In sum, a shift in the balance of power against S_2 has two competing effects. First, S_2 is more likely to accept any specific demand. But, second, more will be demanded of it than would have been had it been more powerful. The first effect tends to make war less likely, while the second tends to make war more likely. The net result of these two opposing effects on the probability of war is unclear: whereas the preponderance-of-power school emphasizes the first effect, the balance-of-power school underscores the second effect. The game-theoretic model of bargaining developed here makes it possible to weigh the interaction of these two competing effects more formally.

Rubinstein's seminal analysis of an alternating-offer, infinite-horizon bargaining game provides a point of departure for the present analysis.[12] In Rubinstein's model two actors are trying to divide a pie. One actor begins the game by proposing a division of the pie to the other,

[10] Organski (fn. 1), 294.
[11] Wolfers (fn. 1), 120. See also Claude (fn. 1), 62.
[12] Ariel Rubinstein, "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50 (January 1982).

who can either accept or reject the offer. Acceptance ends the game and the pie is divided as agreed. If the second actor rejects the initial proposal, it makes a counteroffer that the first actor can either accept or reject. Acceptance again ends the game. If the first actor rejects the proposal, it makes a new offer. The actors alternate making offers until they agree on a division.[13]

The present model adds the option of using force to Rubinstein's basic structure. As illustrated in Figure 1, S_1 begins the game by proposing a division of the status quo. More formally, S_1 demands some fraction x of the disputed territory where x lies between zero and one. If, for example, S_1 sets $x = 1$, then S_1 is demanding all of the territory for itself; if S_1 sets $x = 1/2$, it is demanding half of the territory; and if S_1 sets $x = 0$, it is offering S_2 all of the territory. (The arc between "0" and "1" in Figure 1 indicates that a state can demand any fraction between 0 and 1.) S_2 can accept this offer by saying yes, which is denoted by Y; reject it by saying no, N, and subsequently make a counteroffer; or force a settlement by attacking S_1 with A. The game ends if S_2 accepts or attacks. If S_2 rejects, it then proposes a new territorial division. S_1 can now accept, reject, or attack. As before, accepting or attacking ends the game. If S_1 rejects, it again suggests another division. The bargaining continues, possibly forever, with offers alternating back and forth until one of the states accepts a proposal or attacks.

To specify the states' payoffs, let $(q, 1 - q)$ denote the status quo division where q and $1 - q$ are S_1's and S_2's respective shares. Now suppose that the states agree to divide the territory according to $(x, 1 - x)$ where S_1 and S_2 receive x and $1 - x$, respectively. Assume further that this agreement is reached at time t, that is, after $t + 1$ offers. (For notational convenience, the first offer is made at time $t = 0$.) If agreement is reached at time t, then S_1 had q and S_2 had $1 - q$ from the beginning of the game until time $t - 1$. The payoff that S_1 derives from having a flow of q for the first t periods will be taken to be $U_1(q) + \delta U_1(q) + \delta^2 U_1(q) + \cdots + \delta^{t-1} U_1(q) = (1 - \delta^t) U_1(q)/(1 - \delta)$ where δ is the states' common discount factor and U_1 is S_1's utility function for benefits. S_1 will also be assumed to prefer larger allocations to smaller allocations but with a nonincreasing marginal value.[14] Once agreement is reached at time t, S_1's allocation from time t forward is x. The present value of having x from time t forward is $\delta^t U_1(x) + \delta^{t+1} U_1(x) + \delta^{t+2} U_1(x) + \cdots = \delta^t U_1(x)/(1 - \delta)$. Thus, S_1's payoff to agreeing to $(x, 1 - x)$ at time t is $(1 - \delta^t) U_1(q)/(1 - \delta) + \delta^t U_1(x)/(1 - \delta)$. Similarly, S_2's payoff to agree-

[13] Rubinstein (fn. 12) showed that this game has a unique subgame-perfect equilibrium.

[14] More formally, U_1 is assumed to be twice differentiable with $U_1' > 0$ and $U_1'' \leq 0$. S_1 is risk neutral if $U_1'' = 0$ and risk averse if $U_1'' < 0$.

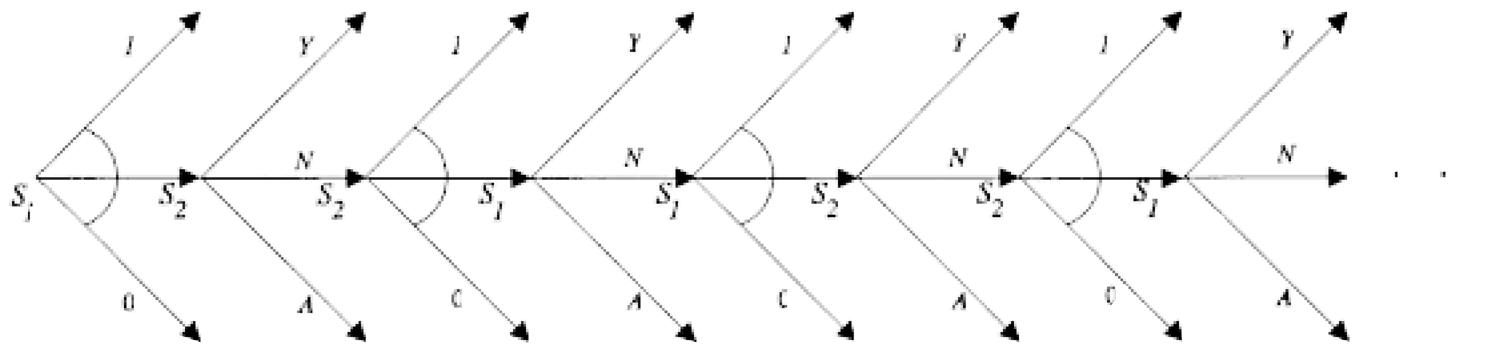


FIGURE 1
AN INFINITE-HORIZON, ALTERNATING-OFFER GAME WITH THE OPTION OF USING FORCE

ing to $(x, 1-x)$ at time t is $(1-\delta^t)U_2(1-q)/(1-\delta) + \delta^t U_2(1-x)/(1-\delta)$ where U_2 is S_2's utility.

To define the states' payoffs if they fail to reach an agreement and one of them uses force, suppose that they fight at time t. As before, S_1 derives $(1-\delta^t)U_1(q)/(1-\delta)$ from having q during the first t periods. To simplify the specification of S_1's payoff following an attack, war will be assumed to end in one of only two ways. Either S_1 will be completely victorious, or S_2 will be. (The limitations of this simplification are discussed below.) If S_1 prevails, it takes all of the territory at a cost of fighting c_1. The payoff to having all of the territory less the cost of fighting from time t forward is $\delta^t(U_1(1)-c_1) + \delta^{t+1}(U_1(1)-c_1) + \delta^{t+2}(U_1(1)-c_1) + \ldots = \delta^t(U_1(1)-c_1)/(1-\delta)$. If S_1 loses, S_2 captures all of the territory which leaves S_1 with nothing but the cost of fighting. The payoff to this outcome is $\delta^t(U_1(0)-c_1) + \delta^{t+1}(U_1(0)-c_1) + \delta^{t+2}(U_1(0)-c_1) + \ldots = \delta^t(U_1(0)-c_1)/(1-\delta)$.

Going to war is an uncertain venture. Let p reflect the balance of power or the distribution of capabilities between S_1 and S_2. That is, p, which is assumed to be common knowledge, is the probability that S_1 will prevail if force is used.[15] Then S_1's expected payoff to a bargaining process that ends in war at time t is the value of living with the status quo until time $t-1$ plus the payoff to prevailing at time t times the probability of prevailing, p, plus the payoff to losing at time t times the probability of losing, $1-p$:

$$\frac{(1-\delta^t)U_1(q)}{1-\delta} + p\,\frac{\delta^t(U_1(1)-c_1)}{1-\delta} + (1-p)\,\frac{\delta^t(U_1(0)-c_1)}{1-\delta}\;.$$

This payoff can be simplified by normalizing the utility function by setting the payoff to having all of the territory equal to one and the payoff to having no territory equal to zero (that is, $U_1(1) = 1$ and $U_1(0) = 0$). Letting $W_1(t)$ and $W_2(t)$ denote S_1's and S_2's normalized payoffs to a bargaining process that ends in war at time t leaves:

$$W_1(t) = \left(\frac{1-\delta^t}{1-\delta}\right) U_1(q) + \frac{\delta^t}{1-\delta}(p-c_1)$$

$$W_2(t) = \left(\frac{1-\delta^t}{1-\delta}\right) U_2(1-q) + \frac{\delta^t}{1-\delta}(1-p-c_2)\;.$$

[15] There are no offensive or defensive advantages in the present formulation, so it makes no difference whether S_1 or S_2 attacks.

The model developed so far has complete information. Each state knows the other state's payoffs. As discussed more fully below, as long as information is complete, bargaining never breaks down, regardless of the distribution of power. In this model incomplete information is a prerequisite to war. In the next section each state is assumed to be uncertain of the other's willingness to use force.

Before turning to incomplete information and an analysis of the game's equilibrium, however, it will be useful to consider some of the limitations and qualifications of the model. The most significant limitation is that there are only two states. One argument about the relation between stability and the distribution of power emphasizes the importance of the number of coalitions that could form to block any state's bid for hegemony. The larger the number of states and the more even the distribution of power, the larger the number of potential blocking coalitions and the less likely is war.[16] As a formal analysis of this argument clearly requires a model in which there are more than two actors, it is beyond the scope of the present analysis.[17]

It is important to emphasize, however, that although some arguments about the relation between the distribution of power and the probability of war presuppose the existence of more than two states, other arguments do not. Wright, for example, explicitly considers the situation in which there are only two states. In this case, balance-of-power theory expects that "there would be great instability unless [the two states] were very nearly equal in power or their frontiers were widely separated or difficult to pass."[18] Similarly, Mearsheimer considers the case of two great powers and claims that a balance of power is more stable: "Power can be more or less equally distributed among the major powers of both bipolar and multipolar systems. *Both* are more peaceful when equality is greatest among the poles."[19] In sum, the conflicting claims of the balance-of-power and preponderance-of-power schools apply to systems in which there are only two states, as well as to systems with more than two states. The present analysis focuses on the former case.

A second significant limitation of the model is that the distribution of power is fixed. But the key to several arguments about the relation

[16] See Levy (fn. 1), 231–32; Mansfield (fn. 2, 1994), 17–18; and Wright (fn. 1), 755.

[17] For efforts to examine the relation between stability and the distribution of power in a setting in which there are more than two states, see Emerson Niou and Peter Ordeshook, "Stability in Anarchic Internatonal Systems," *American Journal of Political Science* 84 (December 1990); Wagner (fn. 1); and idem, "The Theory of Games and the Balance of Power," *World Politics* 38 (July 1986).

[18] Wright (fn. 1), 755.

[19] Mearsheimer (fn. 1), 18; emphasis added.

between the likelihood of war and the distribution of power is that this distribution shifts over time, perhaps because of uneven economic growth.[20] This raises an important problem, which is the effects of changes in the distribution of power over time on the probability of war. Unfortunately, it is beyond the scope of the present model and must await future work.[21]

Another seeming limitation is that war can result in only two possible outcomes: a state either wins all of the territory or loses all of it. This limitation is more apparent than real, however. The assumption that war can end in only two ways simplifies the notation and discussion, but it is not central to the analysis of the model. One might assume instead that war could result in any territorial division. Changes in the distribution of power would still be reflected in changes in the probability distribution over the now larger set of possible outcomes.

Furthermore, both states are assumed to agree on the distribution of power (that is, both states know p). This assumption not only simplifies the analysis but has some substantive import for international relations theory as well. Blainey and others argue that war results from uncertainty about the distribution of power.[22] While this uncertainty may be an important cause of war, it is not a necessary condition as Blainey suggests. The model serves as a counterexample to this claim: states agree on the distribution of power and yet there is war.[23]

Finally, the terms of the agreement do not affect the distribution of power. So, for example, an agreement that transfers a large amount of territory to one state does not make that state more powerful. Wagner argues that this may be an important cause of war, and extending the model to allow for this possibility would be relatively straightforward.[24]

THE EQUILIBRIUM

This section characterizes the game's equilibrium. To summarize the results, a state will be called dissatisfied if it prefers fighting to the status quo and satisfied if it prefers the status quo to fighting. If both states

[20] See, for example, Charles Doran and Wes Parsons, "War and the Cycle of Relative Power," *American Political Science Review* 74 (December 1980); Gilpin (fn. 5); Organski (fn. 1); and Organski and Kugler (fn. 1).

[21] For efforts in this direction, see Fearon (fn. 4); Woosang Kim and James Morrow, "When Do Power Transitions Lead to War?" *American Journal of Political Science* 36 (November 1992); and Powell, "Appeasement as a Game of Timing" (Manuscript, Department of Political Science, University of California, Berkeley, July 1995).

[22] Blainey (fn. 1).

[23] Fearon (fn. 4) first makes and develops this point.

[24] Wagner (fn. 1).

are satisfied, then the status quo remains unchanged and the probability of war is zero, as neither state can credibly threaten to use force to overturn the status quo. If one of the states is dissatisfied, then the satisfied state makes its optimal offer to the dissatisfied state given the satisfied state's beliefs about the willingness of the dissatisfied state to use force. Although the dissatisfied state can always reject this offer and make a counteroffer, it never does so in equilibrium. Rather, the dissatisfied state either accepts this offer or attacks.[25] The next section uses these equilibrium strategies to calculate the probability of war as a function of the distribution of power. (Readers less interested in the derivation of these results may omit the rest of this section.)

Three preliminaries are needed before the equilibrium strategies can be specified. First, a dissatisfied state must be defined more precisely. Second, the equilibrium of the complete-information game must be described. Finally, incomplete information needs to be introduced.

A state is dissatisfied if it prefers fighting to living with the status quo. In symbols, S_1's payoff to fighting now (that is, at $t = 0$) is $W_1(0) = (p - c_1)/(1 - \delta)$. S_1's payoff to living with the status quo is $U_1(q)/(1 - \delta)$. Thus, S_1 is dissatisfied if $p - c_1 > U_1(q)$. Similarly, S_2 is dissatisfied if $1 - p - c_2 > U_2(1 - q)$. At most only one state can be dissatisfied.[26]

The complete-information game has a very simple equilibrium.[27] If both states are satisfied, then neither state can credibly threaten to use force to change the status quo. The status quo remains unaltered. If one of the states is dissatisfied, then the satisfied state offers just enough in equilibrium to the dissatisfied state to ensure that it will never find fighting worthwhile. In effect, the satisfied state appeases the dissatisfied state by offering it just enough to leave it indifferent between accepting the offer and fighting. In symbols, suppose that S_1 is the dissatisfied state. Then the satisfied state, S_2, will offer just enough, say x, so that the dissatisfied state's payoff to accepting x, $U_1(x)/(1 - \delta)$, is equal to the payoff to fighting, $(p - c_1)/(1 - \delta)$. Thus, x satisfies $U_1(x) = p - c_1$. To see that this is S_2's optimal offer, note that S_2 would never

[25] As will be seen, at most only one state can be dissatisfied.

[26] This follows from the assumptions that the states are risk neutral or risk averse, that they agree on the distribution of power, and that fighting is costly. Because the states are risk neutral or risk averse, the utility functions are concave, which implies $U_1(q) \geq q$ and $U_2(1 - q) \geq 1 - q$. If both states are dissatisfied, then $p - c_1 > U_1(q) \geq q$ and $1 - p - c_2 > U_2(1 - q) \geq 1 - q$. Adding these inequalities gives $1 - c_1 - c_2 > 1$ or $0 > c_1 + c_2$. But fighting is costly, so $c_1 \geq 0$ and $c_2 \geq 0$. Thus, $c_1 + c_2$ cannot be less than zero, and this contradiction implies that at least one state must be satisfied.

[27] The complete-information game is a straightforward modification of Rubinstein's (fn. 12) model with an outside option. For an analysis of the unique subgame-perfect equilibrium of this game, see Martin Osborne and Ariel Rubinstein, *Bargaining and Markets* (New York: Academic Press, 1990), 54–58.

offer more than the minimum needed to appease S_1 and to ensure that it will not attack. Nor would S_2 offer less than this amount. S_1 would reject a smaller offer and fight, which would leave S_2 with a payoff of $(1 - p - c_2)/(1 - \delta)$. S_2, however, prefers the payoff of offering x to S_1, which is $U_2(1 - x)/(1 - \delta)$.[28]

With complete information, bargaining never results in war regardless of the distribution of power. The satisfied state always offers the dissatisfied state just enough to appease it. By contrast, bargaining can end in war if there is incomplete information, because the satisfied state is uncertain of what is needed to appease the dissatisfied state. The more the satisfied state offers, the less likely the dissatisfied state will be to attack; but the more the satisfied state offers, the lower its payoff will be should the offer be accepted. The satisfied state must therefore balance these two effects when deciding how much to offer. Often the satisfied state will choose to accept some risk of war. With incomplete information, the relation between stability and the distribution of power is a live issue.

To introduce incomplete information formally, each state will be assumed to be uncertain of the other state's cost of fighting. In particular, S_1 is unsure of S_2's cost of fighting c_2 but believes that this cost is at least $\underline{c}_2$ and not more than $\bar{c}_2$. The value $\underline{c}_2$, which is assumed to be nonnegative, is the lowest cost of fighting that S_1 believes S_2 might have. That is, if S_2's cost is $\underline{c}_2$, S_1 is facing the toughest type of S_2. If S_2's cost is $\bar{c}_2$, S_1 is facing the weakest possible type of S_2, that is, the type of S_2 for which fighting is most costly. Similarly, S_2 is unsure of S_1's cost of fighting but believes that it is at least $\underline{c}_1$ and no more that $\bar{c}_1$. These beliefs are represented formally as probability distributions $F_1(c_1)$ and $F_2(c_2)$.[29]

Although c_1 and c_2 have been described as the states' costs of fighting, these variables also have a more general interpretation. The lower c_1, the higher S_1's payoff to fighting. Accordingly, c_1 can be interpreted more generally as a measure of S_1's willingness to use force or of S_1's resolve. That is, the lower c_1, the more willing S_1 is to use force and the greater its resolve. Thus, the game may be seen more broadly as a model of bargaining between two states that are unsure of each other's willingness to use force.

The introduction of incomplete information necessitates a more refined definition of what it means to be dissatisfied. From S_2's perspec-

[28] If S_2 preferred fighting to offering x, then $1 - p - c_2 > U_2(1 - x)$ where $p - c_1 = U_1(x)$. Adding these relations and recalling that U_1 and U_2 are concave leave the contradiction $0 > c_1 + c_2$.

[29] These cumulative distributions functions are assumed to have continuous densities that are positive over the intervals $(\underline{c}_1, \bar{c}_1)$ and $(\underline{c}_2, \bar{c}_2)$. These distributions are also common knowledge.

tive, S_1 can be any one of a continuum of types with costs ranging from $\underline{c}_1$ to $\bar{c}_1$. That is, S_2 is unsure of S_1's type where S_1's type is its cost of fighting. Similarly, S_1 is uncertain of S_2's type where S_2's type is its cost of fighting. Accordingly, a player type is dissatisfied if it prefers fighting to the status quo. So, type c'_1 of S_1 is dissatisfied if $p - c'_1 > U_1(q)$, and type c'_2 of S_2 is dissatisfied if $1 - p - c'_2 > U_2(1 - q)$. A player is potentially dissatisfied if there is some chance that it prefers fighting to the status quo. That is, a player is potentially dissatisfied if its toughest type is dissatisfied. Accordingly, S_1 is potentially dissatisfied if $p - \underline{c}_1 > U_1(q)$, and S_2 is potentially dissatisfied if $1 - p - \underline{c}_2 > U_2(1 - q)$. Paralleling the complete-information case, at most only one state can be potentially dissatisfied.

The incomplete-information game has a unique perfect Bayesian equilibrium outcome.[30] If both states are satisfied, then the outcome is trivial. Although each state is uncertain of the other's exact cost of fighting, each is sure the other's cost is so high that the other state will not fight to overturn the status quo. Thus, neither state can credibly threaten to use force to change the status quo. In equilibrium, the status quo goes unchanged, and the probability of war is zero.

The derivation of the equilibrium in the case in which one of the states is potentially dissatisfied is long and very detailed and is presented in full detail elsewhere.[31] The present discussion focuses on the central ideas underlying the proof. There are three major steps in the derivation. Assume without loss of generality that S_1 is the potentially dissatisfied state. Then the first step is to show that in equilibrium no dissatisfied type of S_1 will ever reject an offer from S_2 in order to make a counteroffer. A dissatisfied type will either accept the offer on the table or fight, depending on which alternative gives the higher payoff. Second, the fact that no dissatisfied type will ever make a counteroffer implies that all satisfied types will accept any offer larger than the status quo. The first two steps thus characterize the dissatisfied state's response to an offer from the satisfied state: If the satisfied state S_2 offers an x that is larger than the dissatisfied state's status quo share of q, then

[30] Although the outcome is unique, there are multiple equilibria because different off-the-equilibrium-path beliefs will support this equilibrium. The fact that there is a unique outcome is surprising. Typically in bargaining games in which an informed bargainer (i.e., a bargainer with private information) can make offers, there is a multiplicity of equilibrium outcomes. For an excellent introduction to bargaining models, see Drew Fudenberg and Jean Tirole, *Game Theory* (Cambridge: MIT Press, 1991). A good survey is also found in John Kennan and Robert Wilson, "Bargaining with Private Information," *Journal of Economic Literature* 31 (March 1993).

[31] Robert Powell, "Bargaining in the Shadow of Power," *Games and Economic Behavior* (forthcoming).

all types of S_1 that prefer x to fighting accept x and all other types attack. Given this response, the final step in the derivation of the equilibrium is to describe the satisfied states's optimal offer given its cost of fighting, its beliefs about S_1's cost of fighting, and S_1's response.

> Step 1. A dissatisfied type will never reject an offer in order to make a counteroffer. It will either accept the offer on the table or fight.

A preliminary to establishing this claim is to put upper bounds on the satisfied state's offers and acceptances. Let S_2 be the satisfied state and suppose that S_2 is deciding what to offer S_1 at any point in the bargaining game. S_2 will have updated its initial beliefs about S_1's cost in light of the offers S_1 has previously made. That is, S_2 will have revised its initial beliefs about S_1's cost, which are represented by the probability distribution F_1, in light of S_1's previous actions. These updated beliefs can also be represented by a new probability distribution.

Let $\hat{s}_1$ denote the toughest type of S_1 that S_2 believes it might be facing following a sequence of offers and counteroffers denoted by h_t. To put an upper bound on what the satisfied state might offer at any subsequent time, let $\hat{c}_1$ be $\hat{s}_1$'s cost of fighting. If $\hat{s}_1$ is a dissatisfied type, then the satisfied state S_2 will never offer more than what it takes to appease the toughest type that it might be facing. This is the smallest offer that ensures that $\hat{s}_1$ will not attack. Any higher offer would also ensure that S_2 would not be attacked but would mean a lower payoff if the offer were accepted. Formally, if $\hat{s}_1$ is dissatisfied, S_2 will never offer more than $\hat{x}$ where $\hat{x}$ satisfies $U_1(\hat{x}) = p - \hat{c}_1$. If $\hat{s}_1$ is satisfied, then S_2, although uncertain of the other state's exact cost of fighting, is sure that the other state is unwilling to use force to alter the status quo. In these circumstances, S_2 would never offer to revise the status quo in the other state's favor. In sum, S_2 would never offer more than the larger of $\hat{x}$ and q at any time after the initial sequence of offers and counteroffers h_t.

To put an upper bound on the demands to which S_2 might accede following h_t, two cases must also be considered. First, suppose that the toughest type that S_2 might be facing, $\hat{s}_1$, is satisfied and therefore unwilling to use force to overturn the status quo. In this case, S_2 will find any threat to use force incredible. Hence, the largest demand that S_2 might accept in this case is a demand of q, which is really a "demand" to ratify the status quo. (Because S_2 is satisfied, it is also unwilling to revise the status quo forcibly.)

In the second case, $\hat{s}_1$ is dissatisfied. In these circumstances, if S_2 rejects the dissatisfied state's demand and counters with its maximal offer of $\hat{x}$, this counter will be accepted immediately.[32] Given the immediate acceptance of $\hat{x}$, S_2's payoff to rejecting a demand in order to counter with $\hat{x}$ is the payoff to having $1 - q$ for one more period, that is, the period in which it rejects the offer, and then having the share $1 - \hat{x}$ forever. This payoff is $U_2(1 - q) + \delta U_2(1 - \hat{x})/(1 - \delta)$. Thus, S_2 would never agree to a demand from S_1 that left S_2 with a less than $U_2(1 - q) + \delta U_2 (1 - \hat{x})/(1 - \delta)$, for S_2 could do better by rejecting the demand and countering with $\hat{x}$. Accordingly, the only demands y which S_2 might accept must satisfy $U_2(1 - y)/(1 - \delta) \geq U_2(1 - q) + \delta U_2(1 - \hat{x})/(1 - \delta)$. Solving for y gives $y \leq \hat{y} = 1 - U_2^{-1}((1 - \delta)U_2(1 - q) + \delta U_2(1 - \hat{x}))$. $\hat{y}$, therefore, is an upper bound on the demands S_2 might accept when it believes that the toughest type it might be facing is $\hat{s}_1$.

Summarizing, S_2 will never offer more than the larger of q and $\hat{x}$. Nor will S_2 agree to any demand giving the dissatisfied state more than the larger of q and $\hat{y}$.

These upper bounds imply that no dissatisfied type of S_1 will reject an offer to make a counteroffer. This claim will be established by arguing by contradiction. That is, a dissatisfied type of S_1 will be assumed to reject an offer in order to make a counteroffer, and this assumption will be shown to lead to a contradiction.

Suppose there is an equilibrium in which, following a sequence of offers and counteroffers h_t which ends in an offer of, say, x from S_2, a dissatisfied type of S_1 rejects x in order to make a counteroffer of some y. Let $\hat{s}_1$ denote the toughest type of S_1 that S_2 believes it might be facing after receiving the demand y, and let $\hat{c}_1$ be $\hat{s}_1$'s cost of fighting. In equilibrium, $\hat{s}_1$ must have been willing to make the counteroffer of y; otherwise S_2 would not believe that there is some chance that it might be facing $\hat{s}_1$. But $\hat{s}_1$ would be willing to counter with y only if there were some chance that doing so would bring it at least as much as it could have had by simply attacking.

When considering whether to attack or counter with y, $\hat{s}_1$'s payoff to attacking is $W_1(0) = (p - \hat{c}_1)/(1 - \delta)$. As will be seen, the upper bounds on S_2's offers and acceptances show that the payoff to countering is always less than $W_1(0)$. The fact that $\hat{s}_1$ could have done better by attacking rather than countering contradicts the assumption that $\hat{s}_1$ rejects an offer in equilibrium in order to make a counteroffer, for no type can

[32] To see that this offer will be accepted immediately, recall that $\hat{x}$ is designed to ensure that the dissatisfied state cannot do better than accepting $\hat{x}$ by fighting instead. In equilibrium, the dissatisfied state will never reject $\hat{x}$ in order to fight. And since S_2 will never offer more than $\hat{x}$, the dissatisfied state can gain nothing by holding out for a better offer. Indeed, the dissatisfied state will lose by not reaping the benefits from a favorable shift in the status quo from q to $\hat{x}$ as soon as possible. Thus, the dissatisfied state will accept a counteroffer of $\hat{x}$ immediately.

have a positive incentive to deviate from its equilibrium strategy. This contradiction will establish the claim made in step 1.

If $\hat{s}_1$ rejects an offer to make a counteroffer, the game could end in only one of three ways following the counter. First, the game could end in war in some future period. But the payoff to a dissatisfied type to living with the status quo for a while and then fighting is strictly less than the payoff to fighting now. In symbols, $W_1(t) < W_1(0)$ for $t \geq 1$. Thus, $\hat{s}_1$ strictly prefers not to counter if the game is eventually going to end in war.

The second way that the game could end is that $\hat{s}_1$ could ultimately accept an offer from S_2. But as shown above, S_2 never offers more than $\hat{x}$. Consequently, $\hat{s}_1$'s maximum payoff to countering with y if the game subsequently ends with $\hat{s}_1$'s acceptance of an offer of $\hat{x}$ is $\hat{s}_1$'s payoff to living with the status quo for two periods—namely, the period in which $\hat{s}_1$ counters with y and then the period in which S_2 rejects this offer in order to counter with $\hat{x}$—and then having $\hat{x}$ forever. In symbols, this payoff is $U_1(q) + \delta U_1(q) + \delta^2 U_1(1 - \hat{x})/(1 - \delta)$. But this payoff is less than $W_1(0) = (p - \hat{c}_1)/(1 - \delta)$, because $U_1(\hat{x}) = p - \hat{c}_1$ and, since $\hat{s}_1$ is dissatisfied, $p - \hat{c}_1 > U_1(q)$. Again, $\hat{s}_1$ does strictly better by fighting rather than countering if the game ends with $\hat{s}_1$'s ultimately accepting an offer from S_2.

The final way that the game could end is with S_2's agreeing to a demand from $\hat{s}_1$. $\hat{s}_1$'s maximum payoff to this outcome occurs if $\hat{s}_1$ immediately counters with the maximal acceptable demand $\hat{y}$ and S_2 accepts. This would leave $\hat{s}_1$ with the payoff to living with the status quo for the period in which it rejects x to counter with $\hat{y}$ and then having $\hat{y}$ forever. In symbols, $\hat{s}_1$'s payoff is $U_1(q) + \delta U_1(\hat{y})/(1 - \delta)$. This payoff is strictly less than $\hat{s}_1$'s payoff to fighting, $(p - \hat{c}_1)/(1 - \delta)$.[33] As before, $\hat{s}_1$ does strictly better by fighting rather than countering if the game ends with S_2 accepting $\hat{s}_1$'s demand.

In sum, $\hat{s}_1$ can increase its payoff by attacking rather than making a counteroffer regardless of how the game ends following $\hat{s}_1$'s counter. Accordingly, $\hat{s}_1$ would be able to increase its payoff by deviating from its equilibrium strategy if its equilibrium strategy were to reject an offer in order to make a counteroffer. This, however, is a contradiction because the definition of an equilibrium requires that no actor can improve its

[33] To see that $\hat{s}_1$'s payoff is strictly less that $W_1(0)$, note that the definition of $\hat{y}$ implies that S_2 is indifferent to accepting $\hat{y}$ now, which leaves S_2 with $1 - \hat{y}$ forever, and settling on $\hat{x}$ in the next period, which leaves S_2 with $1 - \hat{x}$ forever. Thus $U_2(1 - \hat{y})/(1 - \delta) = U_2(1 - q) + \delta U_2(1 - \hat{x})/(1 - \delta)$. But S_2 prefers the status quo share of $1 - q$ to what it will have after settling on either $1 - \hat{y}$ or $1 - \hat{x}$. That is, $1 - q > 1 - \hat{y}$. Given $1 - q > 1 - \hat{y}$, then the previous equality implies $1 - \hat{y} > 1 - \hat{x}$ or, equivalently, $\hat{x} > \hat{y}$. Finally, the fact that $\hat{x} > \hat{y} > q$ leaves $W_1(0) = (p - \hat{c}_1)/(1 - \delta) = U_1(\hat{x})/(1 - \delta) > U_1(q) + \delta\, U_1(\hat{y})/(1 - \delta)$.

payoff by deviating from its equilibrium strategy. This contradiction thus establishes the claim that no dissatisfied type can reject an offer in equilibrium in order to make a counteroffer. If, moreover, no dissatisfied type will reject an offer to make a counter, then a dissatisfied type must choose one of the two remaining alternatives. It will either accept the offer on the table or attack, depending on which alternative yields the higher payoff.

Step 2. All satisfied types accept any offer larger than the status quo.

Assume as before that S_1 is the potentially dissatisfied state and that the satisfied state has offered to revise the status quo in favor of S_1. More formally, S_2 has offered S_1 a share x where $x > q$. As shown in step 1, a dissatisfied type of S_1 rejects x if this type's payoff to fighting is higher than the payoff to accepting x and accepts x otherwise.

Now consider the decision facing a satisfied type of S_1, which is denoted by $\tilde{s}_1$. $\tilde{s}_1$ never attacks. To see that $\tilde{s}_1$ does not attack, note that the payoff to accepting x is greater than the payoff to living with the status quo because $x > q$, and the payoff to living with the status quo is larger than the payoff to fighting because $\tilde{s}_1$ is satisfied. Thus, the payoff to accepting x is greater than the payoff to fighting, and therefore $\tilde{s}_1$ will not attack.

Given that $\tilde{s}_1$ will not attack, its decision reduces to accepting x or rejecting it in order to make a counteroffer. $\tilde{s}_1$ will accept x because it offers the higher payoff. To calculate the payoff to rejecting x, recall that no dissatisfied type of S_1 will reject x in order to make a counteroffer. If, therefore, x is rejected, this rejection is in effect a signal to S_2 that it is facing only satisfied types of S_1. But as soon as S_2 becomes convinced that it is facing only satisfied types of S_1, it will find any threat from S_1 to use force to overturn the status quo inherently incredible. Thus, S_2 will never agree to revise the status quo in S_1's favor. The status quo, therefore, will not be revised following the rejection of x. Hence, $\tilde{s}_1$'s payoff to rejecting x is the payoff to living with the status quo. But $x > q$, so the payoff to agreeing to x is higher than the payoff to living with the status quo. Accordingly, satisfied types of the potentially dissatisfied state will accept any offer $x > q$.

To see what satisfied types do if S_2 offers or, more aptly, demands a revision of the status quo in its favor, suppose $x < q$.[34] As before, all dissatisfied types of S_1 will either accept x or fight, depending on which al-

[34] The analysis of the case in which $x = q$ is completely analogous.

ternative offers the higher payoff. Indeed, in the case in which S_2 offers less than the status quo, all dissatisfied types fight. (By definition, all dissatisfied types prefer fighting to the status quo, and, with $x < q$, the payoff to accepting x is even less than the payoff to the status quo.) Because all dissatisfied types reject x and fight, S_2 will infer once again that it is facing only satisfied types of S_1 if x is countered. And, as just seen, once S_2 becomes convinced that it is facing only satisfied types, the status quo will not be changed. Thus, the choice confronting a satisfied type when $x < q$ is to attack, accept the offer, or reject it and live with the status quo. Of the three, a satisfied type prefers living with the status quo. The status quo payoff is higher than the payoff to accepting x because $x < q$ and higher than the payoff to fighting because this type is satisfied. In sum, a satisfied type rejects x in equilibrium if $x < q$ and it receives the payoff to living with the status quo.

Step 3: Specifying the satisfied state's optimal offers.

The first two steps in the derivation of the game's equilibrium describe how the types of the potentially dissatisfied state respond to an offer from the satisfied state. The satisfied state, in turn, will make its best offer in light of these responses and its beliefs about the dissatisfied state's costs. Let s_2 denote the type of S_2 with cost c_2. Then given its beliefs about the potentially dissatisfied state's costs, s_2 can determine the probability that S_1 will reject an offer of x and attack. (Assume for now that $x > q$.) To wit, the type of S_1 with cost c_1 will reject x and attack if this type is dissatisfied and prefers fighting to accepting x. In symbols, this type fights if $p - c_1 > U_1(x)$ or, equivalently, $p - U_1(x) > c_1$. But, S_2's beliefs about S_1's costs can be used to calculate the probability that this cost is less than $p - U_1(x)$ and therefore that an offer of x will lead to war. Let $R(x)$ denote the probability that S_1 will attack in response to x.

If a type of S_1 rejects $x > q$ and fights, s_2's payoff will be $(1 - p - c_2)/(1 - \delta)$. If x is accepted, its payoff is $U_2(1 - x)/(1 - \delta)$. (As shown above, x will never be countered when $x > q$.) Thus, s_2's expected payoff to offering x is the payoff if S_1 attacks in reply to x times the probability that S_1 attacks plus the payoff if x is accepted times the probability that x is accepted. Letting $P(x)$ denote this expected payoff leaves:

$$P(x) = \left(\frac{1 - p - c_2}{1 - \delta}\right) R(x) + \left(\frac{U_2(1 - x)}{1 - \delta}\right) (1 - R(x))$$

s_2, therefore, offers the value of x that maximizes $P(x)$ subject to the condition that $x \geq q$.[35]

It is now straightforward to use the results of steps 1, 2, and 3 to describe the unique perfect Bayesian equilibrium outcome of the bargaining game. Suppose that the game begins with the satisfied state's making the initial offer to the potentially dissatisfied state. The potentially dissatisfied state will respond to this initial offer as outlined above: dissatisfied types will accept it or fight, depending on which alternative yields the higher payoff, and satisfied types will accept if the offer is more than the status quo and will reject if it is less than the status quo. The satisfied state will then use its initial beliefs about the potentially dissatisfied state's costs to determine the probability that the potentially dissatisfied state will respond to a particular offer by attacking. And given these probabilities, the satisfied state will make its optimal offer, which is the value of x that maximizes $P(x)$ for $x \geq q$ and where the probability that an initial offer of x is rejected is $R(x) = F_1(p - U_1(x))$.[36] The probability that the bargaining will end in war is, in equilibrium, simply the probability that the potentially dissatisfied state will reply to this optimal initial offer by attacking.[37]

Stability and the Distribution of Power

The equilibrium of the bargaining model makes it possible to calculate the probability of war as a function of the distribution of power. This section calculates this relation in a specific case and then compares these results to the expectations of the balance-of-power and preponderance-of-power schools and to other formal work. The section also offers a possible explanation for the equivocal results of empirical efforts to determine the relation between stability and the distribution of power and briefly considers some of the difficulties of estimating this relation empirically. Finally, the section discusses the relation between the probability of war and the distribution of power in more general cases. As in the specific case analyzed in detail here, the expectations of

[35] The expression for $P(x)$ was derived on the basis of the assumption that $x > q$. If $x \leq q$, s_2's expected payoff to offering x is $P(q)$. Accordingly, s_2 is indifferent to offering any $x \leq q$ and it will suffice to maximize P over the range $x \geq q$.

[36] s_1 rejects x if $p - c_1 > U_1(x)$ or $p - U_1(x) > c_1$. Given that c_1 is distributed according to the cumulative distribution function F_1, the probability that c_1 is less than $p - U_1(x)$ is $F_1(p - U_1(x))$.

[37] Although cumbersome to show, the probability that the bargaining will end in war is essentially the same regardless of whether the satisfied state or the potentially dissatisfied state makes the initial offer. See Powell (fn. 31) for the details.

both the balance-of-power and the preponderance-of-power schools fail to hold in these more general cases.

To simplify the analysis, assume that the states are risk neutral and that each believes the other's cost of fighting is uniformly distributed between zero and one. Risk neutrality and the previous normalizations of the states' utility functions imply that S_1's and S_2's utilities for the division $(x, 1 - x)$ are simply x and $1 - x$, respectively.

The probability of war turns out to be a simple function of the disparity between the distribution of power and the status quo distribution of territory. More specifically, the absolute value of the difference between the distribution of power and the status quo distribution of territory, $|p - q|$, is a measure of the disparity between these two distributions. Letting $\Pi(p,q)$ denote the probability of war as a function of p and q, the appendix then shows:

$$\Pi(p,q) = \begin{cases} |p-q| - |p-q|^2 & \text{if} \quad |p-q| \leq 1/2 \\ 1/4 & \text{if} \quad |p-q| \geq 1/2 \end{cases}$$

These results seem quite intuitive. When the distribution of power mirrors the status quo, that is, when $p = q$, neither state expects to gain by using force to revise the status quo. The probability of war in this case is zero. In symbols, $\Pi(p,q) = 0$ when $p = q$. As the disparity between the distribution of power and the status quo distribution grows, the probability of war rises until it levels off at 1/4.[38]

Figure 2 plots the relation between the probability of war and the distribution of power for three different values of the status quo distribution. S_1's status quo share of territory is 1/3 if $q = 1/3$; 1/2 if $q = 1/2$; and 2/3 if $q = 2/3$. These results clearly contradict the expectations of both the balance-of-power and the preponderance-of-power schools. As illustrated in Figure 3a, the preponderance-of-power school expects the probability of war to be smallest when S_1 or S_2 is predominant (that is, when p is large or small) and to increase as the distribution of power becomes more even (that is, as p approaches 1/2). But in the case analyzed here, the probability of war is smallest at $p = q$ and generally largest when one of the states is preponderant.

[38] The specific value of 1/4 at which the function levels off is a result of the particular assumption that the states' costs are uniformly distributed between the bounds of zero and one. Different assumptions about these bounds would cause the function to level off at a different value.

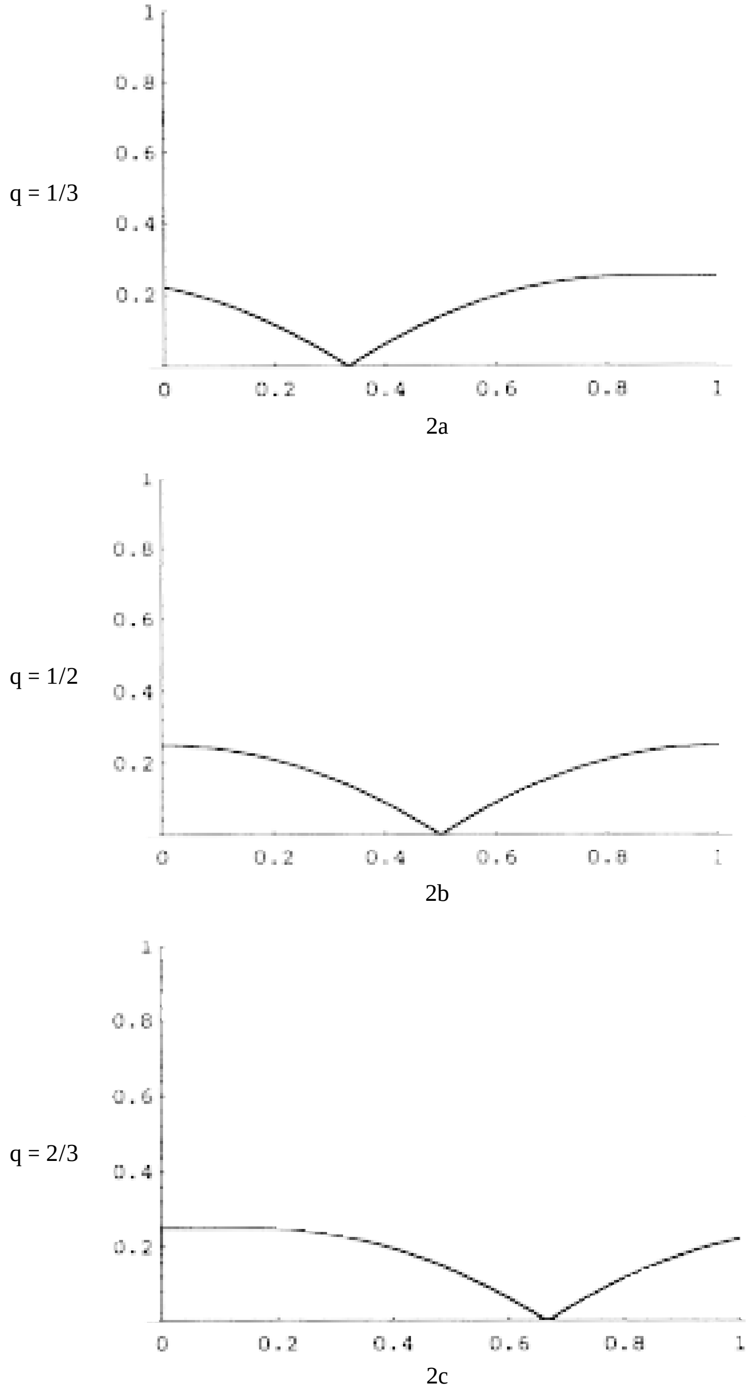
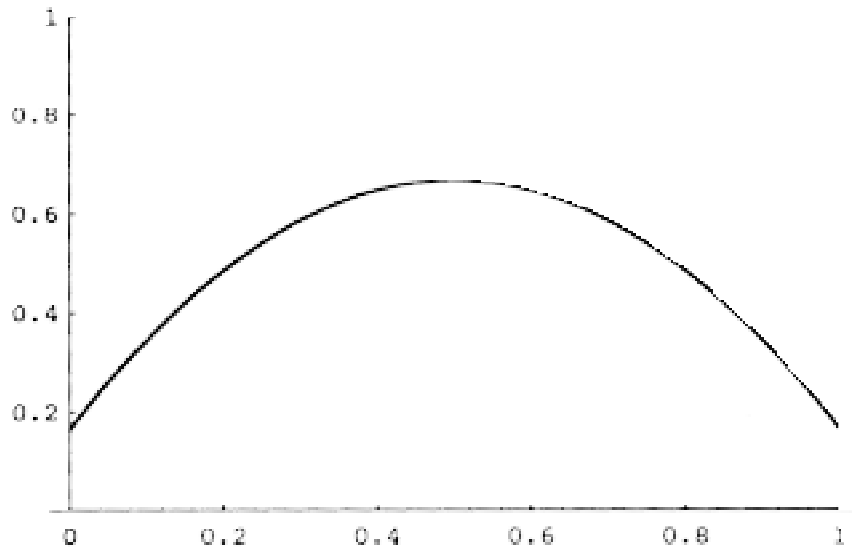
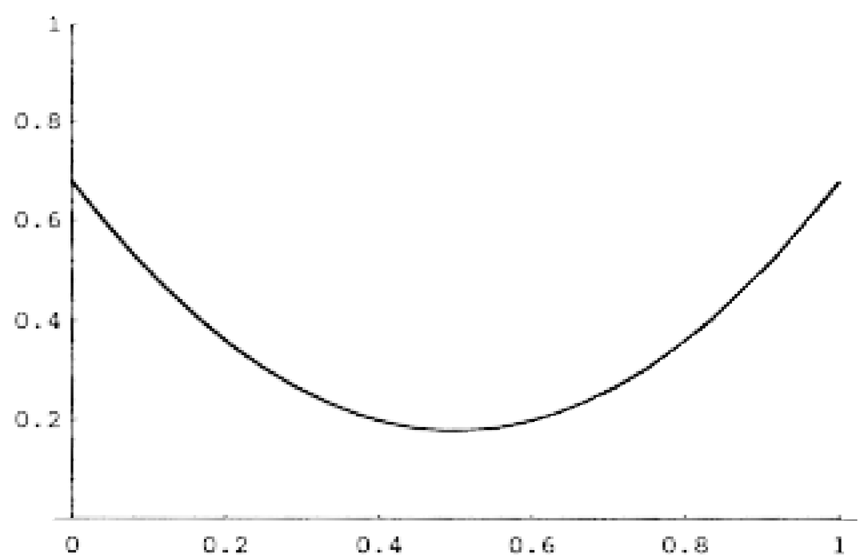


FIGURE 2
THE PROBABILITY OF WAR AND THE DISTRIBUTION OF POWER



(a) PREPONDERANCE OF POWER



(b) BALANCE OF POWER

FIGURE 3
EXPECTATIONS OF THE TWO SCHOOLS

Figure 3b illustrates the balance-of-power school's expectations. The probability of war is smallest when there is a roughly even distribution of power (p is approximately 1/2) and increases as the distribution of power becomes skewed. These expectations correspond to the formal results only in the special circumstance in which the status quo division is even ($q = 1/2$). When the status quo division is uneven, as in Figure 2a and 2c, the probability of war in the model is smallest when the distribution of power mirrors the uneven status quo division and not, as the balance-of-power school asserts, when the distribution is even.

The present analysis also complements and qualifies the results of other formal analyses of the relation between stability and the distribution of power. Bueno de Mesquita and Lalman, for example, study a model in which two states bargain about revising the status quo.[39] In their game each state has the opportunity to make a single demand of the other state. After both states have made their demands, each state decides whether or not to use force to secure its demand. In one version of this game Bueno de Mesquita and Lalman assume that the size of each state's demand is determined outside the model. That is, the amount each state demands is simply one of the exogenously specified initial conditions of the model. In this version of Bueno de Mesquita and Lalman's model, the size of a state's demand is not determined as part of a state's equilibrium strategy. A state does not determine how much it demands by balancing the risk of war against the chance of agreeing on a more favorable settlement. The justification for assuming the states' demands to be exogenous is that these demands are "determined by internal political rules, procedures, norms, and considerations and may or may not be attuned to foreign policy considerations."[40] With the size of the states' demands determined outside the model, Bueno de Mesquita and Lalman find that "dissatisfaction with the status quo is entirely unrelated to the likelihood of war."[41]

The present analysis focuses specifically on how a state balances the greater risk that a larger demand will lead to war against the higher payoff the state will obtain if this larger demand is accepted. This focus requires a model in which the states determine the sizes of their de-

[39] Bueno de Mesquita and Lalman (fn. 2, 1992).

[40] Ibid., 41.

[41] Ibid., 190. In a second version of their game, Bueno de Mesquita and Lalman let the state's demands be determined as part of their equilibrium strategies. But this game has complete information. The combination of endogenous demands and complete information means that the probability of war is zero in their model—as well as in the model analyzed above—regardless of the distribution of power. Bueno de Mesquita and Lalman do not examine the case in which there are both endogenous demands and incomplete information.

mands as part of their equilibrium strategies. The game examined here satisfies this requirement. The probability of war in the present game, unlike in Bueno de Mesquita and Lalman's model, is closely related to the level of dissatisfaction with the status quo.[42] The larger the level of dissatisfaction in the case above, that is, the larger the disparity $|p - q|$, the more likely war.[43]

Fearon also considers the relation between the probability of war and the distribution of power. He studies a game in which one state makes a take-it-or-leave-it offer after which a second state must then either accept this offer or attack. By assumption, this second state cannot make a counteroffer. The size of the state's demand in Fearon's model is determined as part of the state's equilibrium strategy, and he finds that the probability of war is "*completely insensitive*" to the distribution of power.[44]

The current model complements Fearon's formulation. He is studying a stylized situation in which one state presents another state with an ultimatum in the form of a military fait accompli. This fait accompli establishes a new distribution that will become the new international status quo unless the second state goes to war to reverse it.[45] In the stylized situation modeled in this essay, two states are bargaining about revising the status quo and have the opportunity to make offers and counteroffers. There are, however, no military faits accomplis: any use of force to revise the status quo is assumed to lead directly to war. In these circumstances the probability of war is sensitive to the distribution of power and, in particular, to the disparity between the distribution of power and the status quo distribution of territory.

The present analysis provides a possible explanation of the equivocal findings of empirical efforts to determine the relation between stability and the distribution of power. Empirical research has yielded at least three different findings about an even distribution: that it is more stable than a preponderance of power,[46] that it is less stable than a preponderance of power,[47] and, depending on whether one looked at the nineteenth or the twentieth century, that it is either more or less stable.[48]

[42] Bueno de Mesquita and Lalman (fn. 2, 1992), 190.

[43] Strictly speaking, the probability of war increases as $|p - q|$ increases only as long as $|p - q| < 1/2$. After this, the probability of war levels off.

[44] Fearon (fn. 4), 20.

[45] Ibid., 15.

[46] Siverson and Tennefoss (fn. 2).

[47] Kim (fn. 2, 1992); and Moul (fn. 2).

[48] Singer, Bremer, and Stuckey (fn. 2).

Other work has found no significant relation between stability and the distribution of power.[49]

The equilibrium of the bargaining game shows that the probability of war depends on both the distribution of power and the status quo distribution. Thus, any attempt to assess the relation between the probability of war and the distribution of power should control for the status quo. Failing to do so will generally lead to biased estimates if, as might be expected, the distribution of power and the status quo distribution are correlated. But empirical efforts to estimate this relation have generally not controlled for the status quo, and this omission may partially account for the equivocal empirical results.[50]

Finally, it is important to ask about the generality of these results. Do the findings hold only in the specific example considered here in which the states were assumed to be risk neutral and to believe that each other's cost or willingness to use force was uniformly distributed? Or do the results also hold in more general conditions? The central finding that the probability of war is at a minimum when the distribution of power mirrors the status quo distribution is quite robust. It holds as long as the states are risk neutral or risk averse, regardless of the particular shapes of their utility functions or of the particular shapes of the probability distributions representing their beliefs.[51] Thus, the model contradicts the expectations of the balance-of-power and preponderance-of-power schools as long as the states are risk neutral or risk averse.[52]

One might also wonder if this result holds if the uncertainty sur-

[49] Maoz (fn. 2); and Bueno de Mesquita and Lalman (fn. 2, 1992).

[50] Unfortunately, the theoretical results derived here, while indicating the importance of controlling for the status quo, do not identify a means of doing so. The basic problem is to find a way of measuring a state's utility for the status quo. Bueno de Mesquita and Lalman's (fn. 2, 1992) measure of this utility may offer a start in this direction, but their current formulation is inadequate. If the states are risk neutral as in the example above, then Bueno de Mesquita and Lalman's measure reduces to assuming that the status quo distribution is constant in all cases. In terms of the example above, they are in effect assuming that q always equals 1/2. (To obtain $q = 1/2$, assume the state is risk neutral by taking $r_A = 1$ in equation A1.3 in Bueno de Mesquita and Lalman (fn. 2, 1992), 294, and normalize the state's utility to be one if it obtains everything it demands and zero if the other state obtains everything it demands by setting $U^A(\Delta_A) = 1$ and $U^A(\Delta_B) = 0$.) Controlling for the status quo, however, requires the status quo to be treated as a variable across cases, and it is unclear how to do this with their measure as it is currently formulated.

[51] To see this, suppose that the distribution of power closely mirrors the status quo distribution of territory in that $|p - q|$ is less than the minimium of the lowest possible costs of fighting, i.e., the minimum of $\underline{c}_1$ and $\underline{c}_2$. Then, $p - q \leq \underline{c}_1$ and $q - p \leq \underline{c}_2$. These inequalities and the concavity of U_1 and U_2 imply $p - \underline{c}_1 \leq q \leq U_1(q)$ and $1 - p - \underline{c}_2 \leq 1 - q \leq U_2(1 - q)$. Thus, both S_1 and S_2 are satisfied, and the probability of war is zero.

[52] The case of risk-acceptant states introduces nonconvexities and technical difficulties, and the analysis of this case remains a task for future work.

rounding the states' willingness to use force varies with shifts in the distribution of power. More specifically, the costs of fighting in the example examined here are always uniformly distributed between 0 and 1 regardless of the distribution of power. But what would happen if the expected cost and uncertainty about a state's willingness to use force declined as that state became more powerful?[53]

This question requires a two-part response. First, the question itself takes the analysis beyond the balance-of-power and predominance-of-power schools, as neither school bases its argument on the subtleties implicit in this question. Second, the model provides a partial answer to this question. The probability of war is still zero as long as the status quo distribution of benefits mirrors the distribution of power.[54] Once again, the expectations of the balance-of-power and preponderance-of-power schools fail to hold.[55]

CONCLUSION

The relation between the distribution of power and the probability of war is an important and long-debated problem in international relations theory. The balance-of-power school argues that an even distribution of power brings greater stability, while the preponderance-of-power school argues that a preponderance brings greater stability. These opposing claims have been examined in the context of an infinite-horizon bargaining game in which two states alternate making offers about how to revise the international status quo. The bargaining continues until the states agree on a revision or until one of them becomes sufficiently pessimistic about the prospects of reaching a mutually agreeable settlement that it resorts to force in an attempt to impose a new settlement. Contradicting the expectations of both the balance-of-power and preponderance-of-power schools, the probability of war in the model is smallest when the distribution of power mirrors the status quo distribution.

[53] More formally, suppose that the distributions of cost F_1 and F_2 were also a function of p and that the mean and variance of F_1 decrease in p while the mean and variance of F_2 decrease in $1 - p$.

[54] Even if F_1 and F_2 depend on p, the argument in footnote 26 goes through with only minor modification.

[55] The model's answer is only partial because the shape of the function relating the probability of war to the distribution of power, while always zero when $p = q$, is likely to depend on the states' utility functions and on the functional forms of the probability distributions representing their beliefs.

Appendix

There are two cases to consider in calculating the probability of war as a function of the distribution of power. Suppose, first, that $p > q$. Then S_1 is the potentially dissatisfied state, because the expected payoff to fighting for the type of S_1 with the lowest cost of fighting is strictly greater than its status quo payoff: $p - \underline{c}_1 = p - 0 > q$. Then, as shown above, the probability of war in the bargaining game is the probability that the dissatisfied state attacks in response to the satisfied state's optimal initial offer.

To calculate this probability, S_2's optimal initial offer must first be determined. Suppose S_2 offers x to S_1. A dissatisfied type s_1 of S_1 with cost c_1 will reject this offer and attack if its payoff to fighting is higher than its payoff to accepting. That is, s_1 attacks if $p - c_1 > x$ or, equivalently, if $p - x > c_1$. Thus, the probability that S_1 will attack in response to an offer of x is the probability that S_1's cost is less than $p - x$. Given that S_1's costs of fighting are uniformly distributed between 0 and 1, the probability that S_1's cost is less than $p - x$ is just $p - x$. Consequently, S_2's expected payoff to offering x, which will be denoted by $P(x)$, is its payoff to fighting times the probability that S_1 will attack plus the payoff if S_1 accepts x times the probability that S_1 will accept x. In symbols,

$$P(x) = \frac{1 - p - c_2}{1 - \delta}(p - x) + \frac{(1 - x)}{1 - \delta}(1 - (p - x))$$

if S_2's offer is between q and p. If S_2 offers more than p, then even the dissatisfied state with the lowest cost of fighting will prefer to accept the offer to fighting. Consequently, the probability that an offer of $x > p$ will be rejected is 0. S_2's payoff to offering x in this case reduces to $(1 - x)/(1 - \delta)$. If S_2 offers $x < q$, then all dissatisfied types fight and all other types reject the demand which leave the status quo unchanged. S_2's payoff is:

$$\frac{1 - p - c_2}{1 - \delta}(p - q) + \frac{(1 - q)}{1 - \delta}(1 - (p - q)).$$

S_2's optimal offer is the value of x that maximizes its expected payoff. To describe the solution to this maximization problem, set the derivative of $P(x)$ equal to zero and let $x^*(c_2)$ denote the solution to this equation. This leaves $x^*_1(c_2) = p - (1 - c_2)/2$. Then the optimal offer of the type of S_2 with cost c_2 is $x^*(c_2)$ as long as $x^*(c_2) \geq q$. S_2's optimal offer is q if $x^*(c_2) < q$.

Now that the equilibrium offers have been specified, the probability of breakdown can be calculated. Let $\pi(c_2)$ be the probability that the type of S_2 with cost c_2 will be attacked. If $x^*(c_2) < q$ for this type, then this type will offer q. The probability that the dissatisfied state will attack in response to this offer is the probability that the cost of fighting c_1 is less than $p - q$ which, given the uniform distribution of costs, is simply $p - q$. Thus, $\pi(c_2) = p - q$ for all types of S_2 for which $x^*(c_2) < q$ or, equivalently, for all types of S_2 for which $c_2 < 1 - 2(p - q)$.

If $x^*(c_2) \geq q$ for a type of S_2, then this type will offer $x^*(c_2)$. The probability that the dissatisfied state attacks in response to this offer is $p - x^*(c_2) = p - (p - (1 - c_2)/2) = (1 - c_2)/2$. Thus, $\pi(c_2) = (1 - c_2)/2$ whenever $x^*(c_2) \geq q$ or $c_2 \geq 1 - 2(p - q)$.

$\pi(c_2)$ specifies the probability that the particular type of S_2 with cost c_2 will fight. But from the perspective of an outside observer, as well as from that of S_1, S_2's cost is uncertain. Thus the probability of war is the expected value of $\pi(c_2)$. Letting $\Pi(p,q)$ denote this expected value, then:

$$\Pi(p,q) = \int_0^1 \pi(c_2)dc_2$$

$$\Pi(p,q) = \begin{cases} (p - q) - (p - q)^2 & \text{if} \quad 0 \leq p - q \leq 1/2 \\ 1/4 & \text{if} \quad 1/2 \leq p - q \end{cases}$$

$\Pi(p, q)$ describes the relation between the probability of breakdown and the distribution of power if $p > q$ and, consequently, S_1 is the dissatisfied state. If $p < q$, then S_2 is the dissatisfied state and the relation between stability and the probability of war must be determined anew. Fortunately, the symmetry of the problem makes it easy to find this relation in this second case. Let $p' = 1 - p$ and $q' = 1 - q$. Then $p < q$ implies $p' > q'$. Moreover, p' now measures the probability that the dissatisfied state will prevail and q' measures the dissatisfied state's status quo payoff just as p and q did in the first case where $p > q$. The symmetry of the example then ensures that the probability of breakdown if $p <$

q is given by $\Pi(p', q')$. But $\Pi(p', q')$ is just $\Pi(1 - p, 1 - q)$ which in turn equals $(q - p)(1 - (q - p))$ if $1 - 2(q - p) \geq 0$ and $1/4$ if $1 - 2(p - q) < 0$.

Combining these results leaves:

$$\Pi(p, q) = \begin{cases} |p - q| - (p - q)^2 & \text{if} \quad |p - q| \leq 1/2 \\ \\ 1/4 & \text{if} \quad |p - q| \geq 1/2 \end{cases}$$