Are Sanctions Effective?

A GAME-THEORETIC ANALYSIS

GEORGE TSEBELIS
University of California at Los Angeles

Although economic sanctions have been quite frequent in the twentieth century, a close examination of the low success rate (33 out of 83 cases) indicates that sender countries are not able to select the appropriate cases. Moreover, analysts sometimes offer contradictory advice for such selection. This article provides a game-theoretic explanation of these phenomena. Six different game-theoretic scenarios lead to the same equilibrium outcome. This is a mixed strategy equilibrium. The success ratio is the outcome of the selection of mixed strategies by both sender and receiver countries. Under a wide range of (specified) circumstances, the size of the sanction has no impact upon the behavior of the target country. Finally, some empirical implications of the game-theoretic analysis are compared to existing empirical generalizations, and further implications for empirical research are discussed.

Dating as far back as 432 B.C., when Pericles enacted his Megarian decree prohibiting Megarians to trade or travel on Athenian land, economic sanctions have been an important ingredient of foreign policy making. In recent history, sanctions have been applied for military purposes, to destabilize foreign governments, protect human rights, and retaliate against terrorist activities. They have been applied collectively by actors such as the League of Nations and the OPEC countries, and unilaterally by individual nation states such as the United States, Soviet Union, United Kingdom, and Canada. In the post-1960 period, sanctions were used as frequently as two or three times per year.

One would expect that such a long and dense experience would provide accurate answers to the questions of whether sanctions work, and if so, under

AUTHOR'S NOTE: Financial support for this project was provided by UCLA's Academic Senate. I would like to thank R. Bates, J. Frieden, J. Grieco, D. Lake, L. Martin, E. Rogers, and S. Telhami for many useful ideas, comments, and bibliographical references.

JOURNAL OF CONFLICT RESOLUTION, Vol. 34 No. 1, March 1990 3-28

what conditions. Answers to these questions would have decreased the frequency of ineffectively applied sanctions: The sender countries would have been able to anticipate the results of their actions and avoid costly, useless, and humiliating economic hostilities.

Nevertheless, the question of the effectiveness of sanctions remains a question that elicits many contradictory answers. A quick look at the two most recent and important books on sanctions illustrates this contradiction. Baldwin (1985: 3) tries to answer the puzzle, "Why do statesmen continue to practice economic statecraft when 'everybody knows' that it does not work?" Hufbauer and Schott (1985), in an impressive empirical study, indicate that in 83 incidents involving economic sanctions after 1914, the success rate was a poor 40%. Therefore, it is clear that either we do not know how to count success and failure (this is Baldwin's position) or that sender countries have not learned to make a successful selection of cases for the application of sanctions. This is the reason Hufbauer and Schott provide nine "commandments" for the architects of sanction policies.

The purpose of this article is to investigate the problem of authority and compliance in an anarchic environment. What are the conditions under which states will abstain from available profitable alternatives because of fear of retaliation by other states? Do economic sanctions work? Under what conditions does their success rate increase? Why are sender countries slow to learn the conditions for successful sanctions, as indicated by the present low success ratio? Finally, why, as we will see, do analysts give contradictory prescriptions?

The article assumes states to be unitary and rational actors. This is a common (although unrealistic) assumption in the relevant literature. Efforts to study theoretically the interconnection between domestic and international politics are very recent (see Putnam, 1988). The reason I do not follow this more realistic path is to keep the argument as clear and simple as possible and to show that we can answer these questions without necessarily referring to domestic politics. However, in several places I discuss the impact of domestic politics or of other players in the sanctions game. I first examine six different ways of conceptualizing the problem of sanctions as a game between the sender and the target country. These games make completely different assumptions about the players: They operate under complete or incomplete information, they have perfect rationality or adaptive behavior, they move simultaneously or sequentially, and they have discrete or continuous options. Regardless of the differences in the assumptions, however, all six scenarios lead to the same equilibrium outcome. Next, I examine the properties of the common equilibrium of these games. Then, the conditions under which sender countries will apply sanctions regardless of their impact, or target countries will disregard sanctions independently of their severity, are analyzed. A structural kind of uncertainty is introduced: One or both countries do not know the "type" of their opponent (whether the opponent is "soft" or "tough"). The third part introduces domestic politics and problems of international cooperation, and compares existing empirical results concerning sanctions with the predictions of the game-theoretic model, explains the low success rate of sanctions, and gives the reasons for policy prescriptions that are poor and sometimes contradictory.

SIX SCENARIOS IN SEARCH OF AN EQUILIBRIUM VALUE

According to Galtung's (1967) definition, sanctions are "[a]ctions initiated by one or more international actors (the 'senders') against one or more others (the 'receivers') with either or both of two purposes: to punish the receivers by depriving them of some value and/or to make the receivers comply with certain norms the senders deem important." The opinion of the academic community in the late 1960s and early 1970s was that economic sanctions are largely ineffective. Galtung (1967: 409) concluded that "the probable effectiveness of economic sanctions is, generally, negative." Doxey (1972: 547) claimed that "the deterrent and coercive force of sanctions is weak on almost every count." Wallensteen (1968: 262) argued that the "general picture is that economic sanctions have been unsuccessful as a means of influence in the international system"; and for Adler-Karlsson (1968: 9), "the overall conclusion that the described embargo policy has been a failure" is inescapable. In addition, Baldwin (1985) quotes several official reports that reached the same conclusion.

There is, however, a discrepancy between the beliefs of the academic community and policies with respect to economic sanctions. Indeed, during the same period, the number of sanction incidents rose from 5 in the 1965-70 period, to 13 between 1970 and 1975, to 22 between 1975 and 1980, and then dropped to 11 between 1979 and 1984 (see Hufbauer and Schott, 1985: 25). This increase is not correlated with effectiveness. In fact, the success ratio before 1973 was almost 45% while after 1973 it dropped to less than one-third. The opinions of both academics and policy makers oscillate between the belief that sanctions are ineffective on the one hand, and that they can have successful policy outcomes on the other. This belief reflects real differences in the impact of sanctions (alas, with a difference of phase of 180 degrees so that these opinions are always in opposition to the facts).

One reason for such diffused and contradictory beliefs is the incredible amount of noise that aggregate-level data contain: uncertainty with respect to how to define whether or not a sanction was successful, how to measure the size of the penalty, how to measure the impact on the economy of both the sender and the target country, and how to calculate the dependence of the target in an interdependent economic system where other countries may fill the gap in economic relations and trade. It seems that each sanctions incident is unique and, therefore, any generalization is impossible.

Compelling as this argument may seem, it did not prevent policy analysts from trying to control for different variables, classifying sanctions into different categories, and providing policy prescriptions for successful sanctions policies. We will do the same from a theoretical point of view. But first, let us develop several stylized scenarios of sanctions. Such scenarios are abstractions and simplifications of real situations, and each one of them will present unrealistic or undesirable features. However, the diversity of the assumptions and the unity of the conclusions of these models will persuade the reader that the results reported here are valid under a wide range of conditions.

Let us assume that there are two extreme options for the receiver (from now on, "target" country): either to violate a law, rule, norm, or standard that is of material and/or normative importance to the sender country, or to comply with it. These extreme options will be termed "violate" and "comply." On the other hand, the sender country can choose between the following two extremes: either to "sanction" at maximum capacity or to "not sanction." Table 1 represents the payoffs to each player in each one of the four possible extreme outcomes of the game. Table 1 is the general payoff matrix that will hold for all the possible variations of the games.

It is reasonable to assume that if the sender will not sanction, the target would prefer not to comply with the standards. In algebraic terms,

$$\text{ASSUMPTION 1: } b_1 > d_1$$

On the other hand, it is reasonable to assume that sanctions also inflict a cost on the sender country. It would then be preferable for the sender not to sanction if none of her interests were hurt. In algebraic terms,

$$\text{ASSUMPTION 2: } d_2 > c_2$$

On can also assume that sanctions at maximum capacity have a deterrent effect, and that the target country prefers to avoid sanctions rather than to violate the standard and be sanctioned. I will relax this assumption in the next section, but for now,

$$\text{ASSUMPTION 3: } c_1 > a_1$$

TABLE 1
General Payoff Matrix of Sanctions Game: Six Scenarios

	Sanction	*No Sanction*
Violate	a_1 a_2	b_1 b_2
Comply	c_1 c_2	d_1 d_2

NOTE: Assumptions are: $b_1 > d_1$, $d_2 > c_2$, $c_1 > a_1$, and $a_2 > b_2$.

Finally, the sender country prefers to react and sanction when its interests are violated rather than remain inactive. Again, in the next part, this assumption will be put under scrutiny, but for the time being,

ASSUMPTION 4: $a_2 > b_2$

Having established the order of the different payoffs for each player, we can concentrate on the different scenarios for economic sanctions.

Scenario 1: Complete information, rationality, continuous choices, and simultaneous moves. Both the sender and the target country know each other's payoffs, and they are unified, perfectly rational players. Each one has the option of choosing the *level* of his strategy. The target country has to decide the level x of its violation (where x = 1 means complete violation and x = 0 means complete compliance). The sender country has to decide the level of sanctions y (where y = 0 means no sanctions and y = 1 means sanctions at maximum capacity). When x or y are 0 or 1, the payoffs for each country are given by Table 1. To write the payoffs of each player for each pair of strategies, one additional and important assumption has to be made.

ASSUMPTION 5: The payoffs of each player are linear functions of the strategies of both players.[1]

Assumption 5, together with the payoffs from Table 1, gives the following utilities for each player as a function of the strategies chosen by each:

$$u_1 = (d_1-c_1-b_1+ a_1)xy + (c_1- d_1)y + (b_1- d_1)x + d_1 \qquad [1]$$

$$u_2 = (d_2- c_2-b_2+ a_2)xy + (b_2- d_2)x + (c_2- d_2)y + d_2 \qquad [2]^2$$

Equations (1) and (2) enable us to calculate the equilibrium strategies of the sanctions game. Equilibrium strategies are defined as a pair of strategies x* and y* that are optimal responses to each other. Indeed, if the players

1. The importance of this assumption will be discussed at the end of this article. It is sufficient here to say that different assumptions would have produced the same or similar outcomes.

2. The reader can verify that for the extreme strategies where x and y are 0 or 1, functions (1) and (2) have the values specified by Table 1. Because functions (1) and (2) are linear with respect to x and y, they are the *only* linear functions with this property.

choose any other pair of strategies, one of them will have the incentive to change his strategy. The other will also modify his strategy in response, generating an infinite cycle of responses.

The only equilibrium pair of strategies for the sanctions game described by (1) and (2) is given by the equations:[3]

$$x^* = (d_2 - c_2) / (d_2 - c_2 + a_2 - b_2) \quad [3]$$

$$y^* = (b_1 - d_1) / (b_1 - d_1 + c_1 - a_1) \quad [4]$$

It can be shown that the equilibrium calculated by equations (3) and (4) presents all the desirable properties of stability required in game theory. In particular, it can be shown to be regular (Harsanyi, 1973), perfect (Selten, 1975), essential (Wu Wen-tsun and Jiang Jia-he, 1962), and proper (Myerson, 1978).[4]

Scenario 2: Perfect information, rationality, discrete choices, and simultaneous moves. Consider now that all the assumptions made in scenario 1 still hold, except for one: Each country only has the two extreme options available. The target country either can "violate" or "comply" with the standard, and the sender country either can "sanction" or "not sanction." No one pair of these strategies is a mutually best response, and the game presented in Table 1 has no pure strategy equilibria. It follows that the only possible equilibrium of the discrete sanctions game is in mixed strategies. The calculation replicates exactly the previous scenario: Each country, instead of calculating the optimum *level* of its strategy, calculates the optimum *frequency* of mixing its two pure strategies. The calculations lead to exactly the same equilibrium as equations (3) and (4): The target country violates the standards with frequency $p^* = x^*$, while the sender country sanctions with frequency $q^* = y^*$ (see Luce and Raiffa, 1957).

Baldwin (1985) argues that the continuity of strategies leads observers to wrong inferences because they have to interpret whether sanctions work or not, while in fact each country does not decide in a discrete scenario but in a continuous one, and sets levels of sanctions or of violations. From our two stylized scenarios it becomes clear that although continuity of strategies may create problems of counting, there is no essential conceptual difference between the continuous and the discrete strategy scenarios. One can move from the one to the other, translating levels into frequencies and vice versa, so in the remaining examples I will drop the distinction and use only the easier case for expositional purposes.

3. For the derivation, see Ordeshook (1986: 131). Technically these equilibrium strategies are computed by setting $\partial u_1/\partial x = 0$ and $\partial u_2/\partial y = 0$.

4. For the proofs of these theorems, see van Damme (1984).

Scenario 3: Perfect information, rationality, continuous choices, and sequential moves. Suppose now that the payoffs for the choice of the extreme strategies are the ones in Table 1, and that all the remaining assumptions hold, but that the two players move sequentially rather than simultaneously. First, the target country decides how much to violate the standard (decides x), and then the sender country decides how much to punish (decides y). From equations (1) to (4), it is clear that if the target country decides to violate the standard by more than x*, then it is in the interest of the sender country to inflict maximum punishment (respond with y = 1), which would make the target country regret its choice. On the other hand, if the target country chooses any x less than x*, the best response of the sender country is to ignore the violation (set y = 0), which will make the target country want to increase its level of violation. From this account, it follows that no value less than or greater than x* can be an equilibrium strategy, and therefore, the equilibrium of the game is given by (3) (the sender can choose any response).

Scenario 4: Adaptive behavior, alternating moves. Consider now that the payoff matrix is as presented in Table 1, but that no country is a perfectly rational player. Both countries demonstrate adaptive behavior and alternate moves. Suppose that in the beginning, the target country chooses level x of violation of the standard, and the sender country adopts a level of sanction y. The payoffs of the interaction are given by equations (1) and (2). Then the target country modifies its behavior to the value of the punishment. Subsequently, the sender country modifies the level of punishment to agree with the level of the offense. Then it is the turn of the target to adjust, and so on. These mutual adjustments are made proportionally to the difference between the current payoff and the maximum possible payoff, *given* the strategy of the opponent.

It is easy to calculate that the optimum strategy for the target country is to violate the standard at the maximum (x = 1) whenever the sender country applies sanctions at a level less than y* [of equation (4)], and to comply with the standard completely (x = 0) whenever the sender country applies sanctions more than y*. Similarly, the optimum strategy of the sender country is to "sanction" at the maximum level (y = 1) whenever the target country violates the standard at a level greater than x* [of equation (3)], and to "not sanction" (y = 0) whenever the target country violates the standard at a level less than x*.

In Appendix A, I show that this process can be formalized by the following differential equations:

$$dx/dt = k(1-x)\,[(c_1 - d_1 + b_1 - a_1)y - (b_1 - d_1)] \text{ if } y < y^* \qquad [5]$$

and $$dx/dt = -kx[(c_1 - d_1 + b_1 - a_1)y - (b_1 - d_1)] \text{ otherwise} \qquad [5']$$

$$dy/dt = -1y[(d_2 - c_2 - b_2 + a_2)x + (c_2 - d_2)] \text{ if } x < x^* \qquad [6]$$

and $$dy/dt = 1\,(1 - y)\,[(d_2 - c_2 - b_2 + a_2)x + (c_2 - d_2)] \text{ otherwise} \qquad [6']$$

where k and 1 are positive constants.

In simple words, under the assumptions A1-A4, if the sender country sanctions, the target country will comply (A3); if the target country complies with the standards, the sender will stop sanctioning (A2); if the sender stops sanctioning, the target will violate the standards (A1); if the target violates the standards, the sender will sanction (A4); if the sender sanctions, the target will stop violating the standards (A3); and so on. No matter which one of the extreme combinations of strategies results from the choices of the two players, one player will have the incentive to modify her choice. I described this cycling process between sanctions and no sanctions, violation of standards and compliance, in extreme terms. Equations (5) and (6) describe smaller mutual adjustments. But the major question is, Where will this process of mutual adjustments equilibrate?

In Appendix A, I prove that the unique equilibrium of this system of differential equations is also given by (3) and (4). The interpretation of this finding is the following: The heroic assumptions of perfect rationality are not required in order to arrive at the same equilibrium in the sanctions game. Even adaptive, myopic behavior leads to the same outcome. This conclusion is very important because the argument can be made that states are not unified actors, and consequently competing coalitions inside each state will try to adopt different policies in order to solve the sanctions game. These different policies can then be modeled as a "tattonnement" process, where each country tries to find solutions that work under the circumstances—that is, *given* what the opponent has done so far.

Scenario 5: Both-sided incomplete information, rationality, discrete choices, and simultaneous moves. To model incomplete information, I will consider that there is random noise in the payoffs of each player, and when they choose, they know their own payoffs but not the payoffs of the opponent. As I will argue in the next section, this noise in the payoffs may be the result of either domestic politics or events relating to international economic competition. Consider the game of Table 2, where $0 < e_1 < 1$, $0 < e_2 < 1$, and x and y are independent and identically distributed, each with uniform distribution over the interval from 0 to 1. When the game is played, player 1 (the target) knows the value of x but not of y, and player 2 (the sender) knows

TABLE 2
Payoff Matrix with Two-Sided Incomplete Information

	Sanction		*No Sanction*	
Violate	$a_1 + e_1x$	$a_2 + e_2y$	$b_1 + e_1x$	b_2
Comply	c_1	$c_2 + e_2y$	d_1	d_2

NOTE: Assumptions are: $b_1 > d_1$, $d_2 > c_2$, $c_1 > a_1$, $a_2 > b_2$; e_1 and e_2 small positive numbers; x, y drawn from uniform distributions in the [0,1] interval.

the value of y but not of x. If e_1 and e_2 are 0, then the game in Table 2 becomes exactly the same as the game with complete information in Table 1. Thus, let us assume that e_1 and e_2 are very small positive numbers. Then x and y can be interpreted as minor factors influencing the players' payoffs when the strategies "violate" and "sanction" are chosen.

In Appendix B, I show that when e_1 and e_2 tend to 0, the unique equilibrium of the game with incomplete information is again given by equations (3) and (4). The interpretation of this finding is straightforward: Even under incomplete information, as the perturbations of the payoffs of each player are reduced or as each player learns more about his opponent (so that e tends to 0), the equilibrium strategies are the same as the game under perfect information.

Scenario 6: One-sided incomplete information, rationality, discrete choices, and simultaneous moves. This is a particular case of the previous game, where the payoffs of one of the opponents are known by both players, while the payoffs of the other player are subject to random shocks. Using the same assumptions as the previous scenario concerning e_1 or e_2, and the random variable x (if the payoffs of the target country are unknown by the sender) or y (if the payoffs of the sender country are unknown by the target), the equilibrium of the game tends to (3) and (4) when e tends to 0 (see Appendix B for the proof).

Six different scenarios led to the same equilibrium. Some assumed simultaneous, others sequential moves; some assumed perfect rationality, others simple adaptive behavior; some perfect information, and others incomplete information by one or both sides; and in some scenarios the countries had simple dichotomous choices, in others a continuum of strategies was available. The convergence of all these models to the same equilibrium should be interpreted as an indication of the robustness of this equilibrium to different plausible specifications of the sanctions problem. It is time now to examine the properties of this equilibrium.

THE ROBINSON CRUSOE FALLACY

The first and probably most important observation from equations (3) and (4) is that the strategy of neither the target nor the sender countries depends on their own payoffs; it depends on the payoffs of the opponent! Strange as this result might seem, it has been confirmed by the analysis of each one of the six sanction scenarios. Whether the strategies are interpreted as discrete or as continuous, and whether x or y represent level or frequency of sanctions or of violations, the result remains the same. The strategy of each player depends exclusively on the payoffs of the opponent. Let me single out this result in the form of two theorems:

THEOREM 1. Under assumptions 1-4, modification of the payoffs of the target country leaves the level or the frequency of violation of the standards (x^*) unchanged. On the contrary, most of the time, it has an impact on the severity or frequency of sanctions (y^*) imposed by the sender country.

THEOREM 2. Under assumptions 1-4, modification of the payoffs of the sender country leaves the level or the frequency of sanctions (y^*) unchanged. On the contrary, most of the time, it has an impact on the level or the frequency of violations (x^*) of the target country.

The proof of both these theorems is very easy. Inspection of equation (3) indicates that x^* does not depend on the payoffs of the target country, while equation (4) indicates a monotonic relationship between the payoffs of the target country and the equilibrium strategy of the sender country.[5] For example, as long as assumptions 1-4 hold, an increase in the level of maximum sanctions (a_1 in our model) will have the impact of reducing not the level or the frequency of violation of the standard, but the level or severity of sanctions by the sender country.

Although the mathematics of this statement are straightforward, it still flies against our intuition. Therefore, what is needed more than the analytical explanation is an intuitive explanation of theorems 1 and 2. Are they mathematical artifacts? If not, why are they so counterintuitive? Why does the conventional wisdom expect that a change in the payoffs of a player induce modification of his or her behavior, while six different game-theoretic models produce the opposite result?

Conventional wisdom examines the problem of sanctions from the perspective of one player only: the target country. It ignores the fact that sanctions as well as violations are the outcome of the interaction of two players—that is, of a game between rational players and not a simple decision against nature. The fallacy of applying decision theory instead of game theory

5. The monotonic relationship can be shown by testing the sign of the first derivatives with respect to the payoffs.

when more than one rational actor is involved I have called the Robinson Crusoe fallacy (see Tsebelis, 1989). This fallacy leads to wrong conclusions, such as the expectation that modification of the incentives of one player will modify his behavior, while in reality it modifies the behavior of the opponent. The reason that conventional wisdom leads to such mistaken results is that: (1) it considers simple decision problems where there is only one decision maker, and (2) it confines itself to short-run analysis. Indeed, as scenario 4 indicates, it is very plausible that, in the short run, as sanctions increase compliance will increase. However, once the sender country realizes this change in the behavior of the target country, it will modify its own strategy—that is, reduce the severity of sanctions—and the target country will respond with further modifications, . . . and the new equilibrium will be the one described by equations (3) and (4), where modifications of the payoffs of the target country have no effect upon its behavior.

There are several corollaries of theorems 1 and 2. To decrease the frequency of violations (increase the frequency of compliance), a modification of the payoffs of the sender country is required: an increase of a_2 (make "sanctions" easier for the sender country to apply) or a decrease in b_2 (make "not sanctions" more difficult once a standard is violated), or a decrease in d_2 (the value for the sender of harmonious relations between itself and the target country) or an increase of c_2 (the value of wrongly applied sanctions). To decrease the frequency of sanctions, a modification of the payoffs of the target country is required: an increase of d_1 (the value for the target country of harmonious relations between itself and the sender country) or a decrease of a_1 (an increase in the cost of sanctions to the target country), or a decrease of b_1 (the value of unpunished violations) or an increase of c_1 (the value of wrongly applied sanctions).

Case 1: Combination of assumptions 1-4. This is the most frequent and most interesting case of the sanctions game. It is possible, however, that modifications of the payoffs of one player modifies the initial assumptions of the model. In particular, while it is always reasonable to assume that if the sender does not sanction, the target prefers to violate the standard rather than to comply (assumption 1), and that if there is no violation the sender prefers not to sanction rather than to sanction (assumption 2), assumptions 3 and 4 are questionable. One can imagine, for example, that sanctions may not provide sufficient incentive for the target country to modify its behavior, or that the sender country prefers not to sanction even when its interests are violated. To these points we now turn.

Case 2: Violation of assumption 3, ($c_1 < a_1$). In this case, the target country prefers to violate the standard no matter what the reaction of the sender country, and the latter will sanction or not sanction according to whether or not $a_2 > b_2$. If $a_2 > b_2$, we will observe unsuccessful sanctions. If, however, $a_2 < b_2$, as I suspect has been the case quite often, no event will be observed. At best, a diplomatic complaint will be the only visible sign of discomfort demonstrated by the sender country.

Case 3: Violation of assumption 4, ($a_2 < b_2$). In this case, the sender country will never sanction. As a result, target countries will always violate the sender's interests (assumption 1). I believe that this is always the case with small sender countries; that is why there are no cases of sanctions where the sender is a small country (see Hufbauer and Schott, 1985).

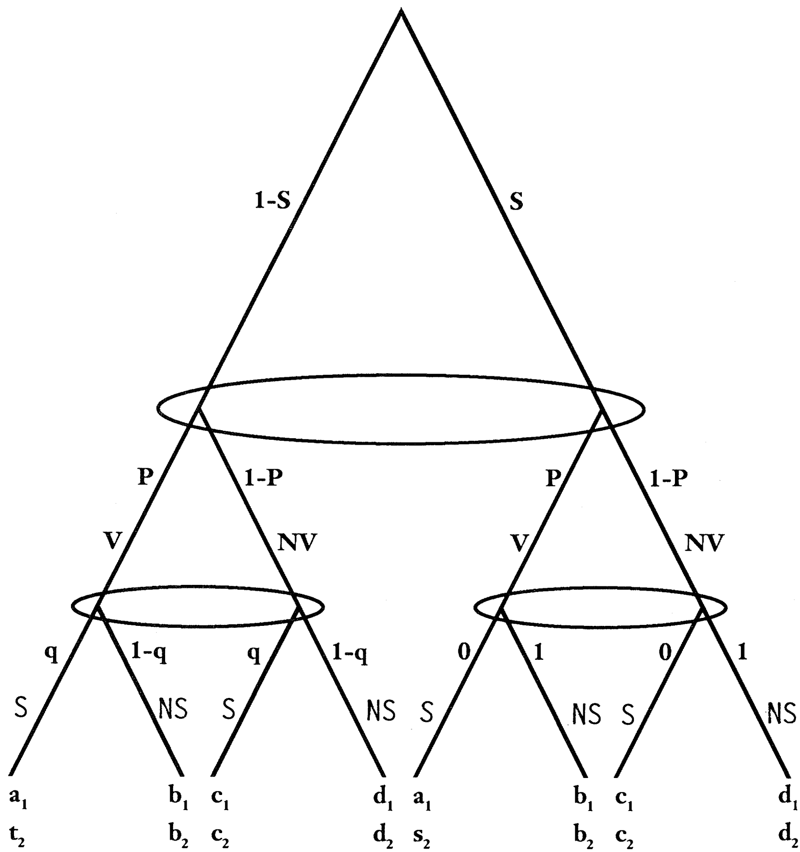
More realistic, however, is the case of a different kind of incomplete information than the one investigated in scenarios 5 and 6.[6] What happens if one or both of the players does not know in which case she finds herself? For example, what if the target country is not sure whether it is in the interest of the sender country to sanction—that is, it cannot discriminate between cases 1 and 3? Or what happens when the sender country does not know whether sanctions would make the target country modify its behavior—that is, the sender country cannot discriminate between cases 1 and 2? Or if combined uncertainty makes the actors believe that they could be in any one of the three cases?

Figure 1 represents a case where the target country does not know whether the sender country is "soft" (that is, never sanctions) or "tough" (that is, prefers to sanction if its interests are violated). The probability of the sender country being soft is s, and this is common knowledge. The countries have to decide their strategies simultaneously, while the target country does not know the type of the sender country (although it knows the probability s). In Appendix C, I examine this case of structural uncertainty, and demonstrate that the two essential propositions of this part—theorems 1 and 2—still hold. The equilibrium strategies of both countries are calculated, and the arguments for why theorems 1 and 2 hold for even more complicated cases of uncertainty are given. The mathematics become more complicated; the functional forms are different from equations (1) and (2), but the essential conclusion remains: Modification of the size of the sanction does not affect the behavior of the target country at equilibrium; it affects the frequency of sanctions.

FROM GAME THEORY TO THE REALITY OF SANCTIONS

Next, I will draw some inferences from the game-theoretic model of sanctions and compare them with the findings of the most massive empirical

6. I thank Shibley Telhami for pointing out this case of "structural," as he called it, uncertainty to me.



Assumptions: $c_1 > a_1$ $b_1 > d_1$
$t_2 > b_2 > s_2$ $d_2 > c_2$

Figure 1

study of sanctions, Hufbauer and Schott's *Economic Sanctions Reconsidered* (1985). I will also draw additional inferences that may be useful for further empirical research. What happens if the sanction potential of the sender country increases? Or what happens if the target country is weak? Or how should the argument that sanctions are noninstrumental and purely symbolic be treated?

TABLE 3
Possible Orders of Payoffs, and Corresponding Outcomes in Sanctions Game

Name	*Conditions*	*Outcome*
Case 1	$b_1 > d_1$ $d_2 > c_2$ $a_2 > b_2$ $c_1 > a_1$	Mixed strategies
Case 2	$b_1 > d_1$ $d_2 > c_2$ $a_1 > c_1$	Violation; sanction if $a_2 > b_2$
Case 3	$b_1 > d_1$ $d_2 > c_2$ $a_2 > b_2$	No sanctions

At this point, looking back at the conditions of cases 1, 2, and 3 of the previous section will be very useful for the subsequent development of the argument. The purpose of Table 3 is to recapitulate the different conditions and outcomes of the sanctions game.

It is time to introduce domestic politics in each one of the countries, and to assess its impact on the sanctions problem. I have examined elsewhere situations in which one actor is involved simultaneously in games in several arenas. I have argued that such cases can be modeled as nested games, where the situation in one arena influences the payoffs of the actors in the principal arena (see Tsebelis, forthcoming, a). According to this "nested games" approach, if one considers the sanctions game as the principal arena, and domestic politics in each country or other international factors such as economic competition between sender countries as the secondary arena, one can take into account the impact of domestic politics or international economic competition on the sanctions game. For example, intransigence on the part of the opposition, or public opinion, may transform the payoffs of a government and make it willing to impose sanctions although it knows very well that they will not be instrumental with respect to the target country. Similarly, international competition may prevent a country from sanctioning another if it knows that its sanctions will not be sufficient to induce the target country to change its behavior. All these contextual factors (with respect to the sanction game) can be taken into account through corresponding changes in the payoffs of the country involved in these nested games. Let us study such factors one at a time.

The interaction between domestic and international politics can be the supporting story for Figure 1. A moderate party competes with an extremist party for power in the sender country, and the outcome of this struggle is uncertain. The moderate party will follow a "soft" line in international politics, while the extremist will be "tough" (for operational definitions of these terms, see Figure 1). The target country has to decide whether to violate a standard without knowing which political line will prevail in the sender country. Or, similarly, an international conference concerning probable sanctions against a target country is held. This conference can succeed, in which case the sender countries will be "tough," or the conference can fail, in which case they will be "soft" and tolerate violations of a standard. The target country has to decide without knowing the outcome of the conference, and the conference has to decide without knowing the decision of the target country. The models presented in this article can help us investigate the strategies of the countries involved in such incidents.

First, let us concentrate on the target country. As noted earlier, it has an interest in violating the standards if they are not going to be enforced by sanctions (A1). Let us now assume that the sender country is small or weak or that it has little leverage, unable to inflict a sufficiently serious penalty on the target country. Or, alternatively, let us assume that the sender country is strong, yet faces other countries that are in economic or political competition with it, so that when the sender applies economic sanctions other countries step in to alleviate the situation for the target country. Or, finally, consider that there is strong opposition against sanctions in the sender country, so that imposing sanctions is politically costly for the government. Under these assumptions, the situation is described as in case 2 in Table 2. The target country will always violate the standard at the maximum level regardless of the reaction of the sender country. The sender country will then either apply unsuccessful sanctions or acquiesce.

But why would a country apply sanctions if it knows the payoffs of both players, and therefore knows that sanctioning will have no result? Such behavior can be explained if other potential violators are taken into account. The sender country may want to use the sanctions in the specific case not because of their effectiveness, but to send a signal to other small countries that similar behavior on their part will not go unpunished.[7]

If, however, the strength of the sender country increases sufficiently, then the situation will be described by case 1 in which the target country will apply a mixed strategy: In a series of cases, the target country will violate the standard some of the time and comply the rest of the time. Similarly, the sender country will apply a mixed strategy, sometimes choosing either to sanction or to not sanction the violator. From this point on, a further increase of the sanction potential of the sender country will *not* have any deterrent effect on the target country. The frequency of violation of the standard will remain unchanged. What will change as the sanction potential increases is the frequency with which sanctions are applied by the sender country. In fact,

7. It requires iterated games with incomplete information. See Kreps and Wilson (1982), and for an application in international relations and the problem of hegemony, see Alt et al. (1988).

as equation (4) indicates, this frequency will be reduced at equilibrium. And as Appendix C indicates, this logic remains the same, even in the case where the sender or the target country do not know whether they are facing a "soft" or a "tough" opponent.

Let us now turn to the sender country. As demonstrated earlier, it will not sanction if no standard is violated (A2). Assume that there is a violation but with such small consequences that the cost of sanctions is higher than the inconvenience from the violated standard. In this case, the sender country will accept the violation of the standard without reaction. The situation is described by case 3 in the previous section in which the observation was made that the sender country has a dominant strategy of no sanction. There are two reasons why this situation might occur: Either the violation is small or the costs of sanctions are very large.

However, if the inconvenience of the sender country increases sufficiently (if it exceeds the costs of sanctioning), the situation will be described by case 1, in which both countries will apply mixed strategies. In a series of incidents, the target country will sometimes violate and sometimes refrain from violation, while the sender country will sometimes sanction and sometimes refrain from sanctioning. The frequencies of the respective strategies are given again by equations (3) and (4). From this point on, any further increase in the importance of the standard will have no impact on the behavior of the sender country. On the contrary, other things being equal, target countries will reduce the frequency of important violations, thinking twice before violating important standards of important senders.

At this point, I have finished the theoretical investigation, and some confrontation with political situations is in order. I will consider the most extensive empirical study of sanctions, which covers the whole universe of sanctions incidents since the First World War (Hufbauer and Schott, 1985; from now on, H-S), and examine its findings in light of this game-theoretic model. The H-S study summarizes its empirical findings in the form of "nine commandments."

Commandment 1 recommends avoiding high policy goals. Indeed, high policy goals were achieved only 16% of the time. Such an outcome is due to the fact that high policy goals occur when the violated standard is considered of vital importance by the target country, in which case it is likely that assumption A4 is violated and the target country prefers to violate the standards regardless of sanctions (case 2 in the previous section).

Commandment 2 recommends attacking the weak. Empirically, there is a correlation between economic weakness and susceptibility to economic sanctions. This commandment is based on the idea that sanctioning strong opponents will lead to case 2; the opponent may be so strong that sanctions do not have a serious impact (violation of A4) and so she will continue to violate the standard.

Commandment 3 recommends attacking allies. The reasoning is similar. Enemies will have structured their economy in such a way that sanctions are not effective upon them; again the situation is described by case 2.

Commandment 4 asks for no incremental application of sanctions. Strictly speaking, there is no time component in the game-theoretic model presented. Therefore, this rule cannot be accounted for directly. However, the reasoning behind it is not time-related. Commandment 4 tries to disallow time for the economy of the target country to adapt to the new situation. In this sense, applying sanctions slowly is equivalent to applying ineffective sanctions: sanctions with costs lower than the benefits of violation of the standard (violating assumption A4), leading again to case 2.

I argue that commandments 5 and 7 are contradictory, which accounts for their discrepancies. That is why they must be examined together. Commandment 5 recommends maximum economic sanctions, while 7 claims that additional military policies are ineffective. It seems a country wants to maximize economic sanctions because it wants to maximize the impact of sanctions on the target country. If this logic is correct, there is no reason why economic sanctions should not be accompanied by diplomatic, military, or other kinds of action. Additional pressure, no matter what its source and nature, should increase the effectiveness of the project. Thus, 5 and 7 are contradictory. However, if we consider the game-theoretic model, this contradiction is explained. Increasing sanctions is likely to transform the game from case 2 to case 1. However, once we are in case 1, any further increase in sanctions has no impact on the equilibrium strategy of the target country, as question (3) indicates. The only impact of higher sanctions is to reduce the frequency of sanctions. That is why military actions do not have any impact. Because application of military action indicates that economic sanctions have already been applied at their maximum force, we are well inside case 1. Consequently, additional sanctions have no impact on the strategy of the opponent (theorems 1 and 2).

Commandment 6 recommends that senders with high costs of sanctions not enter the game. High costs to the sender country places us in case 3 of the sanctions game; therefore, sanctions in this case are a mistake. If, however, costs are barely lower than benefits, the sanctions game is described by case 1, and the frequency of violation will be high as indicated by equation (3).

Commandment 8 claims that effectiveness declines with the number of sender countries. Because only one sender country is assumed in my analysis, this is again outside the framework of our model. However, it is a classic collective-action problem. Each sender country prefers to free ride on the sanctions of the others. Therefore, the overall impact of sanctions is reduced up to the point of switching from case 1 to case 2 where violation of standards is the dominant strategy for the target country.

Commandment 9 links sanctions with broader economic and strategic considerations; however, it does not provide any specific instructions.

These nine commandments were addressed to sender countries. The commandments try to shift the situation from case 2 to case 1. However, even if they were followed to the letter they would not guarantee success, because in case 1 there is *always* a percentage of unsuccessful sanctions specified by the probabilities of the equilibrium strategies of the two opponents. Let us use commandment 1 to demonstrate this point.

Commandment 1 asks the sender country to lower policy goals in order to shift from case 2 to case 1. However, lowering policy goals will not mean automatic success. In case 3, mixed strategies are the only equilibrium strategies. Therefore, we are likely to have cases of nonsanctions as well as cases of unsuccessful sanctions. Moreover, as equations (3) and (4) indicate, lowering the policy goals will have no effect on the frequency of violation by the target country but will reduce the frequency of sanctions by the sender country.

The reason why policy analysts and policy makers in industrialized countries are so slow to learn the conditions of successful sanctions is that they are committing the Robinson Crusoe fallacy. They assume they play against nature while they actually play against rational opponents—that is, against rational actors themselves trying to promote their own goals. Even if all necessary information was collected, there still would be an irreducible part generated by the fact that both parties find themselves most of the time in case 1, resulting in a situation in which mixed strategies must be used. So, in addition to the nine commandments of H-S, which are addressed to the Western policy maker, there are several commandments for the target country that lead to violations without sanctions or to nonviolations (if sanctions are imminent in case of violation). Such cases are not included in the H-S list for the obvious reason that it is very difficult to collect nonevents. The model presented earlier can help us generate the conditions that would push the sanctions game from case 1 to case 3, in which the outcome is no sanctions. The exercise will not be solved here; however, policy analysts and policy makers of target countries follow precisely this strategy. Thus, one of the reasons that sanctions have such a low success rate is that ineffective sanctions are the goal of other rational actors (the target countries).

The argument of selection bias in cases of deterrence has been made convincingly by Achen and Snidal (1989), and attempts at more rigorous empirical study of deterrence are now underway (see Russett and Huth, 1989). Exactly the same logic applies to sanctions. Suppose that some social scientist persuaded by these arguments does her best to collect cases of nonsanctions and add them to the H-S list in order to eliminate the selection bias they present. An empirical study of this extended list would provide much more complete, unbiased, and impartial insights into the sanctions phenomenon.

One natural use of these data for the empirically minded social scientist would be to provide the definitive answer on the effects of sanctions. The standard way to answer this question would be to compare the frequency of compliance with the standards when sanctions are either present or absent. If there is a statistically significant difference between these two frequencies, the conclusion must be that sanctions are effective.

Now assume that we are in the equilibrium situation described by equations (3) and (4) of our model. What would the empirical findings look like? The frequency of compliance with the standards would be exactly the same in either the presence or absence of sanctions. The data would look like random noise! The conclusion would be that there is no impact, yet we know from the previous analysis that those frequencies exist in the first place because sanctions do make a difference. Why is there conflict between these ideal[8] data and theoretical analysis?

There is no discrepancy: The empirical data look like random noise because that is precisely what they are. In fact, each one of the players took precautions to randomize his strategy (or to follow any of the other five scenarios that lead to the same outcome) in order to avoid exploitation by his opponent.

This last theoretic expectation generated by my model has important consequences for both the empirical testability of the model and empirical research. How does one test a model that expects the data to look like random noise? And if one believes the underlying theory, how can one claim that this theory is the real reason that the data look like random noise? There are two points to be made. First, the only way to discriminate between the expectations of my model and the high level of noise existing in empirical data is to design situations where external noise is reduced: laboratory experiments, as opposed to historical cases. Second, empirical tests are not the only possible tests of a theory. Other criteria are excess content, plausibility of assumptions, and congruence with other theories (see Lakatos, 1970). The theory I present has excess content over other theories because it explains why, contrary to expectations, military sanctions do not have any additional impact. Moreover, the assumptions of my model, despite their simplicity, are plausible, and the conclusions hold under reasonable complications.

8. In fact, so ideal as to be imaginary.

However, throughout this article, there were two important simplifications. First, the payoffs of each actor were considered to be independent of the payoffs of the other. Second, even when domestic politics or international events were considered through the nested games approach, the unitary actor assumption was never relaxed. The new actors influenced the payoffs of the two principal actors; they were not considered as capable of strategic action themselves. Such complications will increase greatly the realism of the model.

APPENDIX A

Proof that scenario 4 leads to the same equilibrium:

For reasons of convenience, I repeat equations (1) and (2) of the main text, which give the utilities for each country when they adopt strategies x and y, respectively.

$$u_1 = (d_1 - c_1 - b_1 + a_1)xy + (c_1 - d_1)y + (b_1 - d_1)x + d_1 \quad [1]$$

$$u_2 = (d_2 - c_2 - b_2 + a_2)xy + (b_2 - d_2)x + (c_2 - d_2)y + d_2 \quad [2]$$

Both countries demonstrate adaptive behavior and alternate moves. When the sender country applies sanctions at a level less than y^* [of equation (4)], the optimum strategy for the target country is to violate the standard at the maximum ($x = 1$). In this case, the optimum payoff is given by (1) of the main text by substituting 1 instead of x.

$$\max u_1 = (-b_1 + a_1)y + b_1 \quad [1A]$$

When the sender country applies sanctions greater than y^*, the optimum strategy for the target country is to comply with the standard completely ($x = 0$). In this case, the optimum payoff is given by (1) of the main text by substituting 0 instead of x.

$$\max u_1 = (c_1 - d_1)y + d_1 \quad [2A]$$

Similarly, when the target country violates the standard more than x^*, the optimal strategy for the sender is to sanction at maximum capacity ($y = 1$). In this case, the optimum payoff for the sender is given by (2) of the main text by substituting 1 instead of y.

$$\max u_2 = (a_2 - c_2)x + c_2 \quad [3A]$$

Finally, when the target country violates the standard less than x^*, the optimal strategy for the sender country is to ignore the violation ($y = 0$). In this case, the optimum payoff is given by (2) of the main text by substituting 0 instead of y.

$$\max u_2 = (b_2 - d_2)x + d_2 \quad [4A]$$

The mutual adjustments are made proportionately to the difference between the current payoff and the maximum possible payoff, *given* the strategy of the opponent. Consequently, from (1), (2), and (1A) – (4A) the difference between the levels of violation of the standard by the target country in time t and t+1 is:

$$dx/dt = k(1-x)\,[(d_1-c_1-b_1+a_1)y + (b_1-d_1)] \text{ if } y < y^* \qquad [5A]$$

and $$dx/dt = -kx[(d_1-c_1-b_1+a_1)y + (b_1-d_1)] \text{ otherwise} \qquad [5A']$$

Similarly, the difference between the levels of sanctions by the sender country in time t and t+1 is:

$$dy/dt = -ly[(d_2-c_2-b_2+a_2)x + (c_2-d_2)] \text{ if } x < x^* \qquad [6A]$$

and $$dy/dt = l(1-y)\,[(d_2-c_2-b_2+a_2)x + (c_2-d_2)] \text{ otherwise} \qquad [6A']$$

where k and l are positive constants.
The system of differential equations (5A) to (6A′) is called Lotka Voltera equations in the biology literature (see May, 1963, and Hirsch and Smale, 1974), and they have not yet been solved in their general form. It is relatively easy, however, to calculate the equilibrium of the system if we set the left-hand side of the equations equal to 0.

The unique equilibrium of the system is the one calculated by equations (3) and (4) of the main text. To exclude other possible candidates, one can observe two things: (1) pairs where one of the variables is 0 or 1 are not in equilibrium, because one of the equations (5A) to (6A′) is not 0, and so in the next round the variable that was 0 or 1 will have a different value; and (2) pairs where both variables x and y are 0 or 1 are not in equilibrium, because again, one of the relevant equations is not 0. So, the system of (5A) to (6A′) has a unique equilibrium, and this equilibrium is calculated by equations (3) and (4).

QED

APPENDIX B

Proof that scenario 5 leads to the same equilibrium:

Consider the game of Table 2, where $0 < e_1 < 1$, $0 < e_2 < 1$, and x and y are independent and identically distributed, each with uniform distribution over the interval from 0 to 1. When the game is played, player 1 (the target) knows the value of x but not of y, and player 2 (the sender) knows the value of y but not of x. For every pair of e_1 and e_2 there is a unique equilibrium of the game, which is given by inequalities (1B) and (2B).
If

$$x > [e_1(c_1-a_1) - (b_1-d_1+c_1-a_1)(a_2-b_2)] \,/\, [e_1^2 + (b_1-d_1+c_1-a_1)(d_2-c_2+a_2-b_2)] \qquad [1B]$$

play "violate," otherwise play "comply."
If

$$y > [e_2(a_2-b_2) - (a_2-b_2+d_2-c_2)(c_1-a_1)] / [e_2^2+(b_1-d_1+c_1-a_1)(d_2-c_2+a_2-b_2)] \quad [2B]$$

play "sanction," otherwise play "nonsanction."
PROOF. To calculate these inequalities, we reason as follows: Each player has to maximize his own payoffs, given that he observes the random variable affecting his payoffs but not the payoffs of the opponent. Call p(x) the probability that player 1 (the target) will play "violate" and q(y) the probability that player 2 (the sender) will play "sanction." Player 1 has to choose "violate" when this is the action maximizing his payoffs, no matter what the payoff of his opponent. So, for a value of x greater than some specified value x_0, player 1 will choose "violate." Similarly, player 2 will choose "sanction" for values of y greater than some specified value y_0.
The expected utility of player 1 is given by

$$EU_1 = (a_1+e_1x)pQ + (b_1+e_1x)p(1-Q) + c_1(1-p)Q + d_1(1-p)(1-Q) \quad [3B]$$

where

$$Q = \int_0^1 q(y)dy \quad [4B]$$

EU_1 becomes maximum when $\partial EU_1/\partial p$ is 0, or equivalently when

$$b_1-d_1 + e_1x - (b_1-d_1+c_1-a_1)Q = 0 \quad [5B]$$

So the strategy for player 1 is: Play "violate" when x is such that the left-hand side of (5B) is greater than 0, and "comply" otherwise.
The expected utility of player 2 is given by

$$EU_2 = (a_2+e_2y)Pq + (c_2+e_2y)q(1-P) + b_2(1-q)P + d_2(1-P)(1-q) \quad [6B]$$

where

$$P = \int_0^1 p(x)dx \quad [7B]$$

EU_2 becomes maximum when $\partial EU_2/\partial q$ is 0, or equivalently when

$$c_2- d_2 + e_2y + (a_2- c_2+d_2 -b_2)Q = 0 \quad [8B]$$

So the strategy for player 2 is: Play "sanction" when y is such that the left-hand side of (8B) is greater than 0, and "not sanction" otherwise.
However, because p(x) = 1 when x is greater than the value x_0 calculated from (5B) and 0 otherwise, the integral

$$P = \int_0^1 p(x)dx = 1 - x_0 \quad [9B]$$

Similarly, because $q(y) = 1$ when y is greater than the value y_0 calculated from (8B) and 0 otherwise, the integral

$$Q = \int_0^1 q(y)dy = 1 - y_0 \quad [10B]$$

Substitution of P and Q from (9B) and (10B) to (5B) and (8B) gives a linear system of two equations and two unknowns (x_0 and y_0). The solution of this system is presented by equations (1B) and (2B). It is easy to verify that when $e_1 \rightarrow 0$ and $e_2 \rightarrow 0$, the x_0 and y_0 tend to the following values:

$$x_0 \rightarrow (a_2 - b_2) / (d_2 - c_2 + a_2 - b_2) \quad [11B]$$

$$y_0 \rightarrow (c_1 - a_1) / (b_1 - d_1 + c_1 - a_1) \quad [12B]$$

Equations (11B) and (12B) indicate the limit frequency that each player chooses each one of his pure strategies. Because the distribution of x and y is uniform, x will be greater than x_0 exactly $(1 - x_0)$ of the time. Similarly, y will be greater than y_0 exactly $(1 - y_0)$ of the time. Therefore, "violate" will be chosen with frequency $(1 - x_0)$ and "sanction" with frequency $(1 - y_0)$. The reader can verify that these are the same frequencies as the ones calculated in equations (3) and (4).

QED

Proof that scenario 6 leads to the same equilibrium:

This is a special case of the previous proof, when either e_1 or e_2, but not both, is 0, while the other tends to 0. Let us assume that $e_1 = 0$, and repeat the steps of the previous proof. To facilitate comparisons, the corresponding equations are numbered the same way.

$$EU_1 = a_1pQ + b_1p(1-Q) + c_1(1-p)Q + d_1(1-p)(1-Q) \quad [3B]$$

where

$$Q = \int_0^1 q(y)dy \quad [4B]$$

EU_1 becomes maximum when $\partial EU_1/\partial p$ is 0, or equivalently when

$$b_1 - d_1 - (b_1 - d_1 + c_1 - a_1)\, Q = 0 \quad [5B]$$

So the strategy for player 1 is: Play "violate" when the left-hand side of (5B) is greater than 0, and "comply" otherwise.

The expected utility of player 2 is given by

$$EU_2 = (a_2+e_2y)pq + (c_2+e_2y)q(1-p) + b_2(1-q)p + d_2(1-p)(1-q) \quad [6B]$$

EU_2 becomes maximum when $\partial EU_2/\partial q$ is 0, or equivalently when

$$c_2 - d_2 + e_2y + (a_2 - c_2 + d_2 - b_2)\,Q = 0 \quad [8B]$$

So the strategy for player 2 is: Play "sanction" when y is such that the left-hand side of (8B) is greater than 0, and "not sanction" otherwise.

However, because $q(y) = 1$ when y is greater than the value y_0 calculated from (8B) and zero otherwise, the integral

$$Q = \int_0^1 (y)dy = 1-y_0 \quad [10B]$$

Substitution of Q from (10B) to (5B) and (8B) gives a linear system of two equations and two unknowns (p and Y_0). The limit values of these variables when $e_2 \to 0$ leads to the equilibrium described by equations (3) and (4).

QED

APPENDIX C

Calculation of the equilibrium strategies in Figure 1:

The assumptions are that the sender country is "soft" (with probability s) or "tough" (with probability 1 – s), and that this is common knowledge. So with probability s, the two players play the right-hand part of the game tree, for which $c_1 > a_1$, $b_1 > d_1$, and $b_2 > s_2$, $d_2 > c_2$; with probability (1 – s), the two players play the left-hand side of the game tree, for which $c_1 > a_1$, $b_1 > d_1$, and $t_2 > b_2$, $d_2 > c_2$.

The information sets in Figure 1 indicate that the target country has to choose between violate and not violate the standard without knowing whether its opponent is soft or tough, while the sender country has to choose whether to sanction or not, without knowing the choice of the target country (but knowing, of course, its own type).

The target country has two strategies: to violate or not violate. Assume that it violates with probability p and does not violate with probability (1–p). The sender country has four strategies: To sanction unconditionally (whether it is tough or soft), to not sanction unconditionally, to sanction if tough and not sanction if soft, and to not sanction if tough and to sanction if soft. Of these four strategies, two are dominated and can be eliminated, so the sender country has two undominated strategies: to sanction when tough and to not sanction when soft, and to not sanction unconditionally. Assume that it follows the first with probability q and the second with probability (1–q). We have to calculate the equilibrium values p* and q*.

The expected utilities from each strategy are the following:

$$EV = (1-s)\,[qa_1+(1-q)b_1]+\ wb_1 \qquad [1C]$$

$$ENV = (1-s)\,[qc_1+(1-q)d_1] + wd_1 \qquad [2C]$$

$$ESNS = pt_2 + (1-p)c_2 \qquad [3C]$$

$$ENS = pb_2 + (1-p)d_2 \qquad [4C]$$

To calculate the equilibrium value of q, we equate (1C) and (2C), and solve for q. We find

$$q^* = (d_1-b_1) \,/\, [(1-s)\,(a_1-c_1+d_1-b_1)] \qquad [5C]$$

To calculate the equilibrium value of p, we equate (3C) and (4C), and solve for p. We find

$$p^* = (d_2-c_2) \,/\, (a_2-b_2+d_2-c_2) \qquad [6C]$$

Simple observation of (5C) and (6C) indicates that theorems 2 and 3 hold.

QED

More generally, in two-person games without pure strategy equilibria, the only existing equilibria are in mixed strategies, where the payoffs of each player affect only the behavior of the opponent, so theorems 2 and 3 hold regardless of the nature of uncertainty, as long as one of the subgames is the game of case 1 discussed in the section on the Robinson Crusoe fallacy (see Tsebelis, forthcoming, b).

REFERENCES

ACHEN, C. and D. SNIDAL (1989) "Rational deterrence theory and comparative case studies." World Politics 41: 143-169.

ADLER-KARLSSON, G. (1968) Western Economic Warfare 1947-1967: A Case Study in Foreign Economic Policy. Stockholm: Almqvist & Wiksell.

ALT, J. E., R. L. CALVERT, and B. D. HUMES (1988) "Reputation and hegemonic stability: a game-theoretic analysis." Amer. Pol. Sci. Rev. 82: 445-466.

BALDWIN, D. A. (1985) Economic Statecraft. Princeton, NJ: Princeton Univ. Press.

DOXEY, M. (1972) "International sanctions: a framework for analysis with special reference to the UN and Southern Africa." Int. Organization 26: 527-550.

GALTUNG, J. (1967) "On the effects of international economic sanctions, with examples from the case of Rhodesia." World Politics 19: 378-416.

HARSANYI, J. C. (1973) "Games with randomly distributed payoffs: a new rationale for mixed strategy equilibrium points." Int. J. of Game Theory 2: 1-23.

HIRSCH, M. W. and S. SMALE (1974) Differential Equations Dynamic Systems and Linear Algebra. New York: Academic Press.

HUFBAUER, G. C. and J. J. SCHOTT (1985) Economic Sanctions Reconsidered. Washington, DC: Institute For International Economics.

KREPS, D. M. and R. WILSON (1982) "Sequential equilibria." Econometrica 50: 863-894.

LAKATOS, I. (1970) "The methodology of scientific research programs," pp. 91-196 in I. Lakatos and A. Musgrave (eds.) Criticism and the Growth of Knowledge. Cambridge: Cambridge Univ. Press.

LUCE, R. D. and H. RAIFFA (1957) Games and Decisions. New York: John Wiley and Sons.

MAY, R. M. (1963) Stability and Complexity in Model Ecosystems. Princeton, NJ: Princeton Univ. Press.

MYERSON, R. B. (1978) "Refinements of the Nash equilibrium concept." Int. J. of Game Theory 7: 73-80.

ORDESHOOK, P. C. (1986) Game Theory and Political Theory. Cambridge: Cambridge Univ. Press.

PUTNAM, R. D. (1988) "International cooperation and the logic of 'two-level games.'" Int. Organization 42: 427-460.

RUSSETT, B. and P. HUTH (1989) "Dispute initiation and escalation in international politics," NSF grant proposal.

SELTEN, R. (1975) "Reexamination of the perfectness concept for equilibrium points in extensive games." Int. J. of Game Theory 4: 25-55.

TSEBELIS, G. (1989) "The abuse of probability in political analysis: the Robinson Crusoe fallacy." Amer. Pol. Sci. Rev. 83: 77-92.

TSEBELIS, G. (forthcoming, a) Nested Games: Rational Choice in Comparative Politics. Berkeley: Univ. of California Press.

TSEBELIS, G. (forthcoming, b) "More on the Robinson Crusoe fallacy." Amer. Pol. Sci. Rev. (forthcoming).

Van DAMME, E.E.C. (1984) Refinements of the Nash Equilibrium Concept. New York: Springer-Verlag.

WALLENSTEEN, P. (1968) "Characteristics of economic sanctions." J. of Peace Research 3: 248-267.

WU WEN-TSUN and JIANG JIA-HE (1962) "Essential equilibrium points of n-person non-cooperative games." Sci. Sinica 11: 1307-1322.