

ABSORBING OUTCOMES IN 2 X 2 GAMES

Steven J. Brams  
Department of Politics  
New York University  
New York, NY 10003

Marek Hessel  
College of Business Administration  
Fordham University  
Bronx, NY 10458

December 1980

No. 80-27

## ABSTRACT

### ABSORBING OUTCOMES IN $2 \times 2$ GAMES

Steven J. Brams, New York University  
Marek Hessel, Fordham University

Absorbing outcomes in the 41  $2 \times 2$  ordinal games without nonmyopic equilibria are defined and classified into four mutually exclusive categories. These categories are based on movements rational players would make from worst, next-worst, and next-best outcomes in these games, allowing for the deletability of some movements.

Although absorbing outcomes, at some level, are shown to exist in all 41 games, not all are equally defensible as stable outcomes. Moreover, conditions for determining them, while generally applicable to identifying nonmyopic equilibria in the 37  $2 \times 2$  games that contain them, do not show up the cooperative outcomes, which are nonmyopic equilibria, in Prisoners' Dilemma and Chicken.

Absorbing outcomes do not necessarily coincide with the myopic concept of a Nash equilibrium, which also has an interpretation based on the movements of rational players in a  $2 \times 2$  game. Nevertheless, in games without nonmyopic equilibria, absorbing outcomes do seem to capture the idea of long-term stability better than Nash equilibria when alternating, sequential moves by players are allowed.

ABSORBING OUTCOMES IN  $2 \times 2$  GAMES

Steven J. Brams and Marek Hessel

In a recent paper Brams and Wittman (1980) introduced the concept of a "nonmyopic equilibrium" in  $2 \times 2$  ordinal games. This concept is based on, and applicable to, ordinal games with the following rules of play:

1. Both players simultaneously choose strategies, thereby defining an initial outcome of the game.
2. Once at an initial outcome, either player can unilaterally switch his strategy and change that outcome to a subsequent outcome.
3. The other player can respond by unilaterally switching his strategy, thereby changing the subsequent outcome to a new subsequent outcome.
4. These strictly alternating moves continue until the player with the next move chooses not to switch his strategy. When this happens, the game terminates, and the outcome reached is the final outcome.

This view of a game reflects the idea that most real-life games are ongoing and do not necessarily terminate after the initial strategy choices are made.

In order to analyze ordinal games played according to these rules, some assumptions concerning the players' behavior must be made. In their

paper, Brams and Wittman postulated that both players act on an extended notion of mutually expected rationality in the sense that neither moves from an initial outcome if the move leads to a final outcome which does not improve his payoff. This determination is based on a backward induction analysis, whereby both players analyze the game tree of moves that are possible if either departs from the initial outcome. If this analysis reveals that no moves are rational--that both players would end up worse off if they departed from an initial outcome--the initial outcome is stable, and it is called a nonmyopic equilibrium.

The postulate of rationality described above produces stable outcomes in 37 of the 78 ordinal  $2 \times 2$  games (47 percent). In the 41 games without nonmyopic equilibria, however, the postulate, precisely because it is relatively unrestrictive, proves too weak to produce equilibrium outcomes. In these games, the rational players may "cycle" over the four outcomes indefinitely in an elusive attempt to obtain their best outcome. As a result, no outcome in these games is stable in the nonmyopic sense.

If one is willing to make the weak rationality postulate more restrictive, however, then it is possible to single out some outcomes as "conditionally," or "almost nonmyopically," stable in all the games which do not have nonmyopic equilibria. Our purpose here is to provide (progressively stronger) assumptions concerning the players' behavior so as to obtain (progressively weaker) stability in these games. In this manner, we shall define new concepts of stability applicable to the 41 games (53 percent) lacking nonmyopic equilibria.

These concepts usually coincide with nonmyopic equilibria in the 37 games that have them, but because their stability properties are less stringent than those of nonmyopic equilibria, they can also be used to identify "most stable"--if not the stable--outcomes in all the other games. We call these outcomes absorbing outcomes and show not only that they exist in all 41 ordinal  $2 \times 2$  games without nonmyopic equilibria but also that they do not necessarily coincide with Nash equilibria, the best known and most myopic of equilibrium concepts (to be discussed in the final section).

Unlike the unconditional stability of nonmyopic equilibrium outcomes, the stability of absorbing outcomes depends upon particular assumptions describing the players' behavior that go beyond the weak rationality postulate. We shall provide four conditions sufficient to obtain absorbing outcomes in all 41 games without nonmyopic equilibria. Since these outcomes do not necessarily coincide with Nash equilibria, we shall also give a fifth sufficient condition, comparable to the other four, that produces stability in the Nash sense. The relationship of absorbing outcomes to Nash equilibria and nonmyopic equilibria will be discussed in the final section of this paper.

#### Absorbing Outcomes

To define absorbing outcomes, consider the four previous rules of play. Essentially they allow players to move away from initial outcomes an indefinite number of times. To which outcomes they do move is unanswered by these rules.

The determination of nonmyopic equilibria in Brams and Wittman (1980) assumed players look ahead to ascertain where they will end up if

they depart from an initial outcome. If the final outcome is worse for the departing player(s) than the initial outcome, then there will be no incentive to move, and the initial outcome will be stable.

What if a game contains no outcome that is stable in this sense? Is it still possible that the previous rules of play might induce players to move toward particular outcomes?

To answer this question, we shall assume that players will definitely move away from particular outcomes. Consider the 78  $2 \times 2$  ordinal games, in which two players each have two strategies and strictly order the four resulting outcomes from best to worst. Complete listings of these games are given in Rapoport and Guyer (1966) and Brams (1977), but here we shall concentrate only on those 41 games that have no nonmyopic equilibria.

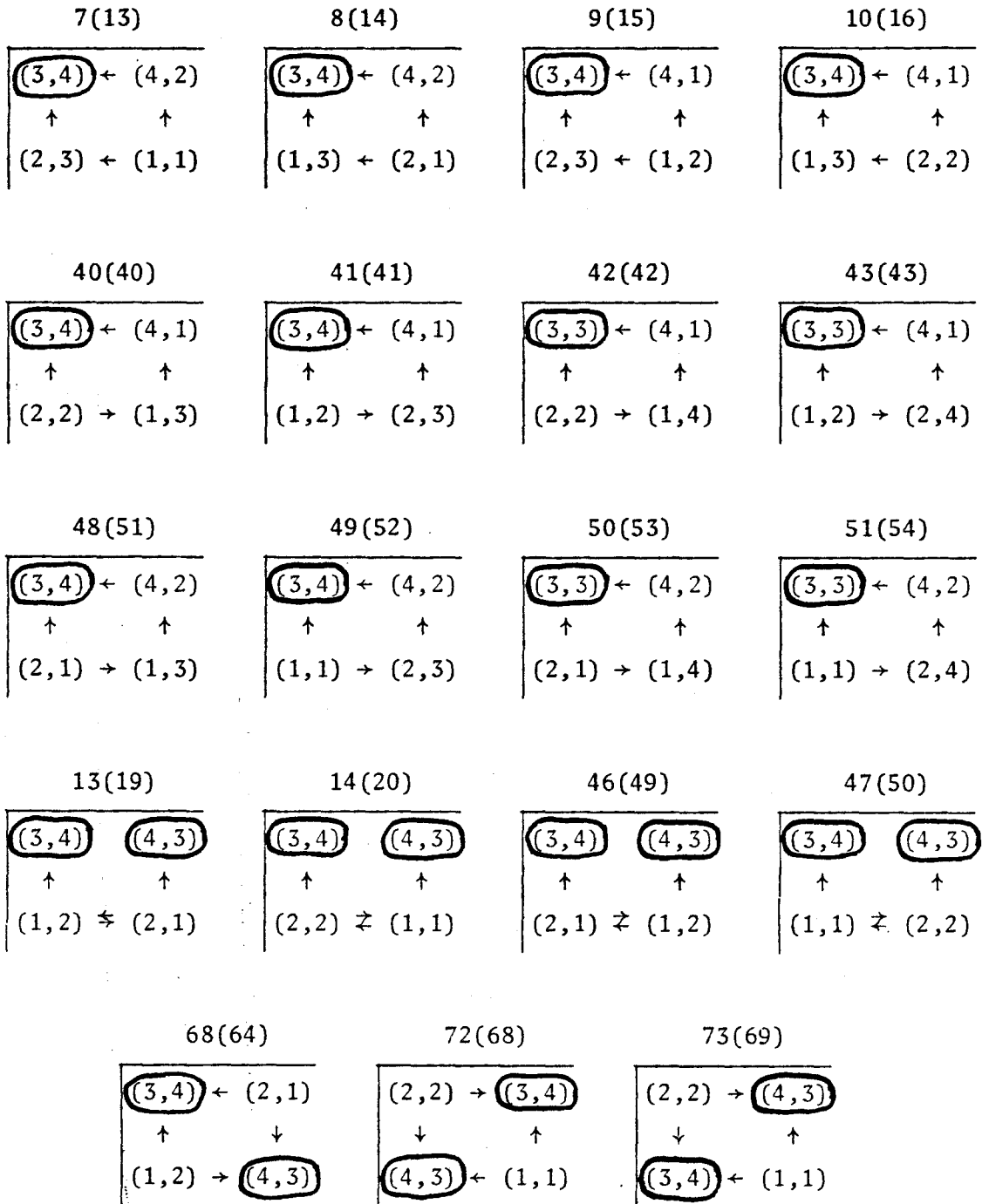
If a game contains only outcomes from which players will have an incentive to depart, looking ahead, it seems reasonable to assume that this incentive will be strongest for outcomes they rank worst and next worst. Accordingly, we begin by offering a sufficient condition, based on this criterion, for an outcome in a  $2 \times 2$  ordinal game to be absorbing:

- I. If departures from worst and next-worst outcomes indicate convergence either to a single outcome or to two outcomes, the converged-upon outcome(s) is (are) absorbing.

Of the 41  $2 \times 2$  ordinal games without nonmyopic equilibria, 19 (46 percent) meet condition I. These games, numbered according to the listings in Brams (1977)--first number on top of games--and Rapoport and Guyer (1966)--second number, in parentheses, on top--are given in Figure 1, and their absorbing outcomes are circled.

FIGURE 1

GAMES WITH ABSORBING OUTCOME(S), BASED ON CONDITION I



Departures from the two worst outcomes are indicated by arrows from these outcomes, which are represented by ordered pairs of numbers in the outcome matrix. The first number in each pair gives row's (R's) ranking of the outcome (4 = best, 3 = next-best, 2 = next-worst, 1 = worst), the second number column's (C's) ranking. Note that the arrows from outcomes ranked 1 and 2 by R and C converge on one outcome in the first twelve games shown in Figure 1, on two outcomes in the remaining seven games.

An outcome is considered converged-upon if the one or two arrows connecting it to other outcomes are all incoming. Hence, once it is reached, there will be no incentive to depart from it in order for a player to avoid one of his two worst outcomes.

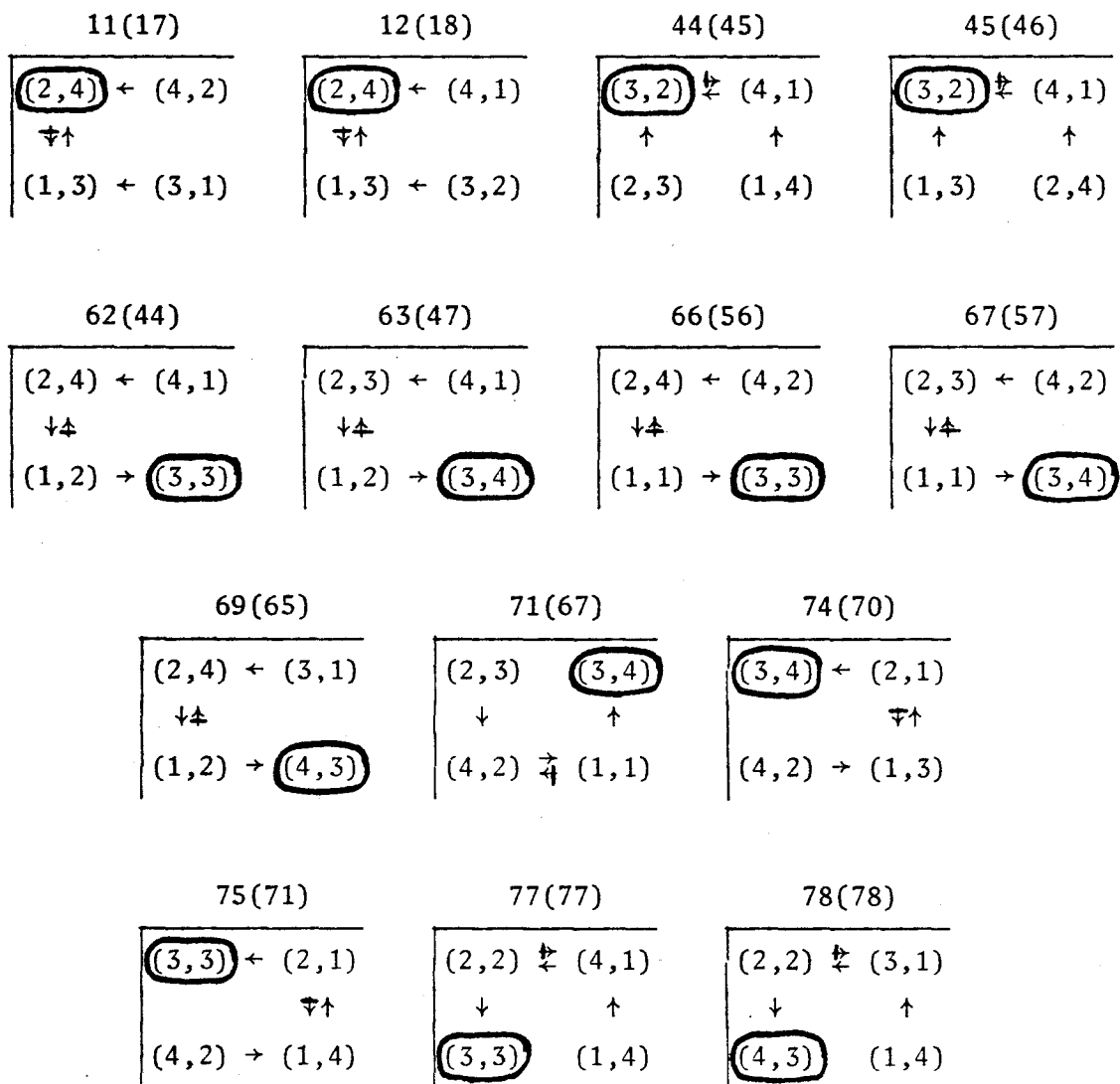
In addition to the games in Figure 1, there are several games without nonmyopic equilibria in which there are no converged-upon outcomes, yet there would be if one arrow were deleted. These games all contain one pair of outcomes connected by two arrows, one arrow going one way and the other going the other way. Is it reasonable to delete one of these two arrows?

An arrow is deletable if it would lead to an outcome worse for a player than would be the case if the movement between the two outcomes in question went in the other direction. The 14 games shown in Figure 2 (34 percent of the 41 games without nonmyopic equilibria) all have deletable arrows, indicated by a line through them, without which there is a converged-upon outcome. Thus, as another sufficient condition for absorbing outcomes, we offer the following:

- II. If departures from worst and next-worst outcomes do not indicate convergence, but there is a deletable arrow whose deletion results in a converged-upon outcome, this outcome is absorbing.

FIGURE 2

GAMES WITH AN ABSORBING OUTCOME, BASED ON CONDITION II



To illustrate this condition and the concept of deletability, consider the first game, #11(17), shown in Figure 2. There are double arrows between outcomes (2,4) and (1,3) because R would like to move from his worst outcome at (1,3) and his next-worst outcome at (2,4). However, to move from (2,4) to (1,3) would be disastrous for R, because C would not subsequently move to his worst outcome at (3,1); rather, he would move away from it; as shown by the arrow from it. (For now, we assume the only movements possible are those shown by the arrows, but this assumption will be relaxed later.) Hence, it is prudent for R to consider only a move from (1,3), because a move to this outcome from (2,4) would terminate the process at R's worst outcome. Drawing a line through the arrow from (2,4) to (1,3), we see that with its deletion the only converged-upon outcome in this game is (2,4), making it the unique absorbing outcome. This outcome, along with other absorbing outcomes, is circled in Figure 2.

The arrows deleted in the games in Figure 2 do not necessarily emanate from a next-worst (2) outcome, as in the game just discussed. Instead, it may, paradoxically, be in the interest of a player to depart from his next-worst outcome to his worst (1) outcome and so delete the arrow from his worst outcome. This happens when a player can anticipate that the other player will move the process to still a better outcome for himself (3 or 4), so he does not, by moving from his worst to next-worst outcome, have to ensure his own security level of 2.

Game #62(44) in Figure 2 is a case in point. R should depart from (2,4) to (1,2), because, once at (1,2), C will move the process to (3,3), better for both players.

But, one might respond, what if the process were at (1,2)? If C, obtaining only his next-worst outcome there, could hold out longer than R, suffering his worst outcome, then he could force R to move to (2,4), C's best outcome. This is certainly possible, but R can readily prevent it from occurring by always choosing, initially, his first strategy. Whatever strategy C chooses, the process will necessarily move to (3,3) if the players are rational, rendering the next-best outcome for both players the unique absorbing outcome.

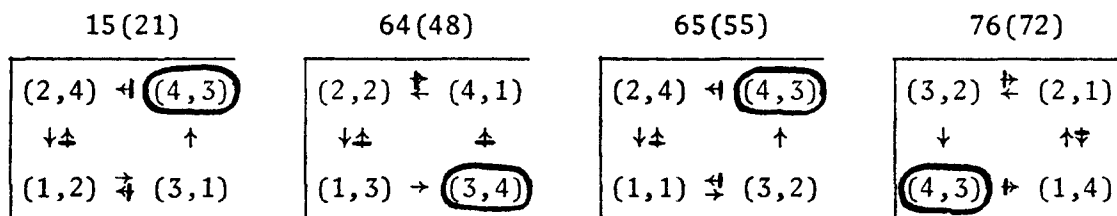
In fact, all absorbing outcomes in Figure 2 are unique, unlike those in Figure 1. Moreover, in four of the fourteen games in Figure 2, convergence is to an outcome next worst (2) for one player. By contrast, all the absorbing outcomes in Figure 1, whether they be unique or not, ensure each player at least his next-best (3) outcome. Clearly, convergence without deletion in general promises players outcomes superior to their security level of 2, whereas convergence with deletion does not.

So far so good. Yet there remain four games without nonmyopic equilibria (10 percent of the 41) in which arrows from the worst and next-worst outcomes connect only three of the four outcomes. Necessarily, these two connections, comprising four arrows, show movements in each of the two directions. The double arrows for the four games shown in Figure 3 illustrate all instances in which this phenomenon occurs (ignore for now the single arrows).

To connect the fourth outcome in these games to the other three, and determine which outcome (if any) is absorbing, assume the players consider movements from their next-best (3) outcomes. These movements

FIGURE 3

GAMES WITH AN ABSORBING OUTCOME, BASED ON CONDITION III



are shown by the single arrows from two outcomes in each game in Figure 3.

Patently, the deletion of no single arrow in these three matrices will result in convergence to either one or two outcomes, from which there are no outgoing arrows. Because there is no possibility of immediate convergence with the deletion of a single arrow, assume the players first consider the possibility of deleting arrow(s) emanating from their next-best (3) outcomes, which at least provide each an outcome better their security levels of 2. If, by such a deletion, they can be assured that the process will stop and they will thereby obtain at least their next-best outcomes, we assume that this deletion will occur. Stoppage will be assured, we assume, if the player not deleting an arrow obtains his best (4) outcome at this point.

By this criterion, there is a single (4,3) or (3,4) outcome in all games in Figure 3 from which an outgoing arrow can be deleted, which is illustrated by a line through one single arrow in each of the four games. Given this deletion, it is now easy to show that the player who makes it can ensure his next-best (3) outcome--and, by implication, two further arrow deletions that also ensure this outcome is converged-upon--given the previous rules and the assumptions we shall now make.

Assume the player who makes the deletion of an arrow from his next-best outcome at (4,3) or (3,4) chooses, initially, his strategy associated with this outcome. For example, in game #15(21) in Figure 3, C would choose his second strategy. If R chose his first strategy, the outcome would be (4,3), and there would be no further movements. If R chose his second strategy, the outcome would be (3,1). Knowing that the deletion from (4,3) to (2,4) had been made, however, R would move the process from (3,1) to (4,3), where it would stay.

Implicitly, this deletion, and the acceptability of (4,3) to C, assumes that C at (3,1) would not try to move the process to (1,2), so the arrow from (3,1) to (1,2) can be deleted. But now the arrow from (1,2) to (2,4) can be deleted because R would move from (2,4) to (1,2), knowing that the process would terminate at (4,3). A similar logic can be applied to the elimination of one of each of the double arrows in the three other games in Figure 3--after the deletion by C of the arrow from (4,3), or R of the arrow from (3,4), in each of these games--leading to a unique converged-upon outcome. We call this process successive deletability, which can be summarized by the following rules:

1. If a game does not contain an absorbing outcome based on Conditions I or II, consider movements from next-best (3) outcomes as well.
2. Delete the arrow from the next-best outcome for a player if there is stoppage--this outcome is best (4) for the other player.
3. Successively delete one of each set of double arrows such that the process moves to the (3,4) or (4,3) outcome.

These rules offer a third sufficient condition for absorbing outcomes:

- III. Given that conditions I and II are not satisfied, a converged-upon outcome found by the rules of successive deletability is absorbing.

Note that, like all absorbing outcomes that satisfy condition I, those that satisfy condition III (circled in Figure 3) ensure that each player will obtain at least his next-best (3) outcome. By comparison, the "simple" deletability of condition II, which ensures a player only his

security level of 2, does not guarantee a player as good an outcome as the more "complex" successive deletability of condition III.

All but four of the 41 games without nonmyopic equilibria are now accounted for by conditions I, II, and III. Arguably, the remaining games also have absorbing outcomes, but the rationale for absorption in these games is weaker. As can be seen in Figure 4, departures from the worst (1) and next-worst (2) outcomes in all these games lead to cycling over the four outcomes.

To break these cycles, one can argue as follows. Given that these games contain no outcomes which are better than next-worst (2) for both players, the players' first priority will be to avoid their worst outcomes. Accordingly, we assume the following rules of worst-outcome avoidance:

1. In games that do not satisfy conditions I, II, and III, players will always depart from their worst outcomes.
2. Players will not depart from their next-worst outcomes if they lead to subsequent worst-outcome departures, and hence to cycling.

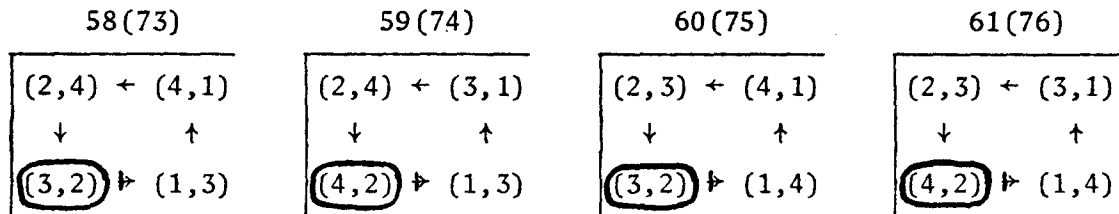
These rules offer a fourth sufficient condition for absorbing outcomes:

- IV. Given that conditions I, II, and III are not satisfied, a converged-upon outcome found by the rules of worst-outcome avoidance is absorbing.

To illustrate this condition, consider the first of the four games in Figure 4. Movement by R from his worst outcome (1,3), and then by C from (4,1), lead to (2,4). Should R move next to (3,2), where he would obtain his next-best (3) outcome?

FIGURE 4

GAMES WITH AN ABSORBING OUTCOME, BASED ON CONDITION IV



Since this move does not lead to a worst-outcome avoidance move from (3,2), we assume it will be made. On the other hand, if C moves subsequently to (1,3), this sets in motion subsequent worst-outcome departures, leading to cycling. Because this latter move violates rule 2 for worst-outcome avoidance, we assume it will not be made, and we therefore draw a line through the arrow leading from (3,2) to (1,3).

Breaking the cycle at (3,2) makes this the converged-upon outcome, rendering (3,2) absorbing by condition IV. Similarly for the other three games in Figure 4, a line through the arrow that leads to worst-outcome departures is indicated, and the prior absorbing outcome is circled.

In summary, we have identified absorbing outcomes in the 41  $2 \times 2$  ordinal games without nonmyopic equilibria that satisfy one of four conditions:

- I. Convergence from two worst outcomes: 19 games (46 percent).
- II. Convergence from two worst outcomes after deletability (of one arrow): 14 games (34 percent).
- III. Convergence from three worst outcomes after successive deletability (of three arrows): 4 games (10 percent).
- IV. Convergence from worst-outcome avoidance by breaking cycle before a worst outcome would be reached: 4 games (10 percent).

Thus, all 41 games contain at least one absorbing outcome, with 34 games having exactly one such outcome and 7 games (all satisfying condition I) having two.

There is, however, a price to pay in designating at least one outcome "absorbing" in every game. The price is that the criteria for absorbingness we have used do not seem equally defensible. The most

compelling justification for an outcome to be absorbing seems to be that embodied in condition I, with the justifications of conditions II and III somewhat weaker. Least defensible is condition IV, but we have given it for the sake of completeness since it reflects a plausible belief some players may harbor of wanting to avoid worst outcomes, and cycles that contain them.

Underlying all four conditions is the possibility players have of alternately switching strategies once an initial outcome is chosen. In real-life games, such switching seems endemic as players seek out better outcomes, if not always stability. As Boulding (1978, p. 145) stated the problem,

What we are struggling with here is the problem of relational sequences. We have two parties, A and B, in some sort of contact. A does something, so B does something, which makes A do something else, then B does something else, A then does something else, and so on in the relational sequence. To describe this, we must have some kind of theory of behavior.

A theory of absorbing outcomes, we believe, complements a theory of nonmyopic equilibria, comparisons of which we shall offer in the final section.

#### Nash Equilibria, Nonmyopic Equilibria, and Absorbing Outcomes

Before comparing absorbing outcomes with nonmyopic equilibrium outcomes, we will relate the former to another equilibrium concept, namely that proposed by Nash (1951). Before defining this concept in a manner consistent with conditions I-IV, we note that absorbing outcomes in the 41 games to which conditions I-IV apply include many outcomes which are not Nash equilibria; conversely, many Nash equilibria are not

absorbing under any of these conditions. The precise breakdown is as follows:

38 outcomes are both absorbing and Nash;

15 outcomes are absorbing but not Nash;

17 outcomes are Nash but not absorbing.

Thus, the bulk of absorbing outcomes are also Nash equilibria, but 32 outcomes (40 percent) meet only one criterion. Manifestly, absorbing outcomes point to an attractiveness of outcomes not coincidental with the ability of a Nash equilibria to "absorb" outcomes only in the row and column of the outcome matrix in which they lie.

In fact, it is possible to provide a condition--in the spirit of conditions I-IV above--which produces "Nash-absorbing" outcomes in all  $2 \times 2$  games. The deletability process on which this condition is based assumes that each player considers moves away from all but his best outcome (i.e., away from his worst, next-worst, and next-best outcomes). The resulting six moves (three for each player) from four outcomes mean that some of the outcomes will be connected by (conflicting) double arrows, as illustrated in Figure 5.

To reduce these double arrows paired against each other to single (one-way) arrows, use the following rule: In every pair of double arrows, delete the arrow which points away from the higher-ranked outcome. This deletability principle, which we call Nash deletability, offers the following condition for stability in the Nash sense:

- V. An outcome is absorbing in the Nash sense iff it is converged upon as a result of Nash deletability.

FIGURE 5

## NASH DELETABILITY IN TWO GAMES

15(21)	58(73)
$(2,4) \leftarrow (4,3)$ $\uparrow \quad \quad \uparrow$ $(1,2) \leftarrow (3,1)$	$(2,4) \leftarrow (4,1)$ $\downarrow \quad \quad \uparrow$ $(3,2) \leftarrow (1,3)$

To illustrate Nash deletability, consider game #15(21) in Figure 5. There are double arrows connecting (1,2) with both (2,4) and (3,1). In the first pair, (1,2) and (2,4), the arrow emanating from (2,4) is deleted because the other arrow in the pair points away from a lower-ranked outcome, namely (2,4). In the pair (1,2) and (3,1), the arrow emanating from (1,2) is deleted because C ranks (3,1) below (1,2). The deletion of these two arrows makes (2,4) absorbing in the Nash sense because (2,4) is a Nash equilibrium in the game: neither player has an incentive to depart unilaterally from (2,4) because he would do (immediately) worse if he did.

Nash deletability does not produce "Nash-absorbing" outcomes in all the games discussed here. Game #58(73) in Figure 5, wherein the arrows remaining after Nash-deletions form a cycle, is an example. Thus, no outcome in this game is absorbing in the Nash sense; and, clearly, game #58(73) has no Nash equilibrium.

Condition V illustrates the principal difference between Nash equilibria and absorbing outcomes: Nash deletability is based on a myopic principle of getting away from worse outcomes, whereas absorbing outcomes are based on a less myopic principle of going toward better outcomes. As long as there is no conflict of interest between the two principles--condition I exemplifies this situation--Nash equilibria are also absorbing outcomes. Whenever the conflict is present, such as in game #5(21) in Figures 3 and 5, the two stability concepts do not coincide.

This conflict in viewpoints on stability is perhaps not surprising in light of the historical development of the different equilibrium concepts. Nash equilibria were developed for static games in which single possible deviations from initial strategy choices were assumed to

terminate play, whereas absorbing outcomes, including nonmyopic equilibria, were developed for dynamic games in which an unspecified number of moves from initial strategy choices were considered possible before play would terminate.

Because absorbing outcomes are based on less myopic considerations than Nash equilibria, they relate in a natural way to nonmyopic equilibria. The logical basis for nonmyopic equilibria is developed in detail in Brams and Wittman (1980) and, in a different context, Brams (1980a, 1980b), so will not be presented here. Suffice it to say that nonmyopic equilibria in general satisfy conditions I, II, or III for absorbing outcomes, subject to some minor qualifications in the case of condition III. However, two of the 37 games with nonmyopic equilibria, Prisoners' Dilemma and Chicken, do not satisfy any of conditions I-IV with respect to their cooperative (3,3) outcomes.

These two games are shown in Figure 6, with arrows emanating from their worst, next-worst, and next-best outcomes. It is easily seen that in neither game is there convergence to the (3,3) nonmyopic equilibrium. Quite the contrary: Both players in these games would apparently move from (3,3), as shown by the arrows. The inapplicability of conditions I-IV to these games testifies to the fact that nonmyopic equilibria are based on a more farsighted view of game play than is provided by "myopic" arrows, including those that yield Nash equilibria as given by condition V.

Absorbing outcomes seem best suited for identifying outcomes that will attract players even if, once reached, these outcomes are not as stable as are nonmyopic equilibria. Thereby they fill a gap between the short-term stability of Nash equilibria and the long-term stability of Brams-Wittman nonmyopic equilibria.

FIGURE 6

## PRISONERS' DILEMMA AND CHICKEN

$(3,3) \rightarrow (1,4)$	$(3,3) \rightarrow (2,4)$
$\downarrow$ $\downarrow\uparrow$	$\downarrow$ $\uparrow\downarrow$
$(4,1) \leftarrow (2,2)$	$(4,2) \leftarrow (1,1)$

It remains to be seen whether absorbing outcomes mirror a stability in real-life games that can be modeled by the  $41 \ 2 \times 2$  games without nonmyopic equilibria. To be sure, the absorbing outcomes in these games that satisfy the four different conditions are not all equally appealing, and one may want to distinguish among them. But wherever one draws the line on "absorbingness," a level-based notion of stability that mirrors the ability of players to move and countermove from initial outcomes in games seems to capture the idea of dynamic, sequential play that is rationally based.

## REFERENCES

- Boulding, Kenneth E. (1978). Ecodynamics: A New Theory of Societal Dynamics. Beverly Hills, CA: Sage.
- Brams, Steven J. (1980a). "Mathematics and Theology: Game-Theoretic Implications of God's Omniscience." Mathematics Magazine, (November):
- (1980b). "A Resolution of the Paradox of Omniscience," mimeographed.
- Brams, Steven J., and Donald Wittman (1980). "Nonmyopic Equilibria in  $2 \times 2$  Games," mimeographed.
- (1977). "Deception in  $2 \times 2$  Games," Journal of Peace Science, 2 (Spring): 171-203.
- Nash, J. (1951). "Non-cooperative Games," Annals of Mathematics, 54: 286-295.
- Rapoport, Anatol, and Melvin Guyer (1966). "A Taxonomy of  $2 \times 2$  Games." General Systems: Yearbook of the Society for General Systems Research, 11: 203-214.