Notes*
II

MANAGEMENT SCIENCE
Vol. 29, No. 11, November 1983
Printed in U.S.A.

SUBJECTIVE PROBABILITY AND THE THEORY OF GAMES: SOME FURTHER COMMENTS†

ALVIN E. ROTH,‡ AND FRANCOISE SCHOUMAKER§

This note reconsiders some of the issues raised by Kadane and Larkey, and Harsanyi, and briefly discusses some relevant empirical results.
(GAME THEORY; RATIONALITY; BAYESIANISM; SUBJECTIVE PROBABILITIES)

Some of the recent discussion of what role subjective probabilities might play in the theory of games (Kadane and Larkey [8], [9]; Harsanyi [6], [7]) seems to have been conducted at cross-purposes. This note reconsiders some of the issues and discusses some relevant empirical results.

Two parallel traditions have coexisted in game theory for some time. On the one hand, theories which use the rationality of the players to restrict the predicted outcomes to some range of possibilities, but not necessarily to a unique outcome, go back at least as far as Edgeworth [2], who considered two-person bargaining over the exchange of commodities. He concluded that when such bargaining was conducted by rational players, the outcome would fall in an interval which he called the contract curve (which is now commonly referred to as the *core* of such a game), but that the precise outcome in this interval was fundamentally indeterminate. Nash [12], on the other hand, considered the same sort of bargaining problem, and constructed a theory which selects a unique outcome from the contract curve.[1] Explorations and variations of Nash's work on bargaining have given birth to a large literature, aspects of which are surveyed in Roth [16]. John Harsanyi, who in Harsanyi [3] demonstrated a close relationship between Nash's work and the earlier work of Zeuthen [21], is the most eloquent and creative modern exponent of the idea that the rationality of the players in a game should allow unique outcomes to be predicted. His work has extended this idea in several directions (see Harsanyi [4], [5]).

There is a close relationship between these two traditional positions and the question of how players' subjective expectations of other players' behavior might enter into a theory of games. A theory which predicts the occurrence of a unique outcome for every game among perfectly rational players is a theory in which each such player must have a unique (and correct) subjective prior distribution on the behavior of the other players. However a theory which predicts that a range of possible outcomes may occur allows the possibility that a player's subjective prior distribution on the behavior of others may be drawn from some range of distributions consistent with the data of the game. (Even in such a theory, considerations of rationality will in general rule out

*All Notes are refereed.

†Accepted by Ambar G. Rao; received May 12, 1982. This Note has been with the authors 1 month for 1 revision.

‡University of Pittsburgh.

§University of Illinois, Champaign.

[1] Nash's paper [12] on the bargaining problem should not be confused with his work on the equilibria of noncooperative games (Nash [13]). In a subsequent paper, Nash [14] reinterpreted the bargaining theory proposed in his 1950 paper as a method for selecting a unique noncooperative equilibrium from the continuum of possible equilibria in a bargaining situation modelled as a noncooperative game.

0025-1909/83/2911/1337$01.25

at least *some* prior distributions: e.g., the subjective expectation that your opponent will very likely choose a dominated strategy is incompatible with the subjective expectation that he is a utility maximizer.)

In what follows, we will briefly consider the empirical content of these alternative positions in the context of bargaining. First we consider a "thought experiment" which will help identify precisely the questions involved, and then we very briefly report some actual experimental results which have been recently obtained. We conclude with some comments on the role of experimental evidence in game theory, viewed both as a descriptive theory and as a prescriptive theory of perfectly rational strategic behavior. We will also argue that the experimental evidence can be interpreted as supporting the view that "perfect rationality" may not always be completely well-defined.

A Thought Experiment

Consider the case of two individuals who must bargain over how to divide $100. The rules are that, after some fixed interval for negotiation, each bargainer simultaneously and privately writes down a demand. If the two demands sum to no more than $100, each bargainer receives his demand, otherwise each receives nothing. If each player is a rational utility maximizer with a known utility function, then the customary game-theoretic assumptions are satisfied. (In general, game-theoretic models of games with complete information assume that the data of the game is *common knowledge*; i.e., that it is not merely known by the players, but also known to be known, etc. cf. Lewis [10]; Aumann [1]; Milgrom [11].)

If we adopt the point of view that only a single outcome of the game is consistent with the perfect rationality of the players, then a theory of games should predict, from their utility functions, how such players will divide the $100. For example, Nash's [12] solution to the bargaining problem is such a theory. Alternatively, if we adopt the view that more than one outcome is consistent with the perfect rationality of the players, then we can consider whether the outcome of this game depends on subjective expectations of the players which are not completely determined by the data of the game (even when the players are perfectly rational).

To see how the expectations of the bargainers might influence the outcome of the game, consider the following "thought experiment." A randomly selected individual plays some very large number of these games. Although he doesn't know it, all of his opponents are confederates of the experimenter, and they all allow him to obtain, say, $80. After he has gone through this experience, *you* have the opportunity of bargaining with him on your own behalf (i.e., not as a confederate). His past success is common knowledge. It will obviously be difficult to bargain with him on an equal basis, since he expects (and has every reason to expect) to receive $80, and since he expects (and has every reason to expect) that you will concede it to him. If this is the *only time* you will be bargaining with him, the fact that this randomly selected individual now expects to get $80 will make it very risky for you to write down a demand of more than $20. (One can imagine the opening salvos of the negotiations. You say "You know, those people you bargained with were all confederates, and I'm not." He replies, "That's what many of my previous opponents said too, but when the time came to write down demands they wrote down $20, and so will you.") Recall that you are dealing with a randomly selected individual, so this state of affairs arises not from his utility function or other personal attributes, but from the expectations created by his unusual experience.

Based on his experience, he is perfectly justified in writing down a demand of $80. If you demand more than $20, you will receive nothing, and since that is the only time you will play the game, your action has no further influence on your future welfare. So

it is rational for you to demand \$20, and therefore, knowing this, it is rational for your opponent to demand \$80. Not only is he maximizing his utility given his expectations, but these expectations continue to be correct. One can imagine a string of new opponents, none of whom are confederates, each of whom plays our randomly selected individual only once, all of whom know his past history, and none of whom therefore ever demands more than \$20. Thus these artificially created expectations are nonetheless stable and self-sustaining.

Some Empirical Results

In order to actually implement experimental tests of game-theoretic models of bargaining, it is necessary to conduct them under conditions which allow the utility of the players for each potential agreement to be known. An experimental design which allows this was introduced in Roth and Malouf [17], and also employed in Roth, Malouf and Murnighan [18], Roth and Murnighan [19], and Roth and Schoumaker [20]. In the experiments reported there, players played "binary lottery games" in which they bargained over the probability each one would have of winning a certain monetary prize. The utility of each player for an agreement in such a game is formally equivalent to the probability which that agreement gives him of winning his prize.

The experiment reported in Roth and Schoumaker [20] employs this design to implement a version of the thought experiment described above. Players seated at computer terminals bargained with programmed opponents designed, in different experimental conditions, to create different expectations. These expectations were then observed to be stable and self-sustaining when play between genuine (nonprogrammed) opponents was begun. Thus the observed results in this experiment were consistent with those suggested above for the thought experiment. That is to say, by manipulating the expectations of the bargainers, the experiment demonstrated that such expectations play a critical role in determining the outcome of such games when played by members of our subject population.

Concluding Remarks

The experimental results cited above lend support to the conclusion that the subjective expectations of players have a role to play in descriptive theories of games. To consider what role such expectations might play in theories of games intended to capture strategic behavior among "perfectly rational" players, we need to consider what is meant by a rational player.

In conventional game-theoretic usage, a rational player is one whose choice behavior can be accurately modelled as utility maximization. The consequences of this kind of "individual" rationality in models like Nash's have been shown (in Roth [15], [16]) to be more far-reaching than was originally anticipated, but it appears that further assumptions about the behavior of players would be needed to uniquely determine the outcome of bargaining problems. In particular, the behavior described in our thought experiment, and observed in the experiment reported in Roth and Schoumaker [20], seems consistent with utility maximization, and we do not see any justification for calling that behavior irrational.

It is quite natural that there should be a two-way interchange between prescriptive theories of rational strategic behavior and descriptive theories of actual strategic behavior. On the one hand, the assumption that players behave in a rational, goal-directed manner not only helps organize and interpret a great deal of empirical observation, it has also proved to be a versatile and powerful explanatory hypothesis, with considerable predictive power in a wide variety of situations. It is difficult to see how any very general descriptive theory could even be formulated except in conjunc-

tion with some kind of prescriptive theory. On the other hand, since the consequence of a given player's choice in a game is partly determined by the behavior of other players, the very definition of rational choice depends on the behavior of others. In situations in which more than one set of players' expectations and behaviors is consistent with utility maximization, knowing how to behave rationally may depend on knowing others' expectations, which in some circumstances may be primarily an empirical question.

In our opinion, the preponderance of (theoretical and empirical) evidence at this time supports the view that there are some situations in which the individual rationality of all the players is insufficient to uniquely determine the outcome. In such situations, the subjective expectations of the players have a potentially decisive role, from the point of view of both prescriptive and descriptive theories.[2]

[2]This work has been supported by NSF Grants SOC 78-09928 and SES 79-15356.

References

1. AUMANN, ROBERT J., "Agreeing to Disagree," *Ann. Statist.*, Vol. 4 (1976), pp. 1236–1239.
2. EDGEWORTH, F. Y., *Mathematical Psychics*, Routledge and Kegan Paul, London, 1881.
3. HARSANYI, JOHN C., "Approaches to the Bargaining Problem and After the Theory of Games: A Critical Discussion of Zeuthen's, Hicks', and Nash's Theories," *Econometrica*, Vol. 24 (1956), pp. 144–157.
4. ———, "The Tracing Procedure: A Bayesian Approach to Defining a Solution for *n*-Person Noncooperative Games," *Internat. J. Game Theory*, Vol. 4 (1975), pp. 61–94.
5. ———, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, Cambridge, 1977.
6. ———, "Subjective Probability and the Theory of Games: Comments on Kadane and Larkey's Paper," *Management Sci.*, Vol. 28 (1982), pp. 120–124.
7. ———, "Rejoinder to Professors Kadane and Larkey," *Management Sci.*, Vol. 28 (1982), pp. 124–125.
8. KADANE, JOSEPH B. AND LARKEY, PATRICK D., "Subjective Probability and the Theory of Games," *Management Sci.*, Vol. 28 (1982), pp. 113–120.
9. ——— AND ———, "Reply to Professor Harsanyi," *Management Sci.*, Vol. 28 (1982), p. 124.
10. LEWIS, DAVID K., *Convention: A Philosophical Study*, Harvard University Press, Cambridge, 1969.
11. MILGROM, PAUL, "An Axiomatic Characterization of Common Knowledge," *Econometrica*, Vol. 49 (1981), pp. 219–221.
12. NASH, JOHN F., "The Bargaining Problem," *Econometrica*, Vol. 28 (1950), pp. 155–162.
13. ———, "Non Cooperative Games," *Ann. of Math.*, Vol. 54 (1951), pp. 286–295.
14. ———, "Two Person Cooperative Games," *Econometrica*, Vol. 21 (1953), pp. 128–140.
15. ROTH, ALVIN E., "Individual Rationality and Nash's Solution to the Bargaining Problem," *Math. Oper. Res.*, Vol. 2 (1977), pp. 64–65.
16. ———, *Axiomatic Models of Bargaining*, Springer, Berlin, 1979.
17. ——— AND MALOUF, MICHAEL W. K., "Game-Theoretic Models and the Role of Information in Bargaining," *Psychological Rev.*, Vol. 86 (1979), pp. 574–594.
18. ———, ——— AND MURNIGHAN, J. KEITH, "Sociological versus Strategic Factors in Bargaining," *J. Econom. Behavior and Organization*, Vol. 21 (1981), pp. 153–177.
19. ——— AND MURNIGHAN, J. KEITH, "The Role of Information in Bargaining: An Experimental Study," *Econometrica*, Vol. 50 (1982), pp. 1123–1142.
20. ——— AND SCHOUMAKER, FRANCOISE, "Expectations and Reputations in Bargaining, An Experimental Study," *Amer. Econom. Rev.*, Vol. 73 (1983), pp. 362–372.
21. ZEUTHEN, FREDERICK, *Problems of Monopoly and Economic Warfare*, G. Routledge, London, 1930.