On the Status of the Nash Type of Noncooperative Equilibrium in Economic Theory

Author(s): Leif Johansen

Source: *The Scandinavian Journal of Economics*, Vol. 84, No. 3 (1982), pp. 421-441
Published by: Wiley on behalf of The Scandinavian Journal of Economics
Stable URL: http://www.jstor.org/stable/3439426
Accessed: 27-06-2016 07:12 UTC

Scand. J. of Economics 84 (3), 421–441, 1982

On the Status of the Nash Type of Noncooperative Equilibrium in Economic Theory

*Leif Johansen**

University of Oslo, Norway

Abstract

In economic theory the Nash type of noncooperative equilibrium is a widely used equilibrium concept. Nevertheless, the concept is frequently misunderstood, and misleading explanations and interpretations are often given. This paper presents a sample of such frequently appearing misleading explanations, and goes on to explain why they are misleading. The case for using the Nash noncooperative equilibrium concept is stronger than what is implied by the many misleading statements found in the literature. A basic issue is whether it can be argued that decisions in accordance with the noncooperative equilibrium theory are individually rational decisions. This issue is not covered by the classical rationality postulate in economic theory. This problem is tackled by stating certain postulates which may be taken as defining individual rationality in situations of noncooperative interaction; it is shown that Nash behavior satisfies these postulates.

I. Introduction

The concept of noncooperative equilibrium of the Nash type is widely used in economic theory. It is therefore important to have a clear view of the nature of this type of equilibrium in order to assess its appropriateness in various contexts and form an opinion about its empirical relevance. It seems to me that there are many misunderstandings about the concept to be found in the literature. The purpose of this paper is to discuss and, hopefully, clarify some of the issues connected with the concept. The discussion may seem pedantic to some readers. However, I think it may be useful as an antidote to routine presentations and uses of the concept.

The gist of my discussion is, on the one hand, that the concept of noncooperative equilibrium of the Nash type is a fundamental solution concept in situations of noncooperative interaction, and that it deserves a special status. It is not merely an example of a solution concept in noncooperative situations; it will be argued that *if we take for granted that there is a natural solution concept in such situations,* then it must be the Nash

* I am grateful to Atle Seierstad for discussions and viewpoints on many parts of this paper.

noncooperative solution. On the other hand, the condition stipulated in this conclusion is not devoid of problems. It may bring us to the border of reasoning where we can no longer say unhesitatingly that our arguments are based firmly on assumptions of rational behavior. Accordingly, the issue is of profound importance for economic theory, which may claim to be strong as long as the theory can be based on meaningful rationality assumptions, but which is on much more shaky ground if rationality assumptions no longer provide clear guidance about how to formulate the behavioral assumptions.

The noncooperative solution concept now generally used was specified and the existence of solutions of this type investigated by Nash (1951). It is described by Nash as a concept relevant to a theory "based on the absence of coalitions in that it is assumed that each participant acts independently, without collaboration or communication with any of the others". It is a generalization of the solution of a two-person zero-sum game to a case with an arbitrary number of players and in general a variable sum. In economic theory this type of equilibrium is often referred to as a Cournot-Nash equilibrium, thereby associating it with the solution to the oligopoly problem proposed by Cournot. For some purposes, and with the appropriate explanation, such an association may be perfectly all right. However, this association may also have given rise to superficial interpretations of what a noncooperative equilibrium is.

In the following I begin by defining a noncooperative equilibrium and discussing some of its properties. Next, I continue the discussion by exposing some misunderstandings and confusing statements which are frequently found in the literature, and explain why I consider them to be superficial or confusing. A special section is devoted to the crucial question: is the decision made by a player in a Nash noncooperative equilibrium an individually rational decision? This question does not seem to have been faced squarely in the literature.

II. Definitions

Let there be n players, and let the decision or action taken by player i be a_i, where a_i belongs to the set A_i of possible actions for player i, $i=1,2,\dots,n$. Let the pay-off or utility of player i be represented by the function

$$W_i = W_i(a_1,\dots,a_i,\dots,a_n) \quad (i=1,\dots,n) \tag{1}$$

so that player i aims at achieving a value W_i which is as high as possible.

The functions in (1) represent a sort of "reduced form". We may imagine a system in which the actions $a_1,\dots,a_n$ jointly determine a state x, where again each player has a preference function as a function of x. The functions in (1) are then to be understood as functions of x, which are again functions of $a_1, \ldots, a_n$.

The players are now in a situation of interdependence or interaction in the sense that the outcome for each player will in general depend on the actions taken by other players as well as on his own action.

We assume the game to be noncooperative as defined in the introduction; this now means that each player i decides about his action a_i without communicating with other players. In other words, when player i is to decide on his action a_i, he has not received any information from the other players about their actions. (On the other hand, we assume that each player is fully informed about the action possibility sets and preferences of all players.)

The decision situation in which the players find themselves can be characterized by quoting the very apt description given by von Neumann and Morgenstern: "Thus each participant attempts to maximize a function of which he does not control all variables. This is certainly no maximum problem, but a peculiar and disconcerting mixture of several conflicting maximum problems. Every participant is guided by another principle and neither determines all variables which affect his interest. This kind of problem is nowhere dealt with in classical mathematics." This warns us that there is something essentially new in the noncooperative game situation as compared with classical assumptions about decision-making in economic theory. In classical theory, each agent has a well-defined maximum problem, for utility or profit. It may be complicated because of the functional forms, the forms of constraints, etc., but in principle it is a straight-forward maximum problem. This is not so in a noncooperative game situation. Thus, the introduction of noncooperative game considerations into economic theory is not just a variation or generalization of classical assumptions about maximizing behavior and individual rationality; it brings in something qualitatively new. Of course, there is still maximizing behavior in the sense that each player tries to maximize his W_i, but the players are interlinked in a different way and the whole setting differs from classical individualistic maximizing behavior.

In the notations introduced above, a noncooperative *equilibrium* can now be defined as follows:

A set of decisions $\hat{a}_1, \ldots, \hat{a}_n$ form a noncooperative equilibrium if and only if the following holds for $i=1,2,\ldots,n$:

$$W_i(\hat{a}_1, \ldots, \hat{a}_i, \ldots, \hat{a}_n) = \underset{a_i \in A_i}{\text{Max}}\, W_i(\hat{a}_1, \ldots, a_i, \ldots, \hat{a}_N). \tag{2}$$

Nash's explanation of the definition is that each player's decision "maximizes his pay-off if the strategies of the others are held fixed. Thus each player's strategy is optimal against those of the others."

We also introduce the concept of a noncooperative *solution*:

Suppose that the conditions in (2) for $i=1,2,...,n$ determine $\hat{a}_1,...,\hat{a}_n$ uniquely. Then we say that $\hat{a}_1,...,\hat{a}_n$ represent a noncooperative *solution*. (3)

The question of uniqueness or nonuniqueness is of particular significance in this context. Suppose that two different sets of decisions $\hat{a}_1,...,\hat{a}_n$ and $a_1^*,...,a_n^*$ satisfy the conditions in (2). Since the decisions are taken by individual players without communication with the other players, it is impossible in this situation to know whether an equilibrium will be reached. A player i will not know whether to decide on $\hat{a}_i$ or a_i^*, or perhaps also some other a_i, since the other players do not necessarily all choose decisions from one and the same of the sets $\hat{a}$ and a^*; and against a mixture of decisions from $\hat{a}$ and a^* by other players, the best decision of player i is not necessarily $\hat{a}_i$ or a_i^*. Now the same consideration also holds for the other players, so the whole decision situation is rather confusing. Clearly there is a fundamental difference between the case of uniqueness in (2) and the case of nonuniqueness. (Now we may have the case of "interchangeable" equilibria. This is the special case in which each W_i takes the same value for all $a_1,...,a_n$ composed of elements from the equilibrium sets such as $\hat{a}$ and a^*. In this case, it does not matter as far as the pay-offs are concerned which actions the various players take, as long as they choose actions from the equilibrium sets. This kind of nonuniqueness is therefore not problematic, and we may say that we have a noncooperative solution if we have a unique noncooperative equilibrium or if we have a set of noncooperative equilibria which are interchangeable. In the following we disregard the case of interchangeable equilibria since it is not of any great importance for the points to be made.)

Before proceeding to further interpretation and discussion of the equilibrium concept, we should—for the sake of completeness—mention that the decisions $a_1,...,a_n$ as introduced above can be interpreted as pure or mixed strategies. An equilibrium in pure strategies may fail to exist in rather natural formulations of economic models. On the other hand, an equilibrium will exist in most naturally formulated cases if mixed strategies are permitted. However, the uniqueness issue still remains, i.e. also when mixed strategies are permitted, we may very well have nonunique equilibria. The problem of nonuniqueness will enhance some of the problems about the rationality concept discussed in the last part of this paper. However, it is not a main concern of the discussion.[1]

[1] By this limitation of the discussion I do not want to belittle the importance of the nonuniqueness problem. Some further remarks are given in Section V.

III. A Sample of Quotations

I now turn to a sample of quotations which I think are rather representative of how noncooperative equilibria of the Nash type are introduced or explained in economic literature. I find them all unsatisfactory in one way or other, as they give misleading explanations of what a noncooperative equilibrium is. All the quotations are taken from respectable sources, but it is unnecessary to give precise references since the literature abounds with similar expressions. I group the expressions according to certain characteristic features. Here and there I have changed a couple of words so as not to draw attention to irrelevant details, but nowhere has the essence been changed.

Actions by other players taken as given. The most common explanation of a Nash noncooperative equilibrium is that it is the set of actions which is such that each player optimizes with respect to his own decision while taking the decisions by other players as given. A typical example is: "In other words, $\hat{a}$ is a noncooperative equilibrium if each player has no interest in changing his action when he considers the actions of the other players as given." Another example: "Any situation where each player takes as given the optimal choice of the other players and where neither can, under that assumption, increase his profits by altering his own strategy is called a Nash-Cournot equilibrium." Many explanations simply say that a Nash equilibrium is a set of decisions in which each player's decision is optimal given the decisions of all other players.

Reasoning about what a player will do "if he knows the strategies chosen by other players". This type of explanation is somewhat similar to the one above about "given" decisions by other players. One example runs as follows: "The conditions of the definition call for a type of circular stability. If the first player is aware that the other players are going to select $\hat{a}_2, \ldots, \hat{a}_n$, the first player will select $a_1=\hat{a}_1$ when he maximizes his own payoff. If the second player is aware that the other players are going to select the alternatives $\hat{a}_1, \hat{a}_3, \ldots, \hat{a}_n$, then the second player will select $a_2=\hat{a}_2$ when he maximizes his payoff; and similarly for the other players." This description is incomplete since it only says what a player will do "*if* he knows ...", and such knowledge is not available. In fact, a paper comments on this aspect and complains that "unfortunately there is no indication of how or why the players will generate expectations about the decisions to be taken by the other players". Another version of this idea is the following: "The Nash strategy is the best strategy for each player *if* it is assumed that all the other players are holding firm to their own Nash strategies."

"*No incentive to alter the strategy.*" The most common explanation of the Nash equilibrium is perhaps that "if the players are at $\hat{a}_1, \ldots, \hat{a}_n$, they are in equilibrium—neither has any incentive unilaterally to alter his strategy". Or similarly: "Thus a Nash equilibrium is a strategy that no individual has an incentive to deviate from, provided he assumes that all other players do not alter their strategies," or, "the actions chosen by the players constitute a Nash noncooperative equilibrium if no player can unilaterally improve his situation as long as others do not change their actions". Some explanations belonging to this family seem to imply a situation which lasts over time or repeats itself so that the actions of other players can actually be observed, while others should perhaps be read more as referring to hypothetical comparisons of alternative decisions.

The types of definitions reviewed above assume that a player considers actions to be taken by other players as "given", or they use more cautious *if*-statements. Other definitions are more explicit about the assumptions or beliefs which a player holds concerning the actions of other players. The Nash equilibrium is often regarded as "one particular member of a class of solution concepts which may be termed *conjectural equilibria*". Such an equilibrium is defined as a position "where no individual wishes to change his strategy, given some beliefs (conjectures) concerning the way other individuals will react to such a change".

One version of this is that *each player assumes he cannot influence the decisions of other players*. One author writes: "The attraction of the Nash-Cournot equilibrium concept is well known: It is the best strategy to follow if the player cannot influence the strategies of the other players." Another quotation from the same family: "The conventional Nash equilibrium assumes that each player believes that the actions of the other players will be unchanged as a result of his action."

A player ignores the effects on other decisions. The quotations just given refer to the beliefs held by a player about influence or noninfluence on other players. Other definitions say that a player "ignores" such effects without saying whether a player believes this to be correct or not. For instance, "Cournot originated and Nash generalized the assumption that each player ignores the effects of his actions on the strategies of the others."

Naive or myopic behavior. When the beliefs about actions to be taken by others, or about the reactions of other players to one's own decision are as simple as those just described, then it is tempting to declare this behavior naive or myopic. This is often done in discussions or interpretations of noncooperative solutions of the Nash type. One author states flatly that "the Nash assumption assumes naivity on the part of the players", and then introduces other assumptions which are meant to represent "a more realistic level of sophistication by the players". Another author writes about the "naive response structure of the Nash equilibrium concept". A third author explicitly mentions that it will be impossible for a player to find out that naive beliefs are incorrect: "Notice that while each player naively believes that he is facing a fixed strategy, in equilibrium it will not be possible for either player to recognize that he holds incorrect beliefs about the other player's behavior." The term "myopic" conveys much of the same idea, for instance in the expression "the myopic spirit of Nash's equilibrium concept" used by one author.

I think this sample of quotations is representative of the ways in which the concept of noncooperative equilibrium of the Nash type is introduced or explained in connection with applications in economic theory. The quotations are chosen more or less at random, and it would be very easy to find many others which convey the same meanings or impressions. I have tried to group them according to some characteristic elements. Such a classification is, of course, not unambiguous, but I think it will help in the following discussion.

Now my contention is that all the types of explanations illustrated by the quotations above, in one way or another, give false impressions of what a noncooperative equilibrium is or of the assumptions underlying the definition of this type of equilibrium. The implication is not necessarily that all the authors of these quotations, and an almost unlimited number of similar expressions, are confused. Some (but not all) of the expressions used correspond in a way to the mathematical formulation given by (2) and can perhaps be defended as short and convenient references. Some authors may have in mind a problem which involves the possibility of an adjustment process in time rather than a problem of noncooperative equilibrium of the Nash type in the strictest sense, and then perhaps some of the expressions may be defensible. However, regardless of whether or not the authors have correct notions in mind, readers may easily become confused.

If the Nash type of noncooperative equilibrium makes arbitrary assumptions about things taken as given; if it assumes false conjectures, or that players ignore some effects of their actions; if it assumes players to be naive or myopic—then presumably there should be other concepts which are more satisfactory. In contrast to this apparently obvious conclusion, one author writes that "if players do, in fact, have complete information, then Nash equilibrium seems virtually the only way to model noncooperative behavior". The arguments to be adduced in the following sections tend to support this view.

Perhaps it should be stressed that I do not think there are many misunderstandings about the mathematical formulation of the definition. Accordingly, the misunderstandings exemplified do not lead to formal errors in the analyses, given that the Nash equilibrium concept is used. The question concerns interpretations and justifications, and is thereby of relevance to the question of whether or not to use the Nash equilibrium theory. In many cases I would argue that the justification for using the Nash equilibrium concept is stronger than the authors in question suggest.

IV. Discussion

I will now discuss the meaning of a noncooperative equilibrium of the Nash type more carefully, and explain why I think the various interpretations illustrated by the quotations in the preceding section are misleading.

The formal definition is given by (2). When we do not explicitly state the opposite, we assume that the conditions in (2) determine the actions $\hat{a}_1, \ldots, \hat{a}_n$ uniquely so that we have a solution according to (3).

Mathematically it is, of course, true that the decision $\hat{a}_i$ maximizes W_i if $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ are given and known to be $\hat{a}_1, \ldots, \hat{a}_{i-1}, \hat{a}_{i+1}, \ldots, \hat{a}_n$. Nevertheless, as a description of the decision-making situation of player i, it is misleading to say that player i chooses a_i so as to maximize W_i, "given the actions of other players". The very essence of the noncooperative situation is that each player has to make his decision "independently, without collaboration or communication with any of the others" (cf. the quotation from Nash in the introduction). Accordingly player i, in deciding on a_i, does not know $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$. He only knows the action possibility sets $A_1, \ldots, A_n$ and the functions $W_1(a_1, \ldots, a_n), \ldots, W_n(a_1, \ldots, a_n)$. In making his own decision, he must use his knowledge about the action possibilities and utility functions. This may take the form of figuring out what the other decisions will be, but they are not "given" in his problem; instead they are part of the problem.

One might perhaps imagine a two-step procedure: 1. Predict the decisions $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ to be taken by the other players. 2. Take these predictions as given, and maximize W_i with respect to a_i. However, such a two-step procedure is illogical. The other players are assumed to be conscious players in similar situations as player i. Accordingly each of the other players must give some thought to what player i is going to do when they decide about their actions. It is then impossible for player i to predict correctly the actions of the other players without having given some thought to what he himself is going to do.

This leads to the conclusion that player i must consider all the decisions $a_1, \ldots, a_n$ simultaneously when deciding about his own action a_i. In other words, he must consider the full problem of determining all $\hat{a}_1, \ldots, \hat{a}_n$ according to the conditions in definition (2).

The contrast between the oversimplified explanation of the Nash equilibrium and the correct notion is clear from the following observation. If it were true that a Nash equilibrium is a situation in which each player considers the actions of the other players as "given", then the analytical problem for each player would simply be to maximize one function with regard to one's own decision. In fact, since a player does not know the decisions to be taken by others, each of them must solve the full analytical task of determining all $\hat{a}_1, ..., \hat{a}_n$ which are jointly or simultaneously determined by the conditions stipulated in (2) for $i=1, ..., n$. In other words, each player independently solves the same full problem, i.e. the "peculiar and disconcerting mixture of several conflicting maximum problems" (cf. the quotation from von Neumann and Morgenstern in Section II).

This discussion has also served to comment on the explanations which define the Nash equilibrium by referring to what a player will do "if he knows the strategies chosen by other players". Strictly speaking, such a definition is incomplete because the very essence of the noncooperative situation is that a player does *not* know the strategies to be chosen by the other players. But the definition implies that the players will generate expectations about the decisions to be taken by the other players. As already explained, each player determines his own action by solving the full problem involved in definition (2), i.e. expectations about decisions to be taken by other players are generated by the same operation which determines one's own action. (This is the answer to the author who complains that "unfortunately there is no indication of how or why the players will generate expectations about the decisions to be taken by other players".)

As illustrated by quotations in the preceding section, many definitions of a Nash equilibrium are rendered in terms of actions which are such that no player has any incentive to change his action unilaterally as long as all other players do not alter their actions. Taken literally, this involves some possibility of inspecting the decisions of other players, which is alien to the concept of a Nash equilibrium as defined by (2). If we read such definitions as referring to hypothetical comparisons of alternative decisions, then they are in a sense incomplete. It is, of course, true that player i will maximize W_i with respect to a_i, and then reach the decision $\hat{a}_i$ *if* he believes that the other players will take actions $\hat{a}_1, ..., \hat{a}_{i-1}, \hat{a}_{i+1}, ..., \hat{a}_n$. But are there any reasons why he will actually believe this? And what does it mean to say that "he assumes that all other players do not alter their strategies"? This way of expressing the concept suggests the possibility of a two-step procedure as mentioned above (first predicting the actions to be taken by other players, and then determining one's own decision), which is illogical.

As also illustrated by a few of the quotations, some explanations of the Nash equilibrium refer to the effects of a player's decisions on the actions of other players. Some say that a player believes that the actions of other players *will be unchanged* as a result of his own actions; others say that a player *cannot influence* the strategies of other players; still others say that a player *ignores* the effects of his action on the strategies of others. Strictly speaking, these explanations are not meaningful. In a noncooperative situation the players make decisions independently, with no communication. A player does not know the actions to be taken by other players until they all reveal their decisions; he can only analyze the situation on the basis of his information about the action possibility sets and utility functions, and on this basis make his own decision. A player's action possibility set A_i and his utility function $W_i(a_1, \ldots, a_n)$ will influence the decisions of other players, but there is no way in which we can speak meaningfully of a player influencing the decisions of other players through his own decision a_i. Other players are influenced by calculations implying a prediction of what a_i will turn out to be, but these calculations are based on information about the action possibility set A_i and the utility function W_i, not about the actual decision concerning a_i. The very definition of the game situation itself is such that a player cannot influence the decisions of other players through his own decision. It is then misleading to say that he *believes* he cannot influence the actions of other players, or that he *ignores* the effects of his action on the strategies of others. To say that the Nash strategy is the best strategy to follow "if the player cannot influence the strategies of the other players", is to say that the Nash strategy is the best strategy, since the if-condition is actually fulfilled by the very definition of a noncooperative game situation; but perhaps this is not how it is meant by authors who use this expression.

In view of the discussion above, I think it is clear that behavior in accordance with the Nash equilibrium theory should not be characterized as naive or myopic. In making his own decision $\hat{a}_i$ a player takes into account all relevant information; he does not assume given decisions of other players, and he does not make any simplistic forecasts of their behavior. On the contrary, he considers the decision-making of other players as intertwined with his own decision-making. As illustrated by quotations in the preceding section, some authors state that a player in a Nash equilibrium "naively believes that he is facing a fixed strategy", but that it will not be possible for the players to discover that they hold incorrect beliefs about the other players. In calculating his own decision $\hat{a}_i$ a player at the same time calculates the decisions to be taken by the other players. When the decisions are implemented, each player will observe the decisions taken by the other players, and all players will see their expectations about other players fulfilled. So far there are no "incorrect beliefs". Now, what does it mean to say that a player believes that he is "facing fixed strategies"? In a game situation with Nash players each player plays his Nash strategy $\hat{a}_i$, and he will discover only $\hat{a}_1, \ldots, \hat{a}_{i-1}, \hat{a}_{i+1}, \ldots, \hat{a}_n$ for decisions by other players, confirming his expectations. If he, for some experimental purpose, should try a strategy other than $\hat{a}_i$, the rest of the players would use the same strategies if they were not aware of the fact that a certain player had now decided to experiment. On the other hand, if they knew that he would experiment, and had some basis for guessing what the experiment would consist of, then the other players would in general choose strategies other than those in the Nash equilibrium. But the more meaningful way of asking whether or not other players keep fixed strategies would be to compare what happens under different preference functions and/or action possibilities. If a player i had a utility function W_i^* and a possibility set A_i^* other than the original W_i and A_i, then not only would the decisions of player i himself be different, but so would, in general the decisions made by all other players. In this sense the strategies of other players are not "fixed". Furthermore, player i is fully and correctly aware of how the decisions to be made by other players depend on his own utility function and possibility set; in other words, he does not naïvely believe that the strategies of other players are given or kept fixed, uninfluenced by his own preferences and/or action possibilities.

Instead of saying that players who behave according to the Nash equilibrium concept are naive or myopic, I think it is more correct to say that they are very sophisticated players. They use all relevant information, they perceive correctly the interrelationships between the players, they make correct predictions of the decisions of all players involved, and in doing so they realize that the decisions made by other players are influenced by the data characterizing one's own situation.

If the Nash equilibrium theory had implied naive or myopic behavior, then the theoretical status of this equilibrium concept would not be very strong, and there would be good reasons for trying to define other types of equilibria which could not be discarded on such grounds. Now according to the discussion above this is not the situation; Nash players are sophisticated players who are fully aware of and take into account the interdependencies in which they are involved. However, we have not yet considered explicitly the fundamental question: if a player in a noncooperative situation wishes to pursue his own interests in a rational manner, should he then choose the strategy which corresponds to the Nash equilibrium? In other words, is $\hat{a}_i$ as defined by (2) a rational decision from the point of view of player i? This question will be considered in the following section.

V. The Individual Rationality of Nash Strategies

In a noncooperative game, the individual player is placed in a decision situation in which he has to take an independent, individual decision. His aim is to arrive at a result which is as good as possible in terms of his own preferences. The question of rationality must then be a question of *individual* rationality. However, the question of what should be meant by rational behavior is much more complicated in this case than in connection with maximizing behavior in standard economic theory. The reason is that the player perceives the game situation correctly and, accordingly, is aware of the fact that other players are also involved and will influence the outcome by their decisions. It is the fact that these other decisions cannot be considered "given" from the viewpoint of the individual decision-maker that makes it somewhat problematic to define what should be meant by rational behavior.

This point is often not made quite clear in the literature. For instance, Gabay & Moulin (1980)[2] refer to "the rationality postulate of noncooperative behavior", but this is stated simply as the assumption that each player chooses the strategy which maximizes his utility level "given the decisions of the other players". The problematic aspect of rationality in the present case is precisely what should be meant by rational behavior when the individual decision-maker *cannot* take the decisions of the other players as given, but must nevertheless somehow take into account the fact that there are other decision-makers who influence the outcome.

We could try to define rational behavior by imposing certain postulates. I propose the list given below.

Postulate 1. A player makes his decision a_i, where $a_i \in A_i$, on the basis of, and *only* on the basis of information concerning the action possibility sets of all players, $A_1, \ldots, A_n$, and the preference functions of all players, $W_1(a_1, \ldots, a_n), \ldots, W_n(a_1, \ldots, a_n)$.

This postulate means, in the first place, that the player does not use irrelevant information for his decision. Furthermore, it implies (consistently with the points made in the discussion above) that the individual player does not possess direct information about decisions of other players. If he wants to predict the decisions of other players, he must do this on the basis of the information included in postulate 1. In this case and in the sequel, I formulate the conditions and arguments without referring explicitly to mixed strategies. However, it is clear that the reasoning will cover the case of mixed strategies. If a mixed strategy is chosen, then the probabilities involved will be determined on the basis of the information concerning possibility sets and preference functions. The use of probabilities in such mixed strategies is not regarded as decisions based on "irrelevant information".

Postulate 2. In choosing his own decision, a player assumes that the other players are rational in the same way as he himself is rational.

This postulate implies a sort of symmetry. All players are rational in the same way, and it is then part of the rational behavior of the individual

[2] In this case I refer to the authors and the paper, since the paper is a very valuable reference concerning the existence of noncooperative equilibria and the stability of such equilibria when they are regarded as having been established through a dynamic process.

player to recognize and take into account the rationality of other players. This is not intended to imply that other situations cannot occur in practice. The question of finding a rational action in a noncooperative game situation is not an easy question, and players may well have different abilities in this respect. If one player knows that other players are less able in this respect, it is of course rational to exploit this fact. However, the theoretically more interesting situation is the one in which all players are equally able to analyze the situation and find a rational strategy (if such strategies can be meaningfully defined). Referring to postulate 1, one player's exploitation of the fact that other players are less able, would mean use of information other than that contained in postulate 1.

Postulate 3. If some decision is the rational decision to make for an individual player, then this decision can be correctly predicted by other players.

It is assumed that all players have the same and full information about the situation, i.e. each player knows the action possibilities and preferences of other players as well as his own. If a player can analyze the situation and find that a certain action is the rational decision on his part, then other players can imagine themselves in his place and duplicate his analysis. Postulate 3 operates in all directions. A player can predict the decisions that will be taken by other players, but he also knows, in conformity with postulate 2, that the other players can predict his own decision.

Postulate 4. Being able to predict the actions to be taken by other players, a player's own decision maximizes his preference function corresponding to the predicted actions of other players.

This postulate corresponds, of course, to the "rationality postulate of noncooperative behavior" mentioned above. Here, however, it is placed in a context where we explicitly recognize that the decisions by other players have to be predicted in such a way that the other postulates 1–3 are also satisfied. It is this whole procedure which must be judged in terms of whether it is individually rational or not.

Postulate 4 is essential for the noncooperative character of the game. It implies that each player unscrupulously pursues his own aim. It prevents behavior whereby some players, by individual decisions, deviate from it, each of them hoping that others will deviate in suitable ways so that everyone benefits from it. Assumption 4 alone implies that if one player predicts that other players will make such an attempt, then he will take advantage of this situation in the best possible way for himself as measured by his own preference function. In this connection it should be kept in mind that no player can influence the decisions of other players by his own actual *decision*. When a player wants to achieve as high a value as possible of his own preference function, then postulate 4 seems unquestionable. On the other hand, these preferences can, of course, reflect some concern for the welfare of other players. So when a player "unscrupulously pursues his own aim", this need not be a selfish aim.

A postulate which gives a rather attractive characterization of rational individual behavior can be stated in the following way:

Postulate 5. A decision is rational if the player, after having observed the decisions taken by other players and the outcome of the game, does not regret the decision he has made.

This definition is perhaps not very operational since the term "regret" is not quite clear. However, some meaning can be attributed to it by interpreting the expression "does not regret" as saying that if the player was again (unexpectedly) put in the same decision situation, he would make the same decision. Regardless of whether or not one considers this to be a sufficiently clear characterization, the intentions behind the formulation in postulate 5 are in any case covered by postulates 1–4 above. It is therefore not a new and independent statement. If a player should regret his decision, it must be either because he made incorrect predictions about the decisions of other players, or because he did not take the consequences of the predictions in an optimal way as defined by postulate 4. In other words, a situation in which a player regrets his decision can occur only if some of the postulates already given are violated. On the other hand, if 1–4 are fulfilled, then a player has no reason to regret his decision. He has then correctly predicted the decisions of other players, and made the best out of the situation.[3]

The question now is what relations exist between a decision according to the Nash equilibrium strategies and the requirements listed above.

First, it is clear that a decision in accordance with the Nash theory satisfies requirements 1–4 (and accordingly also 5). Player i calculates

---

[3] The meaning of the considerations which refer to "regret" are not quite unambiguous when we have an equilibrium in mixed strategies. In such cases a player will not "observe the decisions taken by other players" in the sense that he observes the probabilities constituting the mixed strategies. He will only observe the realizations of the randomized decisions. If he misinterprets these realizations to be a set of pure (nonrandomized) decisions, then he may have reason to regret his own decision. The "regret" consideration in the case of mixed strategies must be formulated in the following way: if a player were informed about the probabilities constituting the mixed strategies of other players, he would not regret his own choice of strategy (pure or mixed); i.e., if he was (unexpectedly) placed in the same situation again, he would behave in the same way. However, since the observations in the case of mixed strategies do not in fact reveal the behavior of other players, the regret considerations in this case are more hypothetical and less operational than in the case of an equilibrium in pure strategies. Thus, the conditions stipulated by postulates 1–4 are more appropriate that 5. In 3 and 4, "prediction" means prediction of the probabilities constituting the mixed strategies of other players, when mixed strategies are actually used. (This note has been stimulated by some remarks made by Professor Karl Borch.)

his decision $\hat{a}_i$ as an individual decision, but does so by calculating the full set of decisions $\hat{a}_1, \dots, \hat{a}_n$ according to definition (2), where $\hat{a}_1, \dots, \hat{a}_{i-1}, \hat{a}_{i+1}, \dots, \hat{a}_n$ are predictions of decisions to be taken by other players. The decision is calculated on the basis of the information listed in postulate 1; it assumes that other players take their decisions in the same way as player i himself; it correctly predicts the decisions of other players, and admits that other players correctly predict player i's own decision; and, corresponding to the predictions of the behavior of other players, $\hat{a}_i$ is the best decision player i can make.

We have assumed that the noncooperative equilibrium is determined uniquely by the requirements in definition (2). Decision $\hat{a}_i$ is then also *the only* decision which satisfies requirements 1–4. Suppose, tentatively, that a decision $a_i' \neq \hat{a}_i$ is a rational decision on the part of player i. Then player i knows that the other players can predict the decision, according to postulates 2 and 3. Player i could again, according to the same conditions, correctly predict what the other players would do, and would accordingly envisage a set of decisions $a_1', \dots, a_n'$, where all a_j' for $j \neq i$ are the predictions of the decisions of players other than i. But since predictions are correct, they are also the actual decisions of other players. Now since $\hat{a}_1, \dots, \hat{a}_n$ is the unique set of decisions determined by (2), the set of decisions $a_1', \dots, a_n'$ cannot satisfy all requirements in (2). Accordingly, for at least one player k, we have that his decision a_k' does not satisfy requirement 4, i.e. that the decision should be the best possible one corresponding to the predictions which refer to the decisions of other players. Player no. k may be player i himself, i.e. $k=i$, in which case postulate 4 is violated by the decision a_i'; or k may represent another player, in which case condition 2, that player i should regard the other players as rational in making predictions of their behaviour, is violated. (We may, of course, have that a_i' violates condition 4 for player i at the same time that the corresponding condition is violated for one or more other players $k \neq i$.) This argument shows that assuming that a_i', which is different from $\hat{a}_i$, is a rational decision for player i, leads to a contradiction.

We can summarize the results above in the following proposition:

Proposition. Suppose that the requirements in definition (2) of a Nash type of noncooperative equilibrum determine the set of decisions $\hat{a}_1, \dots, \hat{a}_n$ uniquely. Then the decision $\hat{a}_i$, considered as an individual decision taken by player i, satisfies the rationality conditions formulated in postulates 1–4 (and accordingly also 5). Furthermore, the decision $\hat{a}_i$ is the only decision which satisfies these requirements.

The question initially raised in this section is whether we can say that an action corresponding to the Nash noncooperative equilibrium is a rational action from an individual point of view. Our proposition provides a provisional answer: *If the rationality conditions stipulated by postulates 1–4 are sufficient for rationality,* and if the Nash equilibrium is unique, then the Nash strategy is an individually rational strategy in a noncooperative game situation where all players are fully informed and rational.

The requirements stipulated may also be considered *necessary* for rationality. Since the Nash strategy is the only strategy which satisfies the requirements, we are then in the following situation. We have to accept the Nash strategy as the individually rational strategy, or we have to give up the idea of finding a solution to the problem of determining a strategy as an outcome of individual rationality.

Personally I find the conditions appealing as necessary and sufficient for individual rationality, i.e. as a definition of the concept. In this case the Nash behavior is *the* individually rational behavior. Accordingly, it is not based on naive beliefs, false conjectures about other players' behavior, ignoring interdependencies or any of the other weaknesses or arbitrary elements suggested by some of the quotations in Section III.[4]

A consideration which has something in common with the conclusion above was suggested in the well-known book by Luce and Raiffa (1957). After having defined and reviewed various aspects of the Nash noncooperative equilibrium theory, they put forward the following consideration: "Nonetheless, we continue to have one very strong argument for equilibrium points: if our noncooperative theory is to lead to an n-tuple of strategy choices and if it is to have the property that knowledge of the theory does not lead one to make a choice different from that dictated by the theory, then the strategies isolated by the theory must be equilibrium points." It would have been interesting to see this argument worked out in detail. There seems to me to be a danger of running into a circular argument by following this line of thought. But suitably interpreted it is similar in nature to the arguments of this paper. "Knowledge of the theory" would show the

[4] After having mentioned so many expressions which give a misleading impression, it should also be noted that there are explanations given in the literature which are more satisfactory. For instance, Friedman (1977*a*) writes about the Cournot equilibrium in oligopoly theory (which in the case mentioned is the same as the Nash equilibrium) in the following terms: "As an equilibrium for a one-period market with simultaneous decision and in which it is assumed impossible for the firms to collude, it is hard to imagine one which fails to be a Cournot equilibrium. The beauty of it lies in its inherent believability. Nowhere is there false information or firms acting on the premise that their rivals are less perceptive than they. ... Put another way, if you impose some standards of reasonableness on your own choice and assume others will use these standards too, then an output vector which is not a Cournot equilibrium will not look reasonable." Although the viewpoint is not spelled out in the same way as in the present paper, Friedman's view appears to be close to the one taken here. See also Friedman (1977*b*), especially pp. 24–25.

players the consequences of Nash behavior, but would only confirm what Nash players already understand and would not induce them to behave differently. This would have been otherwise if the Nash theory had implied naive or myopic players. Knowledge of the theory could then make them less naive or less myopic, and accordingly make them change their behavior.

The conclusions suggested above are based on the assumption that the noncooperative Nash equilibrium is unique. They are equally valid for the case in which the equilibrium is not unique, but where all equilibria are interchangeable. Furthermore, as suggested before, the reasoning is equally valid when we admit mixed strategies as it is for the case of only pure strategies,[5] and it is equally valid for strategies in dynamic or repeated games as for the case of one decision.

If the noncooperative Nash equilibrium is not unique, and the equilibria are not interchangeable, then the reasoning breaks down. When the equilibrium is unique, all players solve the same problem, determining all actions, and they all reach the same result, thus predicting other players' actions correctly in the same operation as they determine their own actions. When there are several equilibria and they are not interchangeable, so that it matters which action is taken from among the equilibrium sets, then each player will be unable to predict the decisions of other players unless we introduce some further elements into the theory. This is a difficult point which Harsanyi in particular has pursued; see for instance Harsanyi (1979). There are very natural game situations which involve this multiplicity problem, as it has been called by Harsanyi. This refers *inter alia* to attempts to determine bargaining strategies by means of noncooperative game theory[6] and cases of repeated game situations regarded as a "supergame". A convincing general theory of what an individually rational action is in cases where we encounter the multiplicity problem, is hard to imagine. Here, we are definitely in a field where the idea of letting a theory be based on assumptions about individual rationality is not a sufficient basis, and perhaps not even a meaningful basis. The classical paradigm of economics based on rational individual decision-making encounters severe difficulties if we try to extend it into such fields of economic interaction.

In concluding this section, a few more words on the classical rationality assumption in economics may be necessary.

In the introduction to the book *Philosophy and Economic Theory* (1979), editors Hahn and Hollis write that "pure theory is deeply committed to an assumption that economic behaviour is rational", and go on to offer the

---

[5] With the reservation discussed in footnote 3.
[6] See, for instance, the discussion in Johansen (1979).

following definition: "The pure economist's definition of rational choice is now this: Given the set of available actions, the agent chooses rationally if there is no other action available to him the consequence of which he prefers to that of the chosen action." They point out that, although the definition looks very simple, "it has striking implications", and it is of course clear that many of the propositions of economic theory are deducible from such an assumption about rational choice. However, considering the noncooperative game situation, the representative "pure economist's definition of rational choice" as given here is insufficient. The reason is that the choice situation of the player is not characterized simply by a set of possible actions, relationships between actions and consequences, and his own preferences. There are other players involved who influence the situation and whose possible actions must be taken into account, without communication and coordination, in interaction with his own action. A definition of rationality must then necessarily involve such elements as we have discussed in this section.

In a paper reproduced in the same book, Simon also discusses the rationality assumption of standard economic theory. He distinguishes between substantive and procedural rationality. Substantive rationality is defined by the following statement: "Behavior is substantively rational when it is appropriate to the achievement of given goals within the limits imposed by given conditions and constraints." He points out that one of the basic assumptions of classical economic analysis is that "the economic actor is substantively rational". This definition is formulated in a more flexible manner, and depending on what one includes under the "given conditions and constraints", the definition may or may not cover the case of noncooperative game situations. In any case, the definition does not give any precise description of what should be meant by rationality in the case of a noncooperative game situation. However, in Simon's opinion, the concept of substantive rationality does not provide an answer to what is rational in the noncooperative duopoly situation. In a discussion of the Cournot problem, he says that Cournot "identified a problem that has become the permanent and ineradicable scandal of economic theory". He points out that the notion of profit maximization is ill-defined in this case: "The choice that would be substantively rational for each actor depends on the choices made by the other actors; none can choose without making assumptions about how others will choose." He maintains the view that "it is generally conceded that no defensible formulation of the theory stays within the framework of profit maximization and substantive rationality", and furthermore that "game theory, initially hailed as a possible way out, provided only a rigorous demonstration of how fundamental the difficulties really are". In Simon's opinion we have to acknowledge "the impossibility of discovering at last 'The Rule' of substantive rational behavior for the oligopolist''. It is clear from these statements that Simon would not agree with the suggestion that the action implied by the noncooperative Nash equilibrium could be characterized as an individually rational action, although he does not explicitly refer to this possibility. Instead, Simon seeks the solution in the direction of ''procedural rationality'', an idea which may be valuable, but which is not pursued here since it implies, in a way, evading the problem posed by a noncooperative game situation.

A paper by Harsanyi (1966) should also be mentioned. It introduces several postulates of rationality which the author considers to be ''natural generalizations of the rationality postulates used in the theory of individual rational behavior''. However, the scope and ambition of the present paper are much more modest than those of Harsanyi's paper, which attempts to give a solution to the problem of rational behavior ''covering *all* game situations'', including cooperative games and bargaining as well as uncertainty, and which proceeds partly along Bayesian lines. More detailed comparisons with Harsanyi's paper would therefore take us beyond the scope and purpose of this paper.

VI. Summary and Concluding Observations

The concept of a noncooperative equilibrium of the Nash type is widely used in economic theory. Sometimes arguments are given in support of the use of this solution concept, and sometimes it is used as one solution concept among several others, without explicit arguments as to whether it is considered to be the most appropriate concept. Very many different types of interpretations or explanations of the noncooperative Nash equilibrium are found in the literature; the quotations in Section III provide examples. The majority of the interpretations or explanations given are unsatisfactory one way or other. These various, more or less misleading, interpretations are discussed in Section IV.

The question of whether a decision in accordance with the Nash equilibrium concept can be regarded as a rational decision is raised in Section V. Since the game underlying the solution concept is a noncooperative game, the question is about *individual* rationality. Some conditions are stipulated for rational decisions, and it is demonstrated that, provided that the equilibrium is unique, decisions in accordance with the Nash noncooperative equilibrium concept satisfy the requirements, and they are the only decisions which do so. It is concluded that if individual rationality is a meaningful concept in the context of noncooperative decision-making, then decisions in accordance with the Nash noncooperative equilibrium are individually rational. The case in which the equilibrium is not unique, which can arise under many natural circumstances, raises a more difficult issue. In this case it is doubtful if a postulate about individual rationality can have any good meaning.

This paper has been concerned mainly with the logical issues and matters of principals. It may be asked whether such decision situations are typical or occur frequently in economic decision-making in practice. In my opinion they are quite typical and frequent. But even if they are not considered very typical, it is important from a methodological and fundamental point of view to have a clear conception of the relationship between individual rationality and solution concepts in noncooperative games.

If there are severe difficulties in deriving individually rational decisions, or even in defining what this means, then decision-makers will in practice often increase the scope for actions and decisions so as to change the situation.[7] Postponing the decision, trying to establish contacts or finding out more about other players (perhaps not necessarily believing that other players are fully rational) and many other possible actions are relevant. Many such attempts to change the situation can be regarded as strategies in a game of a higher order. If this is done, then sometimes the arguments which we have adduced may be applicable at this higher level, but in some cases the higher-order game would be of a different nature in several respects. There are some recent trends in game theory which would be relevant to an extension of the analysis in such directions (and which could perhaps be related to Simon's idea about "procedural rationality"), but it is beyond the scope of this paper to pursue this further.

I have also not pursued the interesting task of discussing noncooperative equilibrium theory in relation to the theory of rational expectations, but I will only note that the recent preoccupation with rational expectations in many branches of economic theory is indirect evidence of the relevance and importance of noncooperative equilibrium theory.[8]

---

[7] When noncooperative behavior according to the Nash equilibrium theory leads to particularly bad results for all players, as in the Prisoner's Dilemma type of games, one may ask whether the players will find ways to change the situation by somehow moving out of the confines of the initially specified rules of the game. For instance, Borch (1968) argues as follows in connection with the outcome of a Prisoner's Dilemma game: "This is not a very satisfactory outcome, and it is difficult to accept it as the final solution to a game played by two rational persons. It seems tempting to assume that the players will somehow find a way to play the game in a cooperative manner." Such considerations have also been put forward by other authors. They may of course be very realistic in many cases. But they do not diminish the importance of clear conceptions about the nature of noncooperative solutions before one embarks on the more complex task of analyzing such higher-order games where the issue of cooperation versus noncooperation must be treated as an endogenous element.

[8] I have presented very briefly some viewpoints on the relationship between rational expectations and noncooperative equilibrium theory in Johansen (1982).

References

Borch, K. A.: *The economics of uncertainty*. Princeton University Press, Princeton, 1968.

Friedman, J. W.: Cournot, Bowley, Stackelberg and Fellner, and the evolution of the reaction functions. In *Economic progress, private values, and public policy. Essays in honor of William Fellner* (ed. B. Balassa and R. Nelson). North-Holland, Amsterdam, 1977 a.

Friedman, J. W.: *Oligopoly and the theory of games*. North-Holland, Amsterdam, 1977 *b*.

Gabay, D. & Moulin, H.: On the uniqueness and stability of Nash equilibria in noncooperative games. In *Applied stochastic control in econometrics and management science* (ed. A. Bensoussan et al.). North-Holland, Amsterdam, 1980.

Hahn, F. & Hollis, M.: Introduction. In *Philosophy and economic theory*. Oxford University Press, Oxford, 1979.

Harsanyi, J. C.: A general theory of rational behavior in game situations. *Econometrica*, 1966.

Harsanyi, J. C.: A new solution concept for both cooperative and noncooperative games. In *Rheinisch-Westfälische Akademie der Wissenschaften, Vorträge N287*. Westdeutscher Verlag, 1979.

Johansen, L.: The bargaining society and the inefficiency of bargaining. *Kyklos*, 1979.

Johansen, L.: Econometric models and economic planning and policy: Some trends and problems. Memorandum from the Institute of Economics, University of Oslo, 1982.

Luce, R. D. & Raiffa, H.: *Games and decisions*. Wiley, New York, 1957.

Nash, J.: Non-cooperative games, *Annals of Mathematics*, 1951.

Simon, H. A.: From substantive to procedural rationality. In *Method and appraisal in economics* (ed. S. Latsis), Cambridge University Press, Cambridge, 1976. Reprinted in *Philosophy and economic theory* (ed. F. Hahn and M. Hollis), Oxford University Press, Oxford, 1979.

First version submitted February 1982;
final version received July 1982.