Rationality and Emotions in Ultimatum Bargaining: Comment

Jean-Louis RULLIÈRE *

Introduction

Mr. Shmuel ZAMIR, who is renowned as both a game theorist, and an experimentalist, focused his lecture, given at the time of the *First Annales Conference*, on the correspondence existing between the proposals forwarded by non-cooperative game theory and the empirical results of work in experimental economics. More precisely, his objective has been to establish the conditions under which the use of experimental data can be made compatible, within the framework of non-cooperative game theory. From such a perspective, this article helps us nourish a debate in which SELTEN, one of the three 1995 Nobel Prize winners in Economics, sought to provide the conclusion by stating: *Game theory is for proving theorems, not for playing games*.

SELTEN's position acknowledges the presence of a divide, which has been widely addressed in the recent literature and which can be illustrated in particular by the opposing points of view expressed by GUTH [2000] and WEIBULL [2000]:

> *Orthodox game theory relies on perfect decision rationality, i.e. the unlimited cognitive and information processing capabilities of players. Even for finite games of perfect information like chess however, it is obvious that these requirements are far beyond what human decision-makers can accomplish […] then the game's theoretical predictions are often not confirmed by experimental observations.*
>
> GUTH [2000, p. 2]

* J.-L. RULLIÈRE: University of Lyon Lumière, Groupe d'Analyse et de Théorie Économique.

I would like to express my gratitude to M.C. VILLEVAL and N. MARCHAND for their valuable comments on the initial draft. Thanks are also to be addressed to D. ENCAOUA and S. ZAMIR for their participation in the *First Annales Conference*, held in July 2000. Any remaining errors are my full responsibility.

> *The author's conjecture is that if a game has been properly identified and made public to the subjects, then a vast majority of subjects would play the subgame perfection equilibrium in such simple game forms, and the same would presumably hold in many of the simplest game forms discussed in the experimental literature.*
>
> WEIBULL [2000, p. 16].

GUTH's position is tantamount to considering that it would be better to leave laboratory experiments to other social scientists, especially cognitive and social psychologists, who are properly trained to run them. This notion will be introduced into behavioral game theory since it often gets mentioned by participants when asked to explain their deliberation process. At the other end of the spectrum, inspired by WEIBULL's statement, the elegance of a traditional game theory defense lies in the normative power of this theory as regards perfectly-rational people playing games, in contrast with a positive theory that predicts actual behavior. As forwarded by NASAR [1998] and GOERE and HOLT [2000, p. 32]: *John NASH, one of the other Nobel recipients, saw no way around this dilemma, and when his experiments were not providing support for the theory, he lost whatever confidence he had in the relevance of game theory and focused on more purely-mathematical topics in his later research*.

In order to explain this irreconcilable split, it is interesting to note that attention generally gets focused on the Ultimatum Bargaining Game. As accurately pointed out by ZAMIR, this symbolic example has led a number of experimentalists to reject wholesale all precepts of the non-cooperative game theory without even seeking to qualify the condemnation among the Nash equilibrium, the Bayesian-Nash equilibrium, the subgame perfect Nash equilibrium and the sequential Nash equilibrium. It is abundantly clear that even under optimal conditions, the experimental results relative to the Ultimatum Bargaining Game can only lend support to refuting the subgame perfectness criterion. The Nash equilibrium has not been challenged herein, which is why ZAMIR considers that *rationality and self-interest behavior are alive and well, even in the Ultimatum Bargaining Game*.

The best means for bolstering this point of view consists of the lowest-level requirement for verifying that human experimental subjects should be placed in a laboratory environment resembling as closely as possible the theoretical description of the game. Yet, the constraints of the experimental method necessitate deviating from the hypothetical environment in which non-cooperative game theory places the players. This deviation gives credence to those game theorists who can thereby dismiss the experimental approach as an instrument for refuting game theory as well as to those game experimentalists able to cite this deviation when focusing upon elements of players' strategic behavior which were not taken into account by game theory.

The study of experimental protocols, which considers various simple game forms in general and the Ultimatum Bargaining Game in particular, has shown that the points of discord revolve around the three following primary elements: the experimental replication of an unrepeated game (*i.e.* one-shot games or repeated games); the interpersonal preference compatibility; and the strategic misrepresentation of preferences.

One-Shot Games or Repeated Games

Both the Nash equilibrium and the subgame perfect Nash equilibrium yield accurate predictions for standard versions of games such as the Ultimatum Bargaining Game. If these games were to be tested just once, this step would make it possible to avoid introducing a mutual learning effect on players' preferences, behavior and beliefs. In contrast, a confusion effect may be encountered precisely because players do not have the opportunity to become familiar with the game's structure through playing (ANDREONI [1995]). Moreover, from an experimental standpoint, this statement presupposes that the panel of subjects is large enough to ensure significance.

These two reasons are typically sufficient to incite experimentalists to prefer repeating the same game between twenty and thirty times by reshuffling the group of subjects at each repeat stage. This method enables the subjects to acquire mutual or quasi-mutual knowledge of one another's rationality, preferences and beliefs. In this respect, the requirement of a rationalistic and static interpretation of the Nash equilibrium is practically resurrected. However, the learning by repetition inherent in the standard game gives way here to the introduction of an incentive to cooperate and the expression of emotions or moral sentiments, like reciprocity, fairness, kindness, altruism or warm-glow.

From the strict perspective of game theory, the tested game is a game repeated a finite number of times and thus no longer corresponds to the game played just once. Furthermore, nothing guarantees that repetition of the standard game represents a repeated game from a pure game theory standpoint: it is necessary for overplayed preferences to be cumulative from one round to the next (WEIBULL [2000]). This is why NASH [1954][1] considers that repetition gives rise to a new set of interpretations (in terms of moral sentiments), which lies very far from the considerations applied to the one-time game: *The flaw in the experiment as a test of equilibrium point theory is that the experiment really amounts to having the players play one large multi-move game. One cannot just as well think of the thing as a sequence of independent games … there is too much interaction … One would have thought them more rational*.

For game experimentalists, on the other hand, these emotions can provide support in interpreting the theoretical prediction failure; such is the case with the Ultimatum Bargaining Game. However, this methodological position is a controversial one. As highlighted by ZAMIR, a sizable body of literature exists to justify the deviation between the theoretical prediction of the Ultimatum Bargaining Game and experimental evidence: the bulk of this literature concentrates on what ZAMIR has combined under the heading *emotion*. In this context however, it is possible to reconcile theory and experimental evidence. ROTH, PRASNIKAR, OKUNO-FUJIWARA and ZAMIR [1991] and ROTH and PRASNIKAR [1992] introduce a competition mechanism among 9 proposers and only one responder. Along the same lines, GUTH, MARCHAND and RULLIERE [1997, 1998] experiment with an Ultimatum Bargaining Game that features just one proposer and one responder, who is the winner of a competi-

1. This citation has been drawn from NASAR [1998, p. 119].

tion held among 5 responders. In both of these experiments, the competition is sufficient to reduce the hypothetical effect of moral sentiments and provides individual incentives to obtain convergence of individual behaviors towards the subgame perfect Nash equilibrium.

Lastly, this previous example demonstrates that the repetition argument remains paradoxical: explanations in terms of emotion are not in and of themselves sufficient to challenge the hypothesis of test subject rationality, yet a change in the game's institutional design reveals that emotions lack the robustness necessary to determine player behavior.

Interpersonal Preference Compatibility

The recurrent observation of a deviation between the theoretical prediction regarding the Ultimatum Bargaining Game (played only once) and experimental results gives rise to a wide array of interpretations. Game experimentalists are in fact faced with the dilemma: how can the notion of rationality be extended? The two possible methodological strategies are as follows: either seek to extend, generalize or select a good candidate among the equilibrium concepts, or provide a more specific characterization to the utility functions.

As noted by ZAMIR, the methodological choice commonly opted for does not tend to favor the first approach; though the subgame perfect Nash equilibrium is being refuted, at the same time other candidate equilibrium concepts have not been included or proposed: *I think that neither NASH, who provided us with his concept of equilibrium, nor SELTEN, who introduced subgame perfectness, would say that subgame Nash equilibrium, even if it is unique, is the prediction of game theory. Even if we are testing game theory as a descriptive theory, there are other solution concepts in addition to subgame perfect equilibrium, some of which, I will argue, may be relevant to the Ultimatum Bargaining data*.

The current predominant choice in experimental literature calls for considering that rationality and self-interest coexist with other motives, which seem to be unrelated to rationality. The principal set of such motives belongs to the family of emotions. The notion of coexistence however remains ambiguous in that no clear-cut border can be drawn between the rationality of self-interest and the role of emotions. Our focus lies in determining whether or not emotions are perfect substitutes/complements to rationality.

It should be noted that the formal machinery of non-cooperative game theory does not require that preferences depend solely upon the rewards obtained at the conclusion of the game. Payoff values need not be identical to the monetary payoffs associated with each game sequence (WEIBULL [2000]). This theoretical observation was used as a basis by, among others, CHARNESS and RABIN [1999], FEHR and SCHMIDT [1999], and BOLTON and OCKENFELS [2000] to propose utility functions in which the subjects are not strictly motivated by their own monetary payoffs. In addition to worrying about their monetary payoffs, subjects are aware of the resultant material rewards to all participants. These kinds of utility functions allow for this analysis, for example, reciprocal fairness within a group of subjects. Nonetheless, this advantageous feature in modeling player preferences also happens to intro-

duce the interpersonal preference comparison. It should be pointed out that the Nash equilibrium enables to identify theoretical solutions whereby the subjects are not assumed to behave by resorting to interpersonal preference comparisons. This bold result of the Nash equilibrium is significant inasmuch as it is not necessary to assume that players rely upon some common norm, whose existence and basis must still be characterized, in order to appreciate their payoffs. This last hurdle is indeed difficult to surmount, as attested by WEIBULL ([2000], p. 9): *If the subject's preferences are mutually dependent, then we face a problem of self-reference in the very definition of a game*.

The previous argument would thus suggest caution whenever interpersonal preference comparisons are authorized from a theoretical perspective. However, it would be every bit as imprudent to consider that the subjects were satisfied with the rationality being presumed by the Nash equilibrium. GOERE and HOLT [2000] conducted experiments on games played only once with experienced subjects: *The most salient feature of the data is extreme sensitivity to a parameter that has no effect on the Nash outcome*. In certain cases, subjects were observed to incorporate elements into their choice which were entirely irrelevant to Nash equilibrium theory. Among these elements, some would clearly display interpersonal preference comparisons. On the whole however, the results presented by GOERE and HOLT [2000] reveal that the role of such irrelevant elements is in no way conclusive. It seems that the rationality of the Nash equilibrium is less heavily refuted both when the structure of the game is symmetrical and when irrational behaviors are relatively costly.

Strategic Misrepresentation of Preferences: to Be Fair VS to Look Fair?

As emphasized by ZAMIR, the deviation between the theoretical prediction of the Ultimatum Bargaining Game played only once and experimental results is often interpreted in terms of fairness. The corollary point is then to determine the extent to which fairness dictates, either partially or fully, subjects' behavior? In the event of behavior being heavily influenced by fairness, then the principle of rationality is not robust enough for this type of game. ZAMIR's reasoning demonstrates, and in a convincing manner, that such is not the case. Nonetheless, this demonstration can be strengthened or even completed.

The relative lack of robustness in the principle of rationality may suggest *a contrario* that fairness therefore corresponds to a robust notion. As shown by GUTH and VAN DAMME [1998], the robustness of fairness is also a debatable topic. These authors experimented on a single Ultimatum Bargaining Game with one proposer, one responder and one dummy responder. They explored, under different information conditions, what transpires when the proposer sends a message that reveals, either partially or completely, this proposal to the non-dummy responder. *The lower the information content of the message, the greedier the demand of the proposer*. (GUTH and VAN DAMME [1998], p. 227). This experiment indicates that the same proposer in an Ultimatum Bargaining Game, in its standardized form, can appear to play fair with the responder. In contrast, the introduction of a dummy player and a partial message to the non-dummy responder shows that this same proposer no longer wants to be fair and, consequently, invokes strategic considerations.

This conclusion coincides with the point of view expressed by ZAMIR, which can be summarized by the following principle: subjects, by their very nature, are not fair, yet may exhibit fair behavior in certain contexts. The next step to enable deriving the same reasoning as ZAMIR consists of ascertaining why subjects are able to play fairly in one context and unfairly in another. ZAMIR concluded that subjects are always capable of playing fairly, provided that this behavior satisfies their self-interests. In an experiment designed as a face-to-face between one principal and one team of agents (with free-riding among team members), we have shown that the principal adopts a fair behavior in the aim of inciting each team member to reduce free-riding within the team: this position thereby allows increasing the size of the pie (MEIDINGER, RULLIERE and VILLEVAL [2001]). Once the principal obtains the desired increase in the size of the pie, we observed that over the subsequent stages the principal is no longer fair: he becomes greedy to a point of keeping the biggest piece of pie all to himself. In this game, experiments have demonstrated that the principle of rationality lies just below the surface of the principal's apparent display of fair behavior.

The two experiments described above serve to reinforce ZAMIR's reasoning in terms of reputation; both sets of results were obtained in a repeated game. Repetition in fact makes it possible for the subject to build a reputation over the time period during which fair behavior constitutes a strategic approach to better defending his self-interests.

Conclusion

The entire discussion presented above calls on refusing to abandon the principle of rationality: this reasoning signifies in particular that game theory is not only normative, but also holds a predictive power. When contrasting game theory with experimental results, the issue at hand is more the modeling power than the principle of rationality: *An ideal theoretical economic paper should identify a simple frame of reasoning, not necessarily close to "reality" in the physical sense, but close to the way in which people reason about situations, Nash did precisely that!* (RUBINSTEIN [1995], p. 11).

The opinion of RUBINSTEIN, which also happens to coincide with that of ZAMIR, could erroneously lead to believe that non-cooperative game theory ultimately offers an adequate approximation of actual human behavior. However, a number of anomalies remain (some of which have been mentioned by GOERE and HOLT [2000]), in which the approximation is too crude. For this reason, experimental economics must continue to produce results, especially in those fields where the game theorist has limited modeling power: *The construction of a theory to deal with an unlimited or very large number of negotiation possibilities is as yet so difficult that it seems desirable to restrict and severely formalize the negotiation procedure to that point where a meaningful theory can be constructed. (…) because of the relatively undeveloped status of the theory, the authors feel that the use of an experimental approach is strongly indicated*. (NASH *et al.* [1954], p. 302). ■

• References

ANDREONI J. [1995]. – « Cooperation in Public-Goods Experiments: Kindness or Confusion? », *American Economic Review*, 85(4), pp. 891-904.

BOLTON G., OCKENFELS A. [2000]. – « ERC: A Theory of Equity, Reciprocity, and Competition », *American Economic Review*, 90 (1), pp. 166-93.

CHARNESS G., RABIN M. [1999]. – « Social Preferences: Some Simple Tests and a New Model », *Working Paper*, Department of Economics, Universitat Pompeu Fabra.

FEHR E., SCHMIDT K. [1999]. – « A Theory of Fairness, Competition and Cooperation », *Quarterly Journal of Economics*, august, pp.817-868.

GOERE J.K., HOLT C.A. [2000]. – « Ten Little Treasures of Game Theory and Ten Intuitive Contradictions », *Working Paper*, Department of Economics, University of Virginia.

GUTH W. [2000]. – « Experimental Game Theory », *Working Paper*, Humboldt University of Berlin.

GUTH W., VAN DAMME E. [1998]. – « Information, Strategic Behavior and Fairness in Ultimatum Bargaining: an Experimental Study », *Journal of Mathematical Psychology*, n° 42, pp. 227-47.

GUTH W., MARCHAND N., RULLIERE J.L. [1998]. – « Équilibration et dépendance du contexte : une évaluation expérimentale du jeu de négociation sous ultimatum », *Revue Économique*, vol. 49, n° 3, pp. 785-94.

GUTH W., MARCHAND N., RULLIERE J.L. [1997]. – « On the Reliability of Reciprocal Fairness: an Experimental Study », *Working Paper*, Humboldt University of Berlin.

MEIDINGER C., RULLIERE J.L., VILLEVAL M.C. [2001]. – « Fairness and Free-Riding in Principal Multi-Agent Relationships: Experimental Evidence », to appear in *Rivista Internazionale di Scienze Sociali*.

NASH J.F., KALISCH G.K., MILNOR J.W., NERING E.D. [1954]. – « Some Experimental n-Person Games », in R.M. Thrall, C.H. Coombs & R.L. Davis (eds), *Decision Processes*, Wiley, New York, pp. 301-27.

NASAR S. [1998]. – *A Beautiful Mind*, a Biography of John Forbes Nash, Touchstone, New York.

ROTH A.E., PRASNIKAR V. [1992]. – « Considerations of Fairness and Strategy: Experimental Data from Sequential Games », *Quarterly Journal of Economics*, august, pp. 865-888.

ROTH A. E., PRASNIKAR V., OKUNO-FUJIWARA M., ZAMIR S. [1991]. – « Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An Experimental Study », *American Economic Review*, 81(5), pp. 1068-95.

RUBINSTEIN A. [1991]. – « John NASH: The Master of Economic Modeling », *Scandinavian Journal of Economics*, vol. 97, pp. 9-13.

WEIBULL J.W. [2000]. – « Testing Game Theory », *Working Paper*, Research Institute of Industrial Economics, Stockholm School of Economics.