Econometrica, Vol. 53, No. 1 (January, 1985)

A NOTE ON RISK AVERSION IN A PERFECT EQUILIBRIUM MODEL OF BARGAINING

BY ALVIN E. ROTH[1]

THIS NOTE CONSIDERS the effect of risk aversion in a noncooperative model of multi-period bargaining studied by Rubinstein [**18**]. It is shown that risk aversion within each period works to a bargainer's disadvantage. Risk aversion within each bargaining period thus has the same qualitative effect on the predicted outcome of bargaining in this strategic model of multi-period bargaining as was observed in a variety of axiomatic models of single period bargaining by Roth [**11**] and Kihlstrom, Roth, and Schmeidler [**5**] who considered the models of Nash [**6**], Kalai and Smorodinsky [**4**], and Perles and Maschler [**8**]. In these axiomatic models, and in the strategic model considered here, the play of the game involves no uncertainty, and the bargainers have complete information, so the risk involved in the game consists entirely of the "strategic risk" inherent in bargaining. Thus the disadvantageous effect of risk aversion observed here constitutes a prediction about how an aversion to probabilistic risk is likely to influence a certain class of strategic interactions that do not themselves involve any probabilistic risk.

The relationship between the predictions of these axiomatic models and of the strategic model considered here will be discussed in the conclusion, as will the intuitive content of these predictions.

THE MODEL

Consider an infinite horizon game in which the players take turns making offers to divide one unit of some desirable commodity. In any period, one player has the opportunity to propose a division of the commodity, and the other player has the opportunity to accept or reject it. The game ends at any period in which an offer is accepted. After any period in which an offer is rejected, the game continues for another period, in which the opportunity to make an offer switches to the other player.

An outcome of the game can be represented by a pair (x, t) in $\{[0, 1]\times N\}\cup\{(0, \infty)\}$, where N is the set of natural numbers. The outcome (x, t) in $[0, 1]\times N$ indicates that the game ended in period t with player 1 receiving a share $s_1 = x$ and player 2 receiving a share $s_2 = 1 - x$. The outcome $(0, \infty)$ indicates that no agreement was reached.

Rubinstein [**18**] considers the perfect equilibria of the games that arise when the preferences of the players over potential outcomes of the game obey certain restrictions. Fishburn and Rubinstein [**2**] show that preferences satisfying these restrictions can be represented by von Neumann Morgenstern utility functions of the form $w_i(x, t) = v_i(s_i, t) = q_i^t u_i(s_i)$ and $w_i(0, \infty) = w_i(0, t) = 0$ where q_i is a discount factor in the interval $(0, 1)$ and u_i is a strictly increasing, continuous real-valued function of s_i. (Recall $s_1 = x$ and $s_2 = 1 - x$.) It is straightforward to verify that preferences that can be represented in this way satisfy the conditions of Rubinstein's model.

Without loss of generality, we will consider utility functions normalized so that $u_i(0) = 0$ and $u_i(1) = 1$. Rubinstein showed that an outcome (x, t) can arise from a perfect equilibrium if and only if x is a fixed point of the compound function defined by $D(x) = d_2(d_1(x))$, where the functions d_1 and d_2 are given by $d_1(x) = y$ such that $u_1(y) = q_1 u_1(x)$, and $d_2(y) = x$ such that $u_2(1 - x) = q_2 u_2(1 - y)$. That is, $d_1(x)$ is the amount y such that player 1 would be indifferent between receiving y now or waiting one period and receiving x,

[1] This work has been supported by grants from the National Science Foundation and the Office of Naval Research, and by Fellowships from the John Simon Guggenheim Memorial Foundation and the Alfred P. Sloan Foundation. The paper has also benefited from stimulating conversation and correspondence with Ariel Rubinstein, Hugo Sonnenschein, and with anonymous referees.

and $d_2(y)$ is the amount x such that player 2 would be indifferent between receiving $1-x$ now or waiting one period and receiving $1-y$.

The function D is an increasing continuous function from $[0, 1]$ to $[0, 1]$ and so it has a nonempty set of fixed points. Rubinstein further showed that the set of such fixed points x corresponding to perfect equilibria is a nonempty closed interval. In the case considered here, he observed that if the players are risk neutral, there is a unique such perfect equilibrium division x. Multiple perfect equilibrium divisions are possible, but when this occurs, arbitrarily small perturbation of the preferences will produce a unique perfect equilibrium division. Note that these results depend only on the ordinal properties of the preferences.

THE EFFECTS OF RISK AVERSION

In order to observe the role played by risk aversion in determining a perfect equilibrium partition of the commodity, we consider the effect of replacing one of the players i ($i = 1, 2$) with a player i^* having the same discount factor, but less willingness to tolerate risk in each bargaining period. Thus the utility of player i^* for an outcome (x, t) giving him a share s_{i^*} is given by the function $v_{i^*}(s_{i^*}, t) = q_i^t u_{i^*}(s_{i^*})$, where $u_{i^*}(s) = k_{i^*}(u_i(s))$ for all s in the interval $[0, 1]$, for some increasing concave function k_{i^*}. If k_{i^*} is strictly concave, i^* will be said to be strictly more risk averse than i in each period. That is, for any period t, the set of lotteries $L^*(x, t)$ over outcomes (y, t) that player i^* would prefer to any particular outcome (x, t) is contained in the set $L(x, t)$ that player i would prefer to (x, t) (cf. Yaari [**20**]). Without loss of generality, u_{i^*} can be normalized so that $u_{i^*}(0) = 0$ and $u_{i^*}(1) = 1$.

The axiomatic models mentioned earlier predict that player i would do better than player i^* against a given opponent in one-period bargaining over the division of a commodity. In order to see what prediction is made by this strategic model about the relative success in multi-period bargaining of two such players, the perfect equilibrium partitions obtained by the players must be examined.

In what follows, we shall denote by $x(1, 2)$, $x(1^*, 2)$, and $x(1, 2^*)$ perfect equilibrium partitions arising in the games played by individuals 1 and 2, 1^* and 2, and 1 and 2^*, respectively. (In all games, player 1 or 1^* makes the first offer.) For games with a unique perfect equilibrium partition, this specification is complete. For games with multiple perfect equilibrium partitions, $x(i, j)$ denotes their maximum (or equally well their minimum, both of which exist since the set of such x is a closed interval).

Recall that $x(i, j)$ is the equilibrium share of the commodity received by the player making the first offer, while the second player receives $1-x(i, j)$. Since these shares are expressed in commodities (not in utilities) it is meaningful to compare $x(1, 2)$, $x(1^*, 2)$, and $x(1, 2^*)$. The disadvantage of being more risk-averse in each period of this bargaining game can now be seen in the following result.

THEOREM: $x(1^*, 2) \leq x(1, 2) \leq x(1, 2^*)$.

Thus more risk averse bargainers get a smaller share at equilibrium[2] than do less risk averse bargainers in the same position. Furthermore, the inequalities are strict in the case of strictly concave functions k_{i^*}: strictly more risk averse bargainers do strictly less well than do less risk averse bargainers in the same position.

PROOF: Let i denote 1 or 1^*, and j denote 2 or 2^*, and recall that $x(i, j) = D_{ij}(x(i, j)) = d_i(d_j(x(i, j)))$, where $d_j(y) = x$ such that $u_j(1-x) = q_2 u_j(1-y)$, and $d_i(x) = y$ such that

[2] In games with multiple perfect equilibrium partitions the theorem states that the entire interval of equilibrium shares to a player shifts downward as he becomes more risk averse.

$u_i(y) = q_1 u_i(x)$. So $d_j(y) = 1 - u_j^{-1}(q_2 u_j(1-y))$, and $d_i(x) = u_i^{-1}(q_1 u_i(x))$, and for any x in $[0, 1]$ the function D_{ij} is given by $D_{ij}(x) = 1 - u_j^{-1}(q_2 u_j(1 - u_i^{-1}(q_1 u_i(x))))$.

Note that $D_{ij}(x)$ is monotone in the term $u_i^{-1}(q_1 u_i(x))$, so we can compare D_{12} and D_{1^*2} by comparing this term for $i = 1$ or $i = 1^*$. Specifically,

$$u_{1^*}^{-1}(q_1 u_{1^*}(x)) = u_1^{-1}(k_{1^*}^{-1}(q_1 k_{1^*}(u_1(x)))).$$

Since k_{1^*} is concave (and $k_{1^*}(0) = 0$) and $q_1 < 1$,

$$q_1 k_{1^*}(u_1(x)) \leq k_{1^*}(q_1 u_1(x)),$$

and hence

$$k_{1^*}^{-1}(q_1 k_{1^*}(u_1(x))) \leq k_{1^*}^{-1}(k_{1^*}(q_1 u_1(x))) = q_1 u_1(x),$$

since $k_{1^*}^{-1}$ is increasing. Since u_1^{-1} is also increasing

$$u_1^{-1}(k_{1^*}^{-1}(q_1 k_{1^*}(u_1(x)))) \leq u_1^{-1}(q_1 u_1(x)),$$

and so $D_{1^*2}(x) \leq D_{12}(x)$ for all x in $[0, 1]$. Thus the (maximum and minimum) fixed point of D_{1^*2} does not exceed the (corresponding) fixed point of D_{12}, so $x(1^*, 2) \leq x(1, 2)$. This proves the first part of the theorem. The proof that $x(1, 2) \leq x(1, 2^*)$ follows in the same way from the fact that $D_{12}(x) \leq D_{12^*}(x)$ for all x in $[0, 1]$.

Before going on, a subtle point concerning the nature of the result proved here should be noted. The qualitative predictions made by this strategic model about the relative bargaining success of two well-defined bargainers i and i^* is the same as the prediction made by a variety of axiomatic models. In this sense, the theorem extends this prediction from the axiomatic models to the model considered here. But because the two kinds of models are stated on quite different domains, this extension must be made carefully. In particular, it is possible to construct a third bargainer, $i^{\#}$, who is treated by the axiomatic models as if he were i^*, and by this strategic model as if he were i.

To see this, note that in the multi-period strategic model studied here, the perfect equilibrium division is determined by the ordinal preferences over outcomes (x, t), for different x and t. The different predictions for a player i^* who is more risk averse within each period than a player i, but who has the same discount factor, results in this model from the different tradeoffs made by i and i^* between the division achieved and the time at which it was achieved. That is, i and i^* may have different preferences between two distinct partitions achieved at different times. (This is why, in describing the relationshp between i and i^*, we compare their risk aversion only *within* each period.)[3] Note that if instead of looking at a utility $v_{i^*}(s, t) = q_i^t k_{i^*}(u_i(s))$ we had looked at one of the form $v_{i^{\#}} = k_{i^{\#}}(q_i^t u_i(s))$, the ordinal preferences, and hence the equilibrium predictions in this model, would be the same for both i and $i^{\#}$. However the one-period axiomatic models predict that $i^{\#}$, like i^*, does worse than i.

DISCUSSION

The theorem above demonstrates, in a strategic model of multi-period bargaining, the same qualitative effects of risk aversion that were observed in a variety of axiomatic models of one-period bargaining. In order to consider the significance of this, it will be helpful to briefly consider the nature of the two kinds of models, their advantages and disadvantages, and their place in the theory of bargaining.

Strategic models of bargaining depend on a relatively detailed description of the strategies available to the bargainers, and have the advantage that they can thus be used

[3] I have been asked to note that one of the referees remained adamantly unreconciled to comparing the risk aversion of i and i^*.

to study the effects of specific changes in the rules by which the bargaining is conducted. A further advantage is that, whatever equilibrium concept (e.g. perfect equilibrium) is employed in the analysis of a strategic model, it embodies in a concrete way fairly specific assumptions about the kind of rational behavior the bargainers are assumed to exhibit. A disadvantage of strategic models is that they often offer little help in distinguishing general phenomena that would persist under a wide variety of related rules of bargaining, from less robust phenomena that are artifacts of the specific rules defining the game.

In contrast, axiomatic models of bargaining rely on a much more schematic description of the bargaining situation and of the behavior of the bargainers. The advantage of these models is that they offer a framework in which it may be possible to identify quite general bargaining phenomena that will be robust to a variety of rule changes. The disadvantage of these models is that, precisely because they abstract away from concrete rules of bargaining and from specific assumptions about the bargainers, they offer no internal check on whether some of their results may prove to be not at all descriptive of either the behavior of actual bargainers or of well-specified "perfectly rational" bargainers.

To determine if the predictions derived from any sort of bargaining model are descriptive of actual behavior, there is of course no substitute for empirical observation.[4] But to determine if a prediction of an axiomatic model is consistent with a given specification of what constitutes rational bargaining behavior, it is sufficient to see if the prediction is fulfilled in the appropriate kinds of strategic models. This note takes a step in that direction.[5]

Note once again that both the strategic model and the axiomatic models discussed here involve no uncertainty or probabilistic risk. The influence of risk aversion on the predicted outcome of bargaining is thus entirely a prediction about situations in which all the risk is strategic, rather than probabilistic.[6] Intuitively, this strategic risk might be reflected in each bargainer's subjective probability that no agreement will be reached in the bargaining period presently underway. In the axiomatic models referred to here, failure to reach agreement in the only bargaining period results in an undesirable disagreement. In the strategic model studied here, failure to reach agreement in the present period results in continued negotiations, in which all potential payoffs are reduced by a discount factor. In both kinds of models, disagreement has costly consequences, and the desire to avoid the risk of disagreement within each period is reflected in the agreement predicted. In neither kind of model is disagreement predicted to occur with positive probability, but in both kinds of models the *possibility* of disagreement, and the cost it imposes on each bargainer, serve to determine the agreement reached.

In bargaining games involving probabilistic risk as well, the effect of risk aversion can be somewhat more complicated, as was demonstrated in Roth and Rothblum [**16**] for Nash's axiomatic model. In this regard, it may be fruitful to systematically investigate the effect of risk aversion in axiomatic and strategic models of games with incomplete information,[7] or in which the feasible agreements are themselves risky.

University of Pittsburgh

[4] In this regard see the series of experimental studies of bargaining reported in Roth and Malouf [**13**], Roth, Malouf, and Murnighan [**14**], Roth and Murnighan [**15**], Roth [**12**], and Roth and Schoumaker [**17**]. These studies have demonstrated some dimensions in which standard predictions of both axiomatic and strategic models are systematically nondescriptive of observed behavior. However these studies have also observed regularities in behavior that suggest directions in which more descriptive bargaining theories might be explored.

[5] For another relationship between Rubinstein's [**18**] model and the axiomatic model of Nash, see Binmore [**1**]. For another discussion of the influence of risk posture in a strategic model of bargaining, see Osborne [**7**], whose conclusions differ from those reached here.

[6] For a discussion of strategic and probabilistic risk in cooperative games with sidepayments, see Roth [**9**, **10**].

[7] Cf. Samuelson [**19**] which makes a start in this direction. There now exist a number of well worked out strategic models of bargaining under specific rules; see, e.g., Fudenberg and Tirole [**3**].

Manuscript received May, 1983; final revision received June, 1984.

REFERENCES

[1] BINMORE, K. G.: "Perfect Equilibria in Bargaining Models," Working Paper, London School of Economics, 1982.

[2] FISHBURN, PETER C., AND ARIEL RUBINSTEIN: "Time Preference," *International Economic Review*, 23(1982), 677-694.

[3] FUDENBERG, DREW, AND JEAN TIROLE: "Sequential Bargaining With Incomplete Information," *Review of Economic Studies*, 50(1983), 221-247.

[4] KALAI, EHUD, AND MEIR SMORODINSKY: "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43(1975), 513-518.

[5] KIHLSTROM, RICHARD, ALVIN E. ROTH, AND DAVID SCHMEIDLER: "Risk Aversion and Solutions to Nash's Bargaining Problem," in *Game Theory and Mathematical Economics*, ed. by O. Moeschlin and D. Pallaschke. Amsterdam: North Holland, 1981.

[6] NASH, JOHN F.: "The Bargaining Problem," *Econometrica*, 18(1950), 155-162.

[7] OSBORNE, MARTIN J.: "The Role of Risk Aversion in a Simple Model of Bargaining," in *Game-Theoretic Models of Bargaining*, ed. by A. E. Roth. Cambridge: Cambridge University Press, 1984.

[8] PERLES, M. A., AND M. MASCHLER: "The Super-Additive Solution for the Nash Bargaining Game," Report No. 1/80, The Institute for Advanced Studies, The Hebrew University of Jerusalem, 1980.

[9] ROTH, ALVIN E.: "The Shapley Value as a von Neumann-Morgenstern Utility Function," *Econometrica*, 45(1977), 657-664.

[10] ———: "Bargaining Ability, The Utility of Playing a Game, and Models of Coalition Formation," *Journal of Mathematical Psychology*, 16(1977), 153-160.

[11] ———: *Axiomatic Models of Bargaining*. Berlin: Springer, 1979.

[12] ———: "Toward a Theory of Bargaining: An Experimental Study in Economics," *Science*, 220(1983), 687-691.

[13] ROTH, ALVIN E., AND MICHAEL W. K. MALOUF: "Game-Theoretic Models and the Role of Information in Bargaining," *Psychological Review*, 86(1979), 574-594.

[14] ROTH, ALVIN E., MICHAEL W. K. MALOUF, AND J. KEITH MURNIGHAN: "Sociological versus Strategic Factors in Bargaining," *Journal of Economic Behavior and Organization*, 2(1981), 153-177.

[15] ROTH, ALVIN E., AND J. KEITH MURNIGHAN: "The Role of Information in Bargaining: An Experimental Study," *Econometrica*, 50(1982), 1123-1142.

[16] ROTH, ALVIN E., AND URIEL G. ROTHBLUM: "Risk Aversion and Nash's Solution for Bargaining Games with Risky Outcomes," *Econometrica*, 50(1982), 639-648.

[17] ROTH, ALVIN E., AND FRANCOISE SCHOUMAKER: "Expectations and Reputations in Bargaining: An Experimental Study," *American Economic Review*, 73(1983), 362-372.

[18] RUBINSTEIN, ARIEL: "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1982), 97-110.

[19] SAMUELSON, WILLIAM: "First Offer Bargains," *Management Science*, 26(1980), 155-164.

[20] YAARI, M. E.: "Some Remarks on Measures of Risk Aversion and On Their Uses," *Journal of Economic Theory*, 1(1969), 315-329.