THE ECONOMIC JOURNAL

SEPTEMBER 1995

The Economic Journal, **105** (*September*), 1099–1109. © Royal Economic Society 1995. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 238 Main Street, Cambridge, MA 02142, USA.

IS BAYESIAN RATIONALITY COMPATIBLE WITH STRATEGIC RATIONALITY?*

Marco Mariotti

In Bayesian game theory each player obeys the Savage axioms and there is common knowledge of this. We show that two of the less controversial axioms (ordering and dominance) are incompatible with some elementary game theoretic principles. We furthermore argue that our impossibility result is the consequence of a more fundamental conflict between the states of nature/acts framework of decision theory and the principles of strategic rationality.

The question in the title is intentionally provocative. In most current game theory Bayesian rationality and strategic rationality are thought to be inseparable from each other. A Bayesian rational agent is an agent whose choices obey Savage's (1954) axioms, or some equivalent set of axioms. Such an agent represents his uncertainty about states of the world by means of a subjective probability measure, evaluates consequences by means of a (von Neumann–Morgenstern) utility function, and chooses the act that maximises expected utility. A game theoretically rational agent is usually thought to obey the same axioms. In game theory, moreover, some form of common knowledge of rationality is assumed (more or less demanding according to the solution concept chosen), which in a Bayesian framework takes the form of common knowledge of the axioms. The most explicit statements we know of this view of game theoretic rationality are, respectively, in Myerson's (1990) influential textbook and in Harsanyi's (1977) treatise:

> 'game theory can be viewed as an extension of decision theory (to the case of two or more decision makers), or as its essential logical fulfilment'[1] (p. 5).
>
> '...our theory of rational behaviour in game situations will represent a generalization of Bayesian decision theory' (p. 47).

The position we aim to discuss in this paper is exposed by Myerson and Harsanyi in their books with unique clarity and precision. But it is fair to say

* I am grateful to Luca Anderlini, Pierpaolo Battigalli, Ken Binmore, Robin Cubitt, Peter Hammond, Martin Hollis, Hamid Sabourian, Bob Sugden, Chris Starmer and participants to seminars at the Universities of Cambridge, East Anglia, Turin, Oslo and London for comments on previous versions of the paper. I also thank Ed McLennen and Wlodeck Rabinowitz for useful and enjoyable discussions on related topics. Of course, I am alone responsible for the views expressed in the paper.

[1] It should be made clear that the decision theory Myerson refers to in this passage is expected utility theory.

that this position is at the basis of the whole theory of solution concepts as it stands now. In the recent textbook by Binmore (1992) one reads:

> '...there is no way that a book on game theory can be written if Bayesian rationality cannot be assumed' (p. 119).

As another example, Kreps and Wilson (1982) have founded the notion of sequential equilibrium (now widely used in applications to economics) on decision theoretic rationality. It is one of their claims that

> '...we have found that by making the idea of beliefs explicit, the concept of sequential equilibrium becomes consonant with the received tradition of single-person decision theory' (p. 863–4).

The title of Aumann's (1987) paper, 'Correlated Equilibrium as an Expression of Bayesian Rationality', is self-explanatory. Bernheim's (1984) justification of Bernheim's and Pierce's (1984) notion of rationalisability identifies explicitly game theoretic rationality with 'Savage rationality plus common knowledge'. And the list of examples could be continued, including Bernheim (1985), Brandenburger and Dekel (1986, 1987), Borgers (1993, 1994) Nau and McCardle (1990), Tan and Werlang (1988), who have sought in various ways to derive game theoretic principles from more primitive Bayesian decision theoretic principles. It should also be mentioned that all refinements of extensive form games are predicated in terms of maximisation with respect to subjective beliefs.

In this paper we provide an argument against the identification of the two types of rationality, and therefore in favour of a more 'classical' view of game theory. Von Neumann and Morgenstern (1944) had in fact claimed that

> 'from the point of view of player 1 who chooses a variable...the other variable can certainly not be considered as a chance event. The other variable...is dependent upon the will of the other player, which must be regarded in the same light of "rationality" as his own' (p. 99).

To be precise on the kind of problem we wish to tackle and avoid misunderstandings, we must sharply distinguish between two different questions:

(*a*) Is the Savage notion of rationality appropriate for (one-person) decision theory?

(*b*) Is the Savage notion of rationality appropriate for situations of strategic interaction?

We shall have nothing to say on question (*a*), while we suggest a negative answer to question (*b*). In order to emphasise the fact that the two questions are totally distinct, we state explicitly that we would in fact give a (qualified) positive answer to question (*a*). Our comments therefore should in no way be taken as an attempt to demote Savage's approach as a theory of decision under uncertainty.[2] Nor will our arguments bear any resemblance to the ones that

[2] Savage was indeed very careful – more careful than many of his followers – to set out precisely the limits of applicability of his theory.

have been used against subjective expected utility theory as a theory of individual decision making.

In the next section, we introduce some formal notation and we state our rationality requirements for games and decision problems. In the section that follows we present an impossibility result that shows precisely one way in which strategic and Bayesian rationality clash (since the criteria of rationality for individual decision making we propose are very basic – ordering and dominance – we will in effect demonstrate the existence of a conflict with a larger class of decision theories). In Section III, we briefly discuss the broader conceptual problems that arise in identifying the two types of rationality.[3] Section IV concludes.

I. SET-UP

We start by trying to pin down formally the alleged equivalence between one-person and strategic decision theory. It should be noted this is not a clear-cut task. What one finds in the literature is just an informal identification between the decision problem of a player in a game and the individual decision problem under uncertainty in which the set of states of nature is represented by, or includes, the strategy sets of the other players.[4] It is not specified how exactly the Savage (or an equivalent set of) axioms are to be applied in this context, where 'common knowledge' of the axioms is assumed; nor, consequently, is it specified what 'common knowledge of the axioms' means exactly. Relying on these axioms and on Savage's results, it is assumed that uncertainty on the play of the opponents is evaluated by means of a subjective probability distribution. A utility function can be constructed by using these probabilities. Finally, according to the strength of common knowledge assumed, various solution concepts may be derived: weak forms of common knowledge yield weak solution concepts such as rationalisability; strong forms of common knowledge yield strong equilibrium concepts such as Nash or correlated equilibria.

In the Savage framework, the basic objects of a decision problem for agent i are a set C_i of consequences, a set H_i of states of the world, a set of A_i of acts (a set of functions from H_i to C_i), and a binary relation $\geqslant_i$ on A_i. So, a decision problem is identified by a four-tuple $D_i = (A_i, C_i H_i \geqslant_i)$.

There is an important remark to be made at this point. Savage demands that an agent should be able to rank *all* functions from H_i to C_i, so that the set of acts is not simply the set of acts available in the decision problem at hand, but is all-inclusive. However, it is not clear what this would mean in the context of

[3] Mariotti (1995) contains further results and arguments on the topic. The excellent discussion of this paper by Battigalli (1995) provides clarifications and counterarguments, and a formalisation similar to the one we use here.

[4] The only exception to this (quite sloppy) way of proceeding we are aware of is some work in progress by Peter Hammond. In Hammond (1994) he attempted to construct the appropriate state-space for games by modelling explicitly the hyerarchies of information partitions in multiperson decision problems. The structure of this space proves to be extremely complex. The next step would be to apply the Bayesian axioms of rationality to such spaces.

a game, where only some acts (strategies) are feasible for each player in a given game and all players make their choices based on the knowledge of a given strategy set for their opponents. Presumably the correct way of proceeding should have been illustrated by the supporters of the Bayesian approach to games. Since this has not been done, in order to avoid controversies we propose to work with a restricted Savage framework, where the players in a game are asked to rank only the strategies available in that game. Note that by doing so we are making it more difficult for us to obtain an impossibility result, for we will not be able to exploit some of the more powerful Savage axioms which demand consistency across 'virtual' decision problems. In the third section we will return on this matter.

Now let $G = (S_i, c_i \geqslant_i)_{i=1,\dots,n}$ denote an n-person game in strategic form, where S_i and c_i denote player is strategy set and consequence function, respectively (the consequence function, which is *not* in terms of utility, specifies the consequence for player i of each strategy n-tuple), and $\geqslant_i$ denotes i's preference ordering on consequences. In the Bayesian approach to game theory, the decision problem of player i in game G is identified with a Savage decision problem $(A_i(G), C_i(G), H_i(G), \geqslant_i(G))$, where:

$H_i(G) = S_{-i}$, the Cartesian product of the strategy spaces of the other players;

$C_i(G) = \{x \mid x = c_i(s) \text{ for some } s \in S\}$, where S is the Cartesian product of the S_is;

$\geqslant_i(G)$ is compatible (in the sense specified below) with the ordering $\geqslant_i$ on consequences.

$a_i(.) \in A_i(G)$ if and only if there is $s_i \in S_i$ with $a_i(s_{-i}) = c_i(s_i, s_{-i})$ for all $s_{-i} \in S_{-i}$.

Notice that each strategy s_i identifies one and only one act. We may therefore identify, without loss of generality, an act in a game with a strategy, and the act set $A_i(G)$ with S_i. Also, notice that an ordering $\geqslant_i$ on acts induces a natural ordering $\geqslant_i$ on C_i, by $c_i \geqslant c_i'$, with $c_i, c_i' \in C_i$, if and only if $s \geqslant s_i'$ where $c_i(s_i, s_{-i}) = c_i$ and $c_i(s_i', s_{-i}) = c_i'$ for all $s_{-i} \in S_{-i}$. From now on we will write $\geqslant_i$ instead of $\geqslant_i(G)$, and we will omit denoting explicitly the preference ordering in a game. So, for a generic game $G = (S_i, c_i(.) \geqslant_i)_{i=1,\dots,n}$ we will just write $G = (S_i, c_i(.))_{i=1,\dots,n}$, and we will often write $\geqslant_i$ to denote an ordering $\geqslant_i(G)$ on strategies compatible with the ordering $\geqslant_i$ on consequences: no confusion will arise.

As anticipated, we do not need to use here the full force of all the Savage axioms on $\geqslant_i$. We only require two axioms which are relatively uncontroversial in decision theory, namely the ordering axiom and the dominance axiom. We use the notation $x >_i x'$ to indicate the strong preference relation.

d1: for all G, $\geqslant_i(G)$ is a complete ordering on S_i.

d2: let $G = (S_i, c_i)_{i=1,\dots,n}$. Suppose that, for some $i \in \{1, \dots n\}$, $c_i(s_i, s_{-i}) <_i c_i(s_i', s_{-1})$ for some $s_i' \in S_i$, for all $s_{-i} \in S_{-i}$. Then $s_i' >_i s_i$.

Next, we come to the game-theoretic principles of rationality. The first principle is very basic: we simply require that eliminating a strongly dominated strategy would not change the ordering on the remaining strategies.

The second principle of game-theoretic rationality says that if a game G

represents an extensive form with perfect information where a strategy, say s_i, is not part of any subgame perfect equilibrium, then s_i cannot be a dominant strategy in G. Again, this is a basic requirement.

The third principle of rationality in games is more complex. As is well known, the consequence space in Bayesian decision theory is unrestricted. In particular, it can (and must) contain consequences which are themselves lotteries. In the present context, we introduce consequences which are themselves games. This is reasonable, for in order to elicit subjective probabilities when acts are strategies in a game one needs to be able to ask the agents questions such as 'do you prefer to play game G (say, as player i) or to receive a payoff of £10 for sure?', or 'do you prefer to play game G or game G'?'[5] Suppose then that a game G' appears as a consequence for a certain strategy combination in a larger game G. Then one can represent this strategic situation as an extensive form game (actually, a multistage game), where the players first make simultaneous choices of strategy for the game G, and then make choices for the game G' at the information sets where G' is played. Now construct the strategic form associated with this extensive form and call it G''. Our game-theoretic rationality requirement is that the choices in G'' should be compatible with the choices in G, in the minimal sense that a strategy which is dominant in G should not be dominated in G''.

Finally, we state a principle concerning the preferences between games and simple consequences. We would like to express in our framework the principle that reasonable preferences should not be ruled out arbitrarily by the other principles. In particular, (the consequences attached to) 'good' Nash equilibria should not be ruled out arbitrarily. By 'good' we will mean strict and not Pareto dominated. Suppose that in a game G player i has a payoff of c_i in such a 'good' equilibrium. Then we demand that it should be possible for a player to prefer getting for sure a consequence better than c_i to playing G_i. Of course, we could impose a similar requirement replacing 'Nash equilibrium' with 'combination of (iteratively) undominated strategies'. Our impossibility result will obviously continue to hold (*a fortiori*) when this requirement, natural in a decision-theoretic setting, is imposed instead.

We now state these principles formally. A strategy $s_i \in S_i$ is strongly $\geqslant$ dominated if $c_i(s_i, s_{-i}) <_i c_i(s'_i, s_{-i})$ for some $s'_i \in S_i$, for all $s_{-i} \in S_{-i}$. A strategy $s_i \in S_i$ is weakly $\geqslant$-dominated if $c_i(s_i, s_{-i}) \leqslant_i c_i(s'_i, s_{-i})$ for some $s'_i \in S_i$, for all $s_{-i} \in S_{-i}$, and strict inequality holds for some $s_{-i} \in S_{-i}$. To avoid being pedantic, in what follows we will omit the '$\geqslant$-' prefix and simply talk of weakly or strongly dominated strategies.

g1: let $G = (S_i, c_i)_{i=1,\dots,n}$. Suppose that there is $s_i \in S_i$ which is strongly dominated. Let $G' = (S'_i, c_i(.))_{i=1,\dots,n}$ be obtained from G by removing s_i, that is $S'_i = S_i \backslash \{s_i\}$, $S'_j = S_j$ for $j \neq i$, and $c_j(.) = c'_j(.)$ for all $j \in \{1, \dots, n\}$. Then $\geqslant_j(G)$ coincides with $\geqslant_j(G')$ for all $j \neq i$ and $\geqslant_i(G)$ coincides with $\geqslant_i(G')$ on S'_i.

In order to avoid complicating the notation further with a formal definition of extensive games and subgame perfection, we give a limited version of the

[5] Battigalli (1995) describes one ingenious procedure to elicit probabilities in games.

second principle which is just sufficient for our purposes. Given a game $G = (S_i, c_i(.))$, a Nash Equilibrium is a strategy profile $s^* \in S$ with $c_i(s^*) \geqslant c_i(s'_i, s^*_{-i})$ for all $s'_i \in S_i$ with $s'_i \neq s^*_i$.

g2: let $G = (S_1, S_2, c_1(.), c_2(.))$ be a two-person game and let $S_1 = \{s_1 s'_1\}$. Suppose that $c_1(s'_1, s_2) = c_1$ and $c_2(s'_1, s_2) = c_2$ for all $s_2 \in S_2$ (that is, the consequences are (c_1, c_2) independently of what player 2 does whenever player 1 plays s'_1). Suppose further that there is $s'_2 \in S_2$ which is a unique weakly dominant strategy, and that $c_1(s_1, s'_2) > c_1(s'_1, s'_2)$. Then, $s_1 \geqslant s'_1$.

As discussed, before, this just says that for this particular game G, which represents a perfect information extensive form, a player should not prefer to play a strategy which is part of a Nash equilibrium that is not subgame perfect.

For the third principle we need some further notation. We denote the consequence 'play as player i the game $G' = (S'_i, c'_i)_{i=1,\dots,n}$' simply as G'_i. Furthermore, for a group of players $\{1,\dots,n\}$ we denote the vector $(G'_i)_{i=1,\dots,n}$ (listing the consequences for players $1,\dots,n$ of playing as players $1,\dots,n$, respectively, in game G') by G'. Given a set of players $\{1,\dots,n\}$, take two games $G = (S_i, c_i)_{i=1,\dots,n}$ and $G' = (S'_i, c'_i)_{i=1,\dots,n}$ and suppose that, for some $s^* \in S$, $c_i(s^*) = G'_i$ for all $i \in \{1,\dots,n\}$. Define the new game $GG' = (S''_i, c''_i(.))_{i=1,\dots,n}$ as follows. $S''_i = (S_i \backslash \{s^*_i\}) \cup \{s^*_i s'_i \mid s'_i \in S'_i\}$, where the meaning of $s^*_i s'_i$ will be obvious from the definition of the consequence function, which follows:

if $s_i \in S_i \backslash \{s^*_i\}$ for all i, then $c''_i(s) = c_i(s)$;

if $s_j \notin S_j \backslash \{s^*_j\}$ for some but not all j, and say (renumbering if necessary) $s_j = s^*_j s'_j$ for $j \in \{1,\dots,k\}$, $k < n$, then $c''_i(s) = c_i(s^*_1,\dots,s^*_k, s_{k+1},\dots,s_n)$;

if $s_i \notin S_i \backslash \{s^*_i\}$ for all i, and say $s_i = s^*_i s'_i$, then $c_i(s) = c'_i(s'_1,\dots,s'_n)$.

In other words, GG' represents the strategic form associated with the extensive form implicit in the game G. The strategy profile s^* played in G ensures the consequence "G' is played". The strategies $s^*_i s'_i$ mean that player i plays first the starred strategy in G that ensures playing G' if all other players play their part of s^*, and then plays s'_i in G' if G' is played. If some player does not play the starred strategy, game G' is not played and payoffs are as in G. An illustration is in Fig. 1 (numbers are amounts of money).

Game G

	L	R
T	G'	1,1
B	7,7	6,6

Game G'

	L	R
T	1,2	0,1
B	0,1	8,0

Game GG'

	LL	LR	R
TT	1,2	0,1	1,1
TB	0,1	8,0	1,1
B	7,7	7,7	6,6

Fig. 1.

The games G and GG' represent the *same* strategic situation, only in a different notation. It therefore seems natural to demand some consistency of choice in the two descriptions. A minimal requirement of consistency is that any strategy which is dominant in G should not be inferior to any other strategy in GG' (notice that it is implicit in this requirement that we limit ourselves to check consistency on strategies in $S_i \backslash \{s_i^*\}$). Formally:

g3: let $G = (S_i, c_i)_{i=1,\ldots,n}$ and $G' = (S'_i, c'_i)_{i=1,\ldots,n}$ be two games, and let $GG' = (S''_i, c''_i(.))$. Suppose that there is $i \in \{1, \ldots, n\}$ and $s_i \in S_i$, which strongly dominates any other $s'_i \in S_i$. Then, if $s''_i \in S''_i$, s''_i does not dominate s_i.

Finally, given a game G, a Nash equilibrium s^* is strict if a strict inequality holds in the definition of a Nash equilibrium. A Nash equilibrium s^* is Pareto dominated if there is $s \in S$ with $c_i(s) \geqslant_i c_i(s^*)$ for all i, with strict inequality for some i. Let us say that a preference relation $x \geqslant_i y$ between two consequences in a game G is admissible if it does not violate **d1**, **d2**, **g1**, **g2**, **g3**.

g4: let $G' = (S_i, c_i(.))$ be a consequence in a game G, and suppose that s^* is a strict and Pareto undominated Nash equilibrium of G'. Let c_i be a consequence in G with $c_i >_i c_i(s^*)$. Then the preference $c_i \geqslant G'_i$ is admissible.

For example, in Fig. 1, it seems reasonable to allow player 1 to prefer getting 7 for sure than playing the game G'.

II. AN IMPOSSIBILITY RESULT

It is easy to show that the requirements defined in the previous section are mutually incompatible. More precisely:

PROPOSITION: *there are games in which no preference ordering satisfies* **d1**, **d2**, **g1**, **g2**, **g3** and **g4**.

Proof: consider a game with two players, 1 and 2, and strategy sets $S_1 = \{A, B\}$, $S_2 = \{X, Y\}$. The consequence functions are defined by:

$$c_1(B, X) = c_1(B, Y) = c_2(B, X) = c_2(B, Y) = c_2(A, Y) = c$$

$$c_1(A, Y) = b;$$

$$c_i(A, X) = G'_i, \; i = 1, 2$$

where $$G' = (S'_i, c'_i(.))_{i=1,2}, \text{ with } S'_1 = \{A', B'\}, \; S'_2 = \{X', Y'\} \text{ and}$$

$$c_1(B', Y') = c_2(B', Y') = d;$$

$$c_1(B', X') = c_2(A', Y') = a;$$

$$c_1(A', X') = b; \; c_2(A', X') = e.$$

The preferences on consequences coincide for the two players and are:

$$a <_i < b <_i c <_i d <_i e, \text{ for } i = 1, 2.$$

(The reader may be helped in visualising this game by looking at the case with monetary consequences illustrated in Fig. 2).

Note that in G' there are two undominated and strict Nash equilibria, (A', X') and (B', Y'). Therefore, since $c >_1 b$, the preference $c >_1 G'$ should be

	X	Y
A	G'	1,2
B	2,2	2,2

Game G

	X'	Y'
A'	1,7	0,0
B'	0,0	3,3

Game G'

	XX'	XY'	Y
AA'	1,7	0,0	1,2
AB'	0,0	3,3	1,2
B	2,2	2,2	2,2

Game GG'

Fig. 2.

admissible. However, consider the game GG'. In this game, AA' is dominated by B. By **g1**, the ordering between B and AB' does not change if one removes AA'. But once this is done, it must be $AB' >_1 B$ by **g2**. Now suppose that it is the case that $c >_1 G'$. Then **d2** yields $B >_1 A$. This, together with $AB' >_1 B$, would contradict **g3**. One concludes that the preference $c >_1 G'$ is not admissible, contradicting **g4**.

What is happening in the game of the proof? Essentially, a player's preference for getting a certain consequence for sure over participating in a game, which seems perfectly reasonable (since the sure consequence is better than what the player would get in a 'good' Nash equilibrium), generates an inconsistency when such a preference is used in larger game. This is understandable: preferring something to something else assumes a completely different meaning when this preference is inserted in the context of a game, where other rational players think about, and derive inferences on behaviour from, such a preference. In game G of Fig. 2, preferring a prize of 2 to playing game G' would mean that player 1 is not confident he can get to the Nash equilibrium with payoffs (3, 3) in G'. This seems a sensible and prudent preference. The point is that such a preference cannot be held within a game, in particular in game G, because by playing A, player 1 could signal to player 2 his intention of going for the (3, 3) equilibrium. In turn, it would be in the interest of player 2 to go for (3, 3) once this intention is signalled, and this supports player 1 sending the signal.

This circular reasoning is a straightforward and standard forward induction argument. It seems that a divergence is created between the notion of preference between two consequences in abstract and the notion of preference between the same consequences within a game. But then, how can one reconcile a purely decision theoretic framework, such as Savage's, with a situation of strategic interaction?

Peter Hammond, commenting on this and related examples,[6] had suggested to me that Anscombe and Aumann's horse lotteries/roulette wheels framework might provide a better foundation for game theory than Savage's framework. This will deserve careful scrutiny, although we should point out that in our example the decision theoretic axioms required are very elementary indeed, and especially they do not include any version of the controversial 'sure-thing'

[6] See Proposition 3·1 in Mariotti (1995).

principle. It seems that any decision theoretic framework that requires an ordering on acts and excludes the choice of a dominated act would lead into trouble when applied to games. Note well: this is not to say that in a game players can choose dominated strategies! What we are claiming is that a player is not forced to hold the same preference between (complex) consequences in the two different contexts, the decision-theoretic and the game theoretic one.

III. DISCUSSION

(*a*) *Using games as consequences*. It seems to us that, granted our interpretation of what a consequence may be, the axioms we have proposed in section I should be uncontroversial, as they reflect very basic decision theoretic and game theoretic principles. In other words, anybody who objected to these principles would be objecting against standard decision or game theory, a possibly commendable task, but which has nothing to do with the problem dealt with in this paper. The only possible matter for controversy is likely to lie in treating the play of a game as a consequence in its own right. Referring to a less formal example contained in Mariotti (1995), a critic raised the objection that 'playing a game is just not sufficiently well-specified to be treated as a consequence'. We find this a rather peculiar objection. It is a characteristic of the Bayesian approach that no restriction is imposed on what consequences (and states of nature) may possibly be. The examples Savage himself uses are particularly homely, for example breaking eggs or preparing omelettes. It does not seem to us that breaking an egg is a better-specified consequence than playing a two-by-two game.[7] Moreover, our requirement **g3** can be read as defining precisely, in formal terms, what it means to have a game as a consequence in a larger game: it means playing an extensive form game constructed in the obvious way from the two original games. Finally, we observe that if it is claimed that the decision problem of a player in a game is reducible to an individual decision problem under uncertainty, the claim that a game cannot be used as a consequence is simply untenable, since in a decision problem under uncertainty decision problems can (and indeed must, in the formal derivation of expected utility theory) be used as consequences.[8]

(*b*) *Acts and strategies*. We believe the incompatibility between game theoretic and Bayesian decision theoretic principles of rationality to be more general, and deeper, than perhaps suggested by the simple formal incompatibility result of the proposition in the previous section. As we have already noted, Savage's decision theoretically reasonable requirement to order all acts, that is, all functions from states to consequences, looks if not absurd at least difficult to interpret in the context of a game. In a game, the ranking between strategies must depend on what other strategies are available, because of the possibility of iterated elimination procedures. What does it mean for a player to rank all possible strategies in a game, even those which are not available in that game?

[7] In Swift's account, it was precisely the indeterminateness of the procedures of egg-breaking that led to the war between the peoples of Lilliput and Blefuscu.

[8] The possibility of treating lotteries as consequences is a well-known reason to criticise sure-thing principles (see e.g. Hansson (1988)).

Obviously this cannot mean that the player must imagine a situation where all strategies are actually available (which, on the other hand, does not create any difficulty in a decision problem). But then, just what, and in what game, is he ranking? We must say that this completely baffles us, nor have we found an answer in the Bayesian game theory literature. For this reason our formal discussion has been conducted in the 'limited' Savage framework where only acts in a given game must be ranked. Even so, we obtained a contradiction; but the conceptual problem for the general framework remains.

(*c*) *The ordering axiom.* In the discussion after the proof of the proposition we have already emphasised that there is no reason to suppose that the ordering on consequences, and therefore on acts, is independent of the context. Here, we point to another problem. In a decision theoretic context, the requirement of ordering acts may be considered relatively uncontroversial.[9] In a game, it is not at all clear what an ordering on strategies can really mean. In a decision theoretic context, an interpretation of, say, 'the second best act' is 'the act that would be chosen if the first best was not available' (this is indeed the interpretation given by Savage himself). This interpretation is clearly not possible in a game; it is easy to construct examples where this would lead to inconsistent or absurd behaviour. Moreover, there is really no need for a player to rank all his strategies. Presumably, he will choose the one which he deems best, the 'equilibrium' one. Why should he then proceed to rank all the others? Can it be said that it would be irrational not to do so? These questions, legitimate in a decision theoretic context, are even more poignant in game theory.

IV. CONCLUDING REMARKS

We hope to have shown that there is ground to suspect a fundamental incompatibility between Bayesian decision theory and game theory. We do not deny that standard game theory has provided useful results and interesting intuitions. We maintain, however, that the Savage framework can be imposed on a game only with brute force, and in a way that leads to inconsistencies. Perhaps, treating a player *as if* he was an expected utility maximiser with respect to some subjective probability distribution, and there was common knowledge of this type of behaviour (whatever this means), is still the best possible way to describe (or prescribe) behaviour in situations of strategic interaction. But it is important to realise that proceeding thus is not axiomatically justified; or at least not in the way that it is usually asserted. This is our minimal claim in the light of our results.

A bolder claim with which we would like to conclude is that a divorce is required between game theory and individual decision theory. Too often have game theorists felt the need to seek a 'decision theoretic justification' for their solution concepts. We submit that strategic decision principles may be radically different from individual decision theoretic principles; that 'states of nature' in the Savage sense may have very little in common with the strategies decided

[9] Although it has been criticised even in that context. See e.g. Wolfowitz (1962).

upon by a rational opponent; and that, consequently, game theorists should be devoting more effort to devise a decision theoretic framework which is more suited to the strategic context, and less effort to slavishly borrow principles of action from individual decision theory.

Queen Mary and Westfield College, London

Date of receipt of final typescript: December 1994

References

Anscombe, F. J. and Aumann R. J. (1963). 'A definition of subjective probability,' *Annals of Mathematical Statistics*, vol. 34, pp. 199–205.

Aumann, R. J. (1987). 'Correlated equilibrium as an expression of Bayesian rationality.' *Econometrica*, vol. 55, pp. 1–18.

Battigalli, P. (1995). 'Discussion of Mariotti's; the foundations etc.' In *Rationality in Economics* (eds K. Arrow, E. Colombatto, M. Pearlman and C. Schmidt). London: Macmillan (forthcoming).

Bernheim, D. (1984). 'Rationalizable strategic behaviour,' *Econometrica*, vol. 52, pp. 1007–28.

Bernheim, D. (1985). 'Axiomatic characterizations of rational choice in strategic environments,' *Scandinavian Journal of Economics*, vol. 88, pp. 473–88.

Binmore, K. (1992). *Fun and Games*, Lexington, Massachussets: D. C. Heath,

Borgers, T. (1993). 'Pure strategy dominance,' *Econometrica*, vol. 61, pp. 423–30.

Borgers, T. (1994). 'Weak dominance and approximate common knowledge,' *Journal of Economic Theory*, vol. 64, pp. 265–76.

Brandenburger, A. and Dekel, E. (1986). 'On an axiomatic approach to refinements of Nash equilibrium,' Economic Theory Discussion Paper no. 104, University of Cambridge.

Brandenburger, A. and Dekel, E. (1987). 'Rationalizability and correlated equilibria,' *Econometrica*, vol. 55, pp. 1391–402.

Hammond, P. J. (1994). 'Information partitions over multiperson hierarchies of partial information,' (mimeo), presented at the 1994 ESEM meeting in Maastricht.

Hansson, B. (1988). 'Risk aversion as a problem of conjoint measurement.' In *Decision, Probability and Utility* (eds. P. Gärdenfors and N. E. Sahlin). Cambridge: Cambridge University Press.

Harsanyi, J. (1977). *Rational Players and Bargaining Equilibrium in Games and Social Situations*, Cambridge: Cambridge University Press.

Kreps, D. M. and Wilson, R. (1982). 'Sequential equilibria,' *Econometrica*, vol. 50, pp. 863–94.

Mariotti, M. (1995). 'The decision theoretic foundations of game theory,' In *Rationality in Economics*, (eds K. Arrow, E. Colombatto, M. Pearlman and C. Schmidt), London: Macmillan (forthcoming).

Myerson, R. B. (1990) *Game Theory. Analysis of Conflict*, Cambridge, Massachussets: Harvard University Press.

Nau, R. F. and McCardle, K. F. (1990). 'Coherent behaviour in noncooperative games', *Journal of Economic Theory*, vol. 50, pp. 424–44.

Pearce, D. (1984). 'Rationalizable strategic behaviour and the problem of perfection', *Econometrica*, vol. 52, pp. 1029–50.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley and Sons, New York.

Tan, J. C. C. and Werlang, S. R. C. (1988). 'The Bayesian foundations of solution concepts in games', *Journal of Economic Theory*, vol. 45, pp. 370–91.

Von Neumann, J. and Morgenstern, O. (1947). *Theory of Games and Economic Behaviour*, Princeton, New Jersey: Princeton University Press.

Wolfowitz, J. (1962). 'Bayesian inference and axioms of consistent decision', *Econometrica* vol. 30, pp. 470–9.