

Weak Dominance and Approximate Common Knowledge*

TILMAN BÖRGERS

*Department of Economics, University College London,
Gower Street, London WC1E 6BT, United Kingdom*

Received July 24, 1990; revised June 14, 1993

Considering finite normal-form games, we assume that players maximise expected utility, that they hold beliefs with full support, and that these facts are “approximate common knowledge” among players. We show that players satisfy these assumptions if and only if they choose strategies that survive the following procedure: first, all weakly dominated strategies are eliminated, and then strongly dominated strategies are iteratively eliminated. This procedure is due to E. Dekel and D. Fudenberg [*J. Econ. Theory* 52 (1990), 243–267]. We compare our justification for this procedure to theirs. *Journal of Economic Literature* Classification Number: C72. © 1994 Academic Press, Inc.

1. INTRODUCTION

In the Bayesian approach to noncooperative games (e.g., Bernheim [1] and Pearce [10]), it is assumed that players form subjective beliefs about the other players’ strategies, about their beliefs, about their beliefs about beliefs, etc., and then make a choice that maximises expected utility. Two basic results are: If no assumption is made about players’ beliefs, a strategy can be an expected utility maximising choice if and only if it is not strongly dominated by another strategy; if, on the other hand, it is common knowledge that all players maximise expected utility, then a strategy can be chosen if and only if it survives the iterated deletion of strongly dominated strategies.¹

* This paper is a revised version of parts of my paper “Bayesian Optimisation and Dominance in Normalform Games” (Basel, 1989). I would like to thank Darryl Biggar, Maarten Janssen, Larry Samuelson, Max Stinchcombe, two referees, and, in particular, Drew Fudenberg and an associate editor for their comments. Financial support from the “Schweizerischer Nationalfonds” is gratefully acknowledged.

¹ The first result is Lemma 3 in [10]. The second result is stated as Theorems 5.2 and 5.3 in [13].

The sets of strategies that are not strongly dominated, and also the sets of strategies that survive iterated deletion of strongly dominated strategies, are usually large. More restrictive concepts are obtained if “strong dominance” is replaced by “weak dominance.” Several authors have hence considered strategies that are not weakly dominated, or strategies that survive the iterated removal of weakly dominated strategies (e.g., Kohlberg and Mertens [7] and Moulin [9]).

The question arises whether these concepts also have justifications in Bayesian terms. If the elimination is not iterated, the answer is “yes.” A strategy is not weakly dominated if and only if it maximises expected utility for a subjective belief that assigns strictly positive probability to all strategy-combinations of the other players.² This may be referred to as the case in which players are “cautious” in the sense that they regard no behaviour of other players as “completely impossible.”

As regards iterated elimination of weakly dominated strategies the following difficulty arises: In analogy to the case of strong dominance one might try to derive iterated elimination of weakly dominated strategies from the assumptions that players maximise expected utility using full support beliefs, and that this fact is common knowledge. But these assumptions may be logically inconsistent. The common knowledge assumption implies that every player knows that all other players maximise expected utility using a full support belief. Hence every player assigns zero probability to choices of the other players that do not maximise expected utility for some full support belief. On the other hand, every player must have a full support belief himself, and hence give positive probability to *every* choice of the other players. If some of the other players have weakly dominated strategies these two requirements are incompatible. This difficulty seems to be insurmountable. So far no justification of iterated elimination of weakly dominated strategies in a conventional model of Bayesian optimisation and common knowledge has been given.³

Responding to these difficulties, this paper considers a procedure that is intermediate between iterated deletion of strongly dominated strategies and iterated deletion of weakly dominated strategies, and shows how this procedure can be justified in Bayesian terms. The procedure begins with one round of elimination of weakly dominated strategies and continues with iterated elimination of strongly dominated strategies. We show that this procedure can be derived from the assumption that players maximise expected utility using full support beliefs, and that this fact is *approximate* common knowledge.

² Lemma 4 in [10].

³ Stahl's [11] recent argument for iterated deletion of weakly dominated strategies uses the somewhat less conventional framework of lexicographically ordered sequences of beliefs. See our Section 5.

The notion of “approximate common knowledge” is due to Monderer and Samet [8] and Stinchcombe [12]. Roughly speaking, an event E is “approximate common knowledge” if everybody assigns high probability to E , and if everybody assigns high probability to the event that everybody assigns high probability to E , etc. An important point to note is that by assuming “approximate” common knowledge rather than “full” common knowledge we avoid the difficulty described above in the context of iterated deletion of weakly dominated strategies.

The main result of this paper is that it is approximate common knowledge that players maximize expected utility using full support beliefs if and only if they choose strategies that survive the procedure described above.

This procedure was originally proposed by Dekel and Fudenberg [5]. To motivate the procedure, Dekel and Fudenberg start from the observation that any definition of rationality in games should *at least* include all those strategies that survive iterated deletion of weakly dominated strategies. They then argue that in practice players’ payoffs are never complete common knowledge. They therefore apply iterated deletion of weakly dominated strategies not directly to the game under consideration, but instead to versions of this game in which information about payoffs is incomplete. Taking the limit for the uncertainty tending to zero, and considering all possible specifications of incomplete information about payoffs, they then obtain the iterative procedure which also we obtain.

The framework within which Dekel and Fudenberg justify their procedure is *not* a formal model of Bayesian rationality and common knowledge. In fact, their argument employs a concept—iterated deletion of weakly dominated strategies—for which, as we explained above, no Bayesian interpretation is known. Although Dekel and Fudenberg’s analysis follows a different methodology than ours, there is a close formal relationship between their paper and ours. We shall elaborate on this relationship below.

The organisation of this paper is as follows. In the next section we introduce our assumption. In Section 3, we characterise the strategies that can be chosen under this assumption. In Section 4, we discuss the formal relation between our analysis and that of Dekel and Fudenberg. Section 5 concludes with a discussion of Bernheim’s [1] closely related notion of “perfect rationalizability.” We shall use an example to illustrate how this concept differs from that of Dekel and Fudenberg. Section 5 will also refer to some related work of Brandenburger [4].

2. THE ASSUMPTION

We consider a game played by a finite number of players $i \in I = \{1, 2, \dots, I\}$ ($I \geq 2$). Every player $i \in I$ has a finite set of strategies $s_i \in S_i$. We

define $S \equiv \prod_{i \in I} S_i$ and, for $i \in I$, $S_{-i} \equiv \prod_{j \neq i} S_j$.⁴ Elements of S are denoted by s , and elements of S_{-i} are denoted by s_{-i} . Every player i has a utility function $u_i: S \rightarrow \mathbb{R}$.

We shall assume that every player chooses his strategy on the basis of a hierarchy of beliefs. A hierarchy of beliefs consists of a first order belief, i.e., a probability measure on the Cartesian product of the players' strategy sets, and, for every $n \in \mathbb{N}$ with $n \geq 2$, an n th order belief, i.e., a probability measure on the Cartesian product of the players' strategy sets and the sets of their $(n-1)$ th order beliefs.⁵ We shall not give the formal definition of hierarchies of beliefs. It is given in Tan and Werlang [13, Definition 3.9].

Within this framework we now make a sequence of assumptions. The first is:

(1) Every player chooses a strategy that maximises his expected utility, whereby expected utility is calculated using the player's first order belief. Moreover, every player's first order belief has full support on the Cartesian product of the other players' strategy sets.

Our further assumptions are defined inductively. For every $n \in \mathbb{N}$ with $n \geq 2$ we shall assume that players' n th order beliefs satisfy:

(n) Every player believes with a probability of at least p that assumptions (1) and $(n-1)$ are true.

Here $p \in (0, 1)$ is a given and fixed parameter. Below we shall focus on the case that p is close to 1. The assumptions (n) ($n \geq 2$) taken together can then be interpreted as saying that (1) is "approximate common knowledge."

The above assumptions can be formalised in the same way as the traditional assumption that players maximise expected utility, and that this fact is "full" common knowledge, is formalised. We omit the details. A formal statement of the traditional assumption is, for example, given in [13, pp. 377–378].

Our notion of "approximate common knowledge" is closely related to the concept of "common beliefs" due to Monderer and Samet [8]⁶ and to Stinchcombe's [12] notion of "approximate common knowledge." As Stinchcombe, but unlike Monderer and Samet, we do not assume that players' beliefs are derived from an objective prior. On the other hand, we

⁴ In analogous contexts in the remainder of this paper we shall use the lower index " $-i$ " without giving an explicit definition.

⁵ We permit that players are uncertain about their own choices. This is only to simplify terminology. Players' beliefs about their own choices will not play any role. The same simplification is made in Tan and Werlang [13, p. 373].

⁶ Monderer and Samet do not define their concept explicitly in terms of hierarchies of beliefs. However, their Proposition 3.(II) [11, p. 177] gives a characterisation of common beliefs in terms of such hierarchies. When comparing our assumptions with Monderer and Samet's concept, it is their Proposition 3.(II) to which we refer.

follow Monderer and Samet and require beliefs to be “close to common knowledge” at all levels of the hierarchy, whereas Stinchcombe requires them to be “close to common knowledge” only at a finite, but “large” number of levels. Finally, both in Monderer and Samet’s definition and in Stinchcombe’s definition each assumption (n) ($n \geq 2$) concerns the probability with which players believe assumption ($n-1$) to be true, whereas in our definition assumption (n) concerns the probability with which players believe assumptions (1) and ($n-1$) to be true simultaneously. Among these points, the last two are not essential for our result. The first is, of course, important.

It is useful to consider circumstances under which our assumptions will be satisfied. Suppose, for example, that for every player there is a small possibility that he makes a “mistake” and chooses a strategy at random. Our assumption is satisfied if players maximise expected utility, if the mistake probabilities are small but positive for every possible strategy, and if these two facts are common knowledge among the players.

Alternatively, imagine that there is some incomplete information about players’ payoffs. Assume every player’s beliefs about the other players’ payoffs has a “large” support, in the sense that every strategy of the other players can become strongly dominant. If it is common knowledge that players maximise expected utility, that the uncertainty has the structure just described, and that every player attaches a “large” probability to the other player’s true payoffs, then again the assumptions introduced above will hold.

3. CHARACTERISATION

To characterise strategy choices that are compatible with the assumption of Section 2 we first take $p \in (0, 1)$ to be arbitrary but fixed. Later we then focus on the case that p is close to 1.

We begin with notation: Let $i \in I$, let \tilde{S}_i be a nonempty subset of S_i , and let \tilde{S}_{-i} be a nonempty subset of S_{-i} . We denote by $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ the set of all strategies that maximise expected utility in \tilde{S}_i for some belief of player i with support S_{-i} , which assigns at least probability p to strategies in \tilde{S}_{-i} . Formally, $s_i \in \tilde{S}_i$ is in $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ if and only if there is a probability measure μ_i on S_{-i} such that $\sum_{s_{-i} \in \tilde{S}_{-i}} [u_i(s_i, s_{-i}) \mu_i(s_{-i})] \geq \sum_{s_{-i} \in S_{-i}} [u_i(\tilde{s}_i, s_{-i}) \mu_i(s_{-i})]$ for all $\tilde{s}_i \in \tilde{S}_i$, $\mu_i(s_{-i}) > 0$ for all $s_{-i} \in S_{-i}$, and $\mu_i(\tilde{S}_{-i}) \geq p$.

Assumption (1) implies that every player i chooses a strategy s_i in $D_i^p(S_i, S_{-i})$. We denote this set by $S_i^{p,1}$. Next, assumptions (1) and (2) together imply that every player i chooses a strategy s_i in $D_i^p(S_i, S_{-i}^{p,1})$. Call

this set $S_{-i}^{p,2}$. We can continue like this. In general, assumptions (1) and (n) together imply that every player i chooses a strategy s_i in $D_i^p(S_i, S_{-i}^{p,n-1})$, and we denote these sets by $S_i^{p,n}$. A necessary condition for a strategy to be a rational choice under the above assumption is thus that it is contained in $\bigcap_{n \in \mathbb{N}} S_i^{p,n}$.

Consider the sequence of sets $(S_i^{p,n})_{n \in \mathbb{N}}$. Clearly, this sequence is nested. On the other hand the sets in this sequence cannot be empty. Because the strategy sets are finite, there is hence an $m \in \mathbb{N}$, such that for $n \geq m$ the sets $S_i^{p,n}$ no longer depend on n . Thus, a necessary condition for a strategy to be a possible choice is that it is in $S_i^{p,m}$.

This condition is also sufficient. To show this we first note that by construction the sets $S_i^{p,m}$ satisfy $S_i^{p,m} = D_i^p(S_i, S_{-i}^{p,m})$. Hence a strategy is in $S_i^{p,m}$ if and only if it is a best response to a belief with support S_{-i} that assigns a probability of at least p to $S_{-i}^{p,m}$. Borrowing a term that Pearce [10, p. 1033] used in a similar context, we shall call this property of the sets $S_i^{p,m}$ the “best response property.”

This property implies that every strategy in $S_i^{p,m}$ is a best response to some first order belief with full support, and hence satisfies assumption (1) above. Since this is true for all players i , it follows from the “best response property” that every strategy in $S_i^{p,m}$ is also a best response to a first order belief of a hierarchy that satisfies assumption (2) above. Since this is true for all players i , the “best response property” implies that every strategy is also a best response to a first order belief of a hierarchy that simultaneously satisfies assumptions (2) and (3). We can continue ad infinitum. We conclude that the sets of strategies for every player that are possible choices under our assumption are precisely the sets $S_i^{p,m}$.

We continue by developing a procedure by which the sets $S_i^{p,m}$ can be determined. First, we observe that for every $n \geq 2$ the set $D_i^p(S_i, S_{-i}^{p,n-1})$ is identical to the set $D_i^p(S_i^{p,n-1}, S_{-i}^{p,n-1})$. This is because when looking for an optimal reply to full support beliefs that assign a probability of at least p to $S_{-i}^{p,n-1}$ we can restrict attention to strategies that are optimal replies to full support beliefs that assign a probability of at least p to $S_{-i}^{p,n-2}$ (since this latter set is a superset of $S_{-i}^{p,n-1}$), and hence we can restrict attention to strategies in $S_i^{p,n-1}$.

We can conclude that the sets $S_i^{p,m}$ can be determined as follows: We begin by deleting for every player all strategies that are not a best response to any full support belief of that player. We then delete for every player all strategies that are *among the remaining strategies* not a best response to a full support belief that assigns a probability of at least p to the remaining strategies of the other players. We iterate this step until no further strategies can be deleted for any player.

So far we have not made any assumptions about p . Now we focus on the case in which p is close to 1. We first give a characterisation of the sets

$D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ using dominance notions. Then we use this result to obtain a simplified description of the procedure described in the preceding paragraph.

In the following, we need to take into account that a strategy may be (strongly or weakly) dominated by a mixed strategy only, not by a pure strategy.⁷ A mixed strategy of player i is a probability measure σ_i on S_i . If σ_i is a mixed strategy of player i and if s_{-i} is a pure strategy combination for all other players then we write $u_i(\sigma_i, s_{-i})$ for player i 's expected utility.

LEMMA. *There exists some π with $0 < \pi < 1$ such that for all $i \in I$, all $p \geq \pi$, all nonempty subsets \tilde{S}_i of S_i and all nonempty subsets \tilde{S}_{-i} of S_{-i} a strategy $s_i \in \tilde{S}_i$ is contained in $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ if and only if the following two conditions are satisfied:*

(i) *There is no mixed strategy σ_i which has support in \tilde{S}_i and which weakly dominates s_i on S_{-i} , i.e., which satisfies $u_i(\sigma_i, s_{-i}) \geq u_i(s_i, s_{-i})$ for all $s_{-i} \in S_{-i}$ with strict inequality for at least one $s_{-i} \in S_{-i}$.*

(ii) *There is no mixed strategy σ_i which has support in \tilde{S}_i and which strongly dominates s_i on \tilde{S}_{-i} , i.e., which satisfies $u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i})$ for all $s_{-i} \in \tilde{S}_{-i}$.*

Proof. The Lemma claims that there exists some number π which has the asserted properties for all i , \tilde{S}_i , and \tilde{S}_{-i} . However, it suffices to prove that for any fixed i , \tilde{S}_i , and \tilde{S}_{-i} some π with the asserted properties exists, thus allowing π to depend on i , \tilde{S}_i and \tilde{S}_{-i} . The number π , the existence of which is claimed in the Lemma, can then be defined as the maximum over all i , \tilde{S}_i and \tilde{S}_{-i} of the π 's thus constructed. This maximum exists since there are only finitely many combinations of i , \tilde{S}_i , and \tilde{S}_{-i} .

So let some $i \in I$, some nonempty subset \tilde{S}_i of S_i , and some nonempty subset \tilde{S}_{-i} of S_{-i} be given and fixed. Consider the correspondence which maps every p with $0 < p < 1$ into the set $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$. Clearly this correspondence is nonempty valued. Also, as p increases the set $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ is (in terms of set-inclusion) non-increasing. Taking into account that S_i is finite, we conclude that there is a boundary (strictly between 0 and 1) such that $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ no longer depends on p if p exceeds the boundary. We claim that π can be taken to be this boundary. Letting π denote this boundary our claim is hence that, for $p \in (\pi, 1)$, a strategy s_i is contained in $D_i^p(\tilde{S}_i, \tilde{S}_{-i})$ if and only if it satisfies the two conditions of the Lemma.

We first prove the "only if-part" of this assertion. As far as condition (i) is concerned, this part is trivial. The necessity of condition (ii) can be proved as follows: Let $(p_n)_{n \in \mathbb{N}}$ be a sequence of numbers which are contained in the open interval $(\pi, 1)$, and suppose that this sequence converges

⁷ This possibility is illustrated by an example on p.6 of [6].

to 1. We have for every n that $s_i \in D_i^{p_n}(\tilde{S}_i, \tilde{S}_{-i})$. Hence for every $n \in \mathbb{N}$ there exists a probability measure on S_{-i} which assigns at least probability p_n to \tilde{S}_{-i} such that s_i maximises expected utility within \tilde{S}_i . We can take the limit of this sequence of probability measures (taking a convergent subsequence if necessary). The limit measure is a measure with support in \tilde{S}_{-i} . By continuity, s_i must also maximise expected utility in \tilde{S}_i given the limit belief. Then it obviously cannot be strongly dominated on \tilde{S}_{-i} by a mixed strategy with support in \tilde{S}_i .

We now turn to the "if-part." Suppose that a strategy s_i satisfies the two conditions of the Lemma. By Lemma 4 in [10], condition (i) implies that there exists a probability measure $\mu_{i,1}$ on S_{-i} such that s_i maximises expected utility in \tilde{S}_i given this measure. By Lemma 3 in [10], condition (ii) implies that there exists a probability-measure $\mu_{i,2}$ on S_{-i} that has support in \tilde{S}_{-i} such that s_i maximises expected utility in \tilde{S}_i given this measure. Obviously s_i also maximises expected utility in \tilde{S}_i for any linear combination of the two measures, i.e., for any measure $\mu_i^\lambda = (1 - \lambda) \cdot \mu_{i,1} + \lambda \cdot \mu_{i,2}$ with $\lambda \in (0, 1)$. All such measures have the support S_{-i} . Moreover, the probability which μ_i^λ assigns to \tilde{S}_{-i} is at least λ .

Now let some p with $\pi < p < 1$ be given. The strategy s_i maximises expected utility in \tilde{S}_i for the belief μ_i^p that one obtains if one sets $\lambda \equiv p$. This belief has full support and assigns a probability of at least p to \tilde{S}_{-i} . Hence we can conclude $s_i \in D_i^p(\tilde{S}_i, \tilde{S}_{-i})$, as required. Q.E.D.

The Lemma shows that for large enough p the procedure that we developed to identify the sets $S_i^{p,m}$ is equivalent to the procedure in which at each step all strategies are eliminated that are either weakly dominated in the original game or strongly dominated in the remaining reduced game. In each case, the dominating strategy can either be one of the remaining pure strategies of the player under consideration, or some mixed strategy that randomises over some of these pure strategies.

Considering this procedure, it is moreover obvious that in the first step the eliminated strategies are simply those that are weakly dominated, and that in all later steps the eliminated strategies are those that are strongly dominated in the remaining game. We have thus derived the procedure described in the Introduction.

4. COMPARISON WITH DEKEL AND FUDENBERG'S ARGUMENT

In the previous section, we obtained the procedure of Dekel and Fudenberg [5], using an argument different from theirs. In this section, we shall provide some insight into the question why the two approaches yield the same procedure.

As was explained in Section 1, Dekel and Fudenberg consider incomplete information versions of the given normal-form game, and solve these using iterated deletion of weakly dominated strategies. It is useful to consider a special case of this construction.

Let some finite normal-form game be given, and consider the following incomplete information version of this game. Nature selects one out of $I + 1$ possible states of the world. (I is the number of players.) If nature selects state 0 then all players' payoffs are equal to their "true" payoffs, i.e., to their payoffs in the original game. If nature selects a state i with $1 \leq i \leq I$ then only player i ' payoffs are his true payoffs. All other players are indifferent between all possible outcomes of the game. Nature selects state 0 with probability q and all other states with probability $(1 - q)/I$. These probabilities are commonly known.

Every player is informed about his own payoffs, but not about the other players' payoffs. Hence, if a player learns that his payoffs in the incomplete information game are equal to his true payoffs, he assigns probability $\tilde{q} \equiv q/(q + (1 - q)/I)$ to the event that all other players' payoffs are also equal to their true payoffs, and he assigns the remaining probability to the event that all other players are indifferent between all outcomes of the game.

A player's strategy consists of one choice for the case in which his payoffs are his true payoffs, and of another choice for the case in which he is indifferent between all outcomes of the game.

Suppose that we applied iterated deletion of weakly dominated strategies to this incomplete information game. Clearly, in the case that they are indifferent between all outcomes of the game, players can make arbitrary choices. Hence the deletion procedure will only restrict those choices that players can make if they have their true payoffs. For any player i denote by $S_i(q)$ those choices that player i can make if he chooses a strategy that survives iterated deletion of weakly dominated strategies, and if his payoffs are his true payoffs.

Dekel and Fudenberg's result immediately implies that for q sufficiently close to 1 the sets $S_i(q)$ are contained in the sets of strategies that remain if the elimination procedure of Section 3 is applied to the original normal-form game. In principle, the sets $S_i(q)$ could be smaller than Dekel and Fudenberg's sets because the above construction is only a special case of Dekel and Fudenberg's construction. However, inspection of Dekel and Fudenberg's proof shows that this is irrelevant, and that the sets $S_i(q)$ are in fact equal to Dekel and Fudenberg's solution. Roughly speaking, this is so because indifferent players can exhibit *every* conceivable behaviour.

We can now link this construction to our own result. Consider again the incomplete information game constructed above. Suppose we had carried out a fixed number of iterations of elimination of weakly dominated strategies in that game. In the next iteration a player's choice (for the case

in which he has his true payoffs) can be eliminated if and only if it is not a best response to a convex combination giving weight \tilde{q} to a belief which has full support on the choices that remain for the other players' true types and weight $1 - \tilde{q}$ to a belief which has full support on the set of all choices of the other players. Thereby all beliefs are defined on the other player's strategy sets in the original complete information game.

Equivalently, a choice can be eliminated if and only if it is not a best response to a belief that has full support on the set of all strategies of the other players, and that assigns a probability of at least \tilde{q} to the remaining strategies of the other players (for their true payoffs). Hence, if we focus on the players' true types, iterated elimination of weakly dominated strategies as applied to the incomplete information game with parameter q is just the same as the procedure developed in the preceding section for the approximate common knowledge assumption with parameter \tilde{q} . Moreover, if q tends to 1, \tilde{q} approaches 1, too.

Thus, for q close to 1, iterated deletion of weakly dominated strategies in the incomplete information game is identical both to Dekel and Fudenberg's procedure, and to ours. The two procedures are therefore identical.

5. CONCLUSION

A concept that is related to the concept studied in this paper is due to Bernheim [1]. Bernheim introduces a refinement of "rationalizability" which he calls "perfect rationalizability," defined as follows: Consider any finite normal-form game. Assume that every player has to choose each of his pure strategies with a certain strictly positive minimum probability. Assume that the minimum probabilities are common knowledge. Then apply "rationalizability" to the thus perturbed game. Strategies are "perfectly rationalizable" if they are the limit of rationalizable strategies in perturbed games as the minimum probabilities in these perturbed games converge to zero.

It is immediate that all perfectly rationalizable strategies must be contained in the sets determined by the procedure of Section 3. In some games, Bernheim's concept is, however, more restrictive than ours. There are two reasons for this. The first is that he assumes that different players' trembles are uncorrelated and that players' beliefs about other players' behaviour take the form of product measures. The second is that Bernheim assumes the probabilities of trembles to be common knowledge.

The first point matters only in games with more than two players. The second point matters also in two player games. Consider the example in Fig. 1. The procedure of Section 3 leaves for player 1 the strategies T , M_1 ,

	L	C	R
T	3,0	1,0	0,0
M_1	0,0	1,0	3,0
M_2	2,0	0,0	2,0
M_3	0,3	0,2	0,0
B	0,0	0,2	0,3

FIGURE 1

and M_2 . However, the strategy M_2 is not perfectly rationalizable in Bernheim's sense.

A formal proof of this was given in an earlier version of this paper. The structure of this argument is as follows: Player 2's optimal choice in this game depends only on the probabilities which he attaches to M_3 and B . Player 1 will choose these strategies only if he trembles. Since tremble probabilities are common knowledge, it follows that player 1 knows all important aspects of player 2's beliefs. Now note that there are no beliefs of player 2 which make both L and R (simultaneously) optimal choices. Hence, if tremble probabilities are small, player 1 can never attach large probabilities simultaneously to L and R . However, that is necessary for M_2 to be an optimal choice. Hence, M_2 is not perfectly rationalizable.⁸

Bernheim's paper, as well as this paper, argue within a traditional Bayesian framework. An alternative is a model in which agents' beliefs take the form of a sequence of lexicographically ordered probability measures. Such a model was investigated by Blume *et al.* [2].

Using Blume *et al.*'s model, Brandenburger [4] has recently provided a further justification of Dekel and Fudenberg's procedure. He assumes that the union of the supports of the players' lexicographically ordered beliefs equals the other players' strategy sets. He then introduces a common knowledge assumption. It is characteristic of Brandenburger's result that this assumption is formalised as a requirement referring only to those beliefs that have highest rank in the lexicographic ordering.⁹ It might be interesting to investigate more restrictive concepts that can be obtained if assumptions for lower rank beliefs are also made. Steps in this direction have recently been made by Stahl [11].

⁸ The example is a counterexample to Dekel and Fudenberg's [5, Footnote 4] assertion that in two player games perfect rationalizability coincides with the procedure of Section 3.

⁹ In this respect, Brandenburger's result is related to Propositions 4 and 7 in [3]. These results characterise trembling hand perfect equilibria in the lexicographic framework. The characterisations use a full support assumption, but "knowledge" assumptions are made only for beliefs of the highest rank.

REFERENCES

1. B. D. BERNHEIM, Rationalizable strategic behavior, *Econometrica* **52** (1984), 1007–1028.
2. L. BLUME, A. BRANDENBURGER, AND E. DEKEL, Lexicographic probabilities and choice under uncertainty, *Econometrica* **59** (1991), 61–79.
3. L. BLUME, A. BRANDENBURGER, AND E. DEKEL, Lexicographic probabilities and equilibrium refinements, *Econometrica* **59** (1991), 81–89.
4. A. BRANDENBURGER, Lexicographic probabilities and iterated admissibility, mimeo., Harvard Business School, 1990.
5. E. DEKEL AND D. FUDENBERG, Rational behavior with payoff uncertainty, *J. Econ. Theory* **52** (1990), 243–267.
6. D. FUDENBERG AND J. TIROLE, “Game Theory,” MIT Press, Cambridge, MA, 1991.
7. E. KOHLBERG AND J. F. MERTENS, On the stability of equilibria, *Econometrica* **54** (1986), 1003–1038.
8. D. MONDERER AND D. SAMET, Approximating common knowledge with common beliefs, *Games Econ. Behav.* **1** (1989), 170–190.
9. H. MOULIN, Dominance solvable voting schemes, *Econometrica* **47** (1979), 1337–1351.
10. D. PEARCE, Rationalizable strategic behavior and the problem of perfection, *Econometrica* **52** (1984), 1029–1050.
11. D. STAHL, “Lexicographic Rationality, Common Knowledge, and Iterated Admissibility,” Working Paper 91–10, Center for Economic Research, University of Texas, 1991.
12. M. B. STINCHCOMBE, Approximate common knowledge, mimeo, University of California at San Diego, 1988.
13. T. C. TAN AND S. WERLANG, The bayesian foundations of solution concepts of games, *J. Econ. Theory* **45** (1988), 370–391.