

EBBE GROES, HANS JØRGEN JACOBSEN,
BIRGITTE SLOTH and TORBEN TRANÆS

NASH EQUILIBRIUM WITH LOWER PROBABILITIES

ABSTRACT. We generalize the concept of Nash equilibrium in mixed strategies for strategic form games to allow for ambiguity in the players' expectations. In contrast to other contributions, we model ambiguity by means of so-called lower probability measures or belief functions, which makes it possible to distinguish between a player's *assessment* of ambiguity and his *attitude* towards ambiguity. We also generalize the concept of trembling hand perfect equilibrium. Finally, we demonstrate that for certain attitudes towards ambiguity it is possible to explain cooperation in the one-shot Prisoner's Dilemma in a way that is in accordance with some recent experimental findings.

KEY WORDS: Knightian uncertainty, ambiguity, mixed strategy Nash equilibrium, lower probabilities, belief functions, prisoner's dilemma.

1. INTRODUCTION

Empirical evidence and thought experiments such as the Ellsberg paradox (Ellsberg, 1961) have led to the understanding that decision makers do not always behave as if they were maximizing expected utility with respect to additive probability measures. Several decision-theoretic models have sought to explain the roots and consequences of expectations taking a less precise, non-additive form. However, the traditional expected utility model still prevails in applications. For example, in game theory a Nash equilibrium (in mixed strategies) takes the form of a combination of probability measures on the players' strategy sets. A Nash equilibrium can be interpreted as the players' common theory of what will be played; what is required is that the theory be consistent with each player maximizing expected utility, i.e., only optimal strategies are assigned positive probabilities. In this paper we investigate the consequences for Nash equilibrium and perfect equilibrium of allowing players to hold ambiguous expectations. We model ambiguity by means of lower probability measures, also called belief functions.

As an example, consider the game form below where the players receive monetary payments depending on their strategy choices:

		Player 2	
		<i>L</i>	<i>R</i>
Player 1	<i>U</i>	\$700,\$0	\$0,\$700
	<i>M</i>	\$0,\$700	\$700,\$0
	<i>D</i>	\$200,\$500	\$200,\$500

In all outcomes the players share \$700. Player 1 can play *D* and receive \$200 for sure, leaving \$500 for player 2; or player 1 can play *U* or *M* and thereby enter a game of the “matching pennies”-type with player 2 to determine who wins \$700. Assume that the preferences of the players depend only on the amount of money received and that more money is strictly preferred to less. Then it is straightforward to check that the game has no equilibrium in pure strategies. Hence, to ‘solve’ the game it is necessary to introduce mixed strategies. Here we interpret a mixed strategy of a player as a theory, taking the form of a probability measure, on the player in question. In particular, we assume that there is no possibility of actual randomization between pure strategies, except when explicitly included in the set of pure strategies. Thus, following Harsanyi (1973), we interpret a mixed strategy equilibrium as a common theory on the pure actions of each player, which is best-replay-consistent, i.e., it has the property that each pure strategy in the support of a player’s mixed strategy is a best reply against the theory. For a discussion of this interpretation, see also Rubinstein (1991).

To find mixed strategy equilibria it is necessary to know the players’ preferences with respect to lotteries, i.e., probability measures over the outcomes. Normally it is assumed that these are von Neumann-Morgenstern, such that they are represented by expected utility. Assume that both players have the von Neumann-Morgenstern utility function u , which fulfills: $u(\$0) = 0$, $u(\$200) = 3$, $u(\$500) = 6$, and $u(\$700) = 7$. The marginal utility of money is decreasing, so both players are risk averse, but $U(\frac{1}{2}\$0 \oplus \frac{1}{2}\$700) = \frac{1}{2}u(\$0) + \frac{1}{2}u(\$700) = 7/2 > 3 = u(\$200)$, so players prefer a lottery assign-

ing probability $\frac{1}{2}$ to each of \$0 and \$700 over getting \$200 for sure. In terms of von Neumann-Morgenstern utilities the game is:

		Player 2		
		L	R	
G_1 :	Player 1	M	0,7	7,0
		D	3,6	3,6

It has a unique Nash equilibrium, $(\frac{1}{2}U \oplus \frac{1}{2}M, \frac{1}{2}L \oplus \frac{1}{2}R)$, which is in mixed strategies. Nevertheless, it is our intuition that many people in the position of player 1, even with a risk attitude as embodied in u , would choose the strategy D . Insofar as we conceive of the players as Bayesian expected utility maximizers this is strange. Player 1 should hold a prior, a specific probability distribution expressing his expectation of the action of player 2. Such a prior has the form $\pi_2 = (\pi_L, \pi_R)$, where $\pi_L, \pi_R \geq 0$ and $\pi_L + \pi_R = 1$, and the best reply of 1 can never be D ; if $\pi_L > 3/7$, then $U(U, \pi_2) > U(D, \pi_2)$, and if $\pi_L < 4/7$, then $U(M, \pi_2) > U(D, \pi_2)$ or in other words: D is never a best reply and is thus strictly dominated by, e.g., the mixed strategy $\frac{1}{2}U \oplus \frac{1}{2}M$.

A possible explanation for the choice of D is *ambiguity* (or Knightian uncertainty) in the mind of player 1. Perhaps he has not, as Bayesian rationality requires, formed a specific prior but has a less precise expectation. If player 1 knew that player 2 were using some specific random device, he would indeed calculate the expected utility of each of his three possible strategies, and never choose D . However, player 2 is not a known random device but a human being with a thought process that is unknown to player 1.

Assume that player 1 assigns *lower probabilities* b_L to the event ‘2 plays L ’, b_R to ‘2 plays R ’, and b_{LR} to ‘2 plays either L or R ’, interpreted as the minimal probabilities with which player 1 expects the events ‘ L ’, ‘ R ’, and ‘ LR ’ to occur, where we allow $b_L + b_R < 1$, even though player 1 knows that one of the strategies must be used, i.e., $b_{LR} = 1$. Then each of player 1’s possible choices gives rise to a whole range of possible probability distributions over the strategy combinations. If, for example, player 1 plays U , this is the set of the probability distributions which assign at least probability b_L to (U, L) , at least probability b_R to (U, R) and 0 to all strategy

combinations where player 1 does not play U . Now we will need assumptions on the players' preferences when there is ambiguity. It is no longer sufficient to know the preferences over risky alternatives as expressed by expected utility, and the von Neumann-Morgenstern utility functions. One possibility is that alternatives are ranked by a utility function giving for each pure strategy the lowest possible expected utility, a security level of *expected* utility. If, for instance, the function b satisfies $b_L = b_R = 1/3$, then the utility associated with both U and M is $7/3$, while D gives 3. So this ambiguous belief of player 1, together with the security level utility function, can support the otherwise dominated pure strategy D . In the paper we consider different preference relations in ambiguous environments including, but not restricted to, the security level, or max-min, preference.

Section 2 gives an introduction to lower probability measures and discusses expected utility preferences with respect to lower probability measures over outcomes. Section 3 offers a definition of Nash equilibrium in lower probabilities, where each player plays optimally against a combination of lower probability measures on the other players. Section 4 gives a definition of (trembling hand) perfect equilibrium in lower probabilities. In Section 5 we demonstrate that cooperation can be achieved in an equilibrium in lower probabilities in one-shot Prisoner's Dilemma for certain preference relations over ambiguous outcomes.

Independently of our work, others have suggested extensions of game theoretic concepts to non-additive expectations, cf. Dow and Werlang (1994), Eichberger and Kelsey (1994), Klibanoff (1993), and Lo (1995). In contrast to our approach, all these contributions use variations of Choquet expected utility, where it is not possible to distinguish between a players' *assessment* of ambiguity (which should be equal for all players in equilibrium) and his *attitude* towards ambiguity (which should be allowed to differ among players). This is discussed further in Section 6, Finally, Section 7 concludes and offers suggestions for further work.

2. LOWER PROBABILITIES AND EXPECTED UTILITY

In this paper we consider equilibria where the players' expectations take the form of lower probability measures. For the application of lower probabilities to games certain concepts are of importance and will therefore be discussed in this section. First, a definition of the support of a lower probability measure is needed. For games with more than two players we also need a definition of the product of lower probability measures, so a player can form expectations on the opponents' strategy combinations from the expectations on each opponent. Finally, we discuss expected utility preferences with respect to lower probability measures over outcomes, and explain how one may distinguish between a player's expectation in terms of a lower probability measure and his attitude towards ambiguity.

Lower probabilities. Let X be a finite set of outcomes, and let Δ be the set of (additive) probability measures on X . Just like an ordinary probability measure, a lower probability measure, sometimes called a belief function or a totally monotone capacity, is a mapping $b : 2^X \rightarrow [0, 1]$, such that $b(\emptyset) = 0$ and $b(X) = 1$. The interpretation is that $b(E)$ is the lowest possible probability the person holding the lower probability b assigns to the event E . The upper probability of E is then $1 - b(X \setminus E)$, and b thus gives a range of possible probabilities to the event E , $[b(E), 1 - b(X \setminus E)]$. To make sense of this interpretation, some properties of b are needed. Indeed, the additivity requirement for a probability measure π , $\pi(E \cup F) = \pi(E) + \pi(F) - \pi(E \cap F)$ for all E, F , is relaxed for a lower probability measure b as follows. First it is required that b is 2-monotone; for any two events E and F , $b(E \cup F) \geq b(E) + b(F) - b(E \cap F)$. The idea is that going to larger sets can only reduce ambiguity. In the same spirit, but for k instead of 2 events, one can require:

$$(2.1) \quad b(E_1 \cup \dots \cup E_k) \geq \sum_{I \subseteq \{1, \dots, k\}} (-1)^{\#I+1} b\left(\bigcap_{i \in I} E_i\right),$$

for any k events E_1, \dots, E_k .

This property is called k -monotonicity. To understand the content of k -monotonicity note that (2.1) with equality is the usual inclusion-exclusion rule for probability measures which follows from additivity. For lower probabilities k -monotonicity must be imposed in

order to have an analog to the usual inclusion–exclusion rule. If b is k -monotone for all $k \geq 2$, i.e., if b is totally monotone, then b is a lower probability measure. Let \mathcal{B} be the set of all lower probability measures on X . Since probability measures are k -monotone for all k , we have $\Delta \subseteq \mathcal{B}$.

One can think of a lower probability measure b as the vector $(b(E))_{E \subseteq X}$. For any two lower probability measures $b_1, b_2 \in \mathcal{B}$ and $\alpha \in [0, 1]$, the convex combination $b = \alpha b_1 + (1 - \alpha)b_2$ is defined by: $b(E) = \alpha b_1(E) + (1 - \alpha)b_2(E)$, for all $E \subseteq X$. It is easy to verify that \mathcal{B} , like Δ , is a convex set.

It is natural to think of an ambiguous environment as given by a so-called mass function. A function $m : 2^X \rightarrow [0, 1]$, is a mass function if $m(\emptyset) = 0$ and $\sum_{E \subseteq X} m(E) = 1$. Note that a mass function is equivalent to a probability measure, not on X , but on 2^X . The interpretation is that for any event E , $m(E)$ is the weight of evidence in support of E which is additional to the weight already assigned to the proper subsets of E . The fact that m has non-negative values captures the idea that going to larger events can only reduce ambiguity. The belief in an event F is naturally defined as $b(F) = \sum_{E: E \subseteq F} m(E)$.

Shafer (1976) shows that if b is defined from a mass function this way, then b is indeed totally monotone, and conversely, any lower probability measure b is given by a uniquely determined mass function m_b . In particular, for $\pi \in \Delta$, $m_\pi(\{x\}) = \pi(x)$ for all $x \in X$, and $m_\pi(E) = 0$ for all E with $\#E > 1$.

Let $\mathcal{P}(X) = \{E \subseteq X \mid E \neq \emptyset\}$. For any $E \in \mathcal{P}(X)$, define the elementary lower probability measure b_E by, $b_E(F) = 1$ if $E \subseteq F$, and $b_E(F) = 0$ otherwise. This expresses that the outcome is going to be in E , but there is total ambiguity with respect to which element of E will occur. For the mass function associated with b_E we have $m_{b_E}(E) = 1$, while $m_{b_E}(F) = 0$ for all $F \neq E$.

From $b(F) = \sum_{E \subseteq F} m_b(E)$ it follows that for all $b \in \mathcal{B}$,

$$(2.2) \quad b = \sum_{E \in \mathcal{P}(X)} m_b(E) b_E,$$

so \mathcal{B} is the convex hull of the set of elementary lower probability measures.

For a lower probability measure b , define $\text{core}(b) = \{\pi \in \Delta \mid \pi(E) \geq b(E) \text{ for all } E\}$, the set of (additive) probability measures fulfilling

the lower probability requirements of b . For an additive probability measure π we have $\text{core}(\pi) = \{\pi\}$. It can be shown that for $b \in \mathcal{B}$ and all $E \in \mathcal{P}(X)$, $b(E) = \min_{\pi \in \text{core}(b)} \pi(E)$, so a lower probability measure is an exact capacity.¹

Support. We define the support of b as the minimal set within which the outcome for sure belongs according to b ,

$$\text{supp } b = \min_{b(F)=1} F.$$

Contrary to the case for probability measures, this is neither equivalent to the set $\{x \in X | b(\{x\}) > 0\}$, nor to the set $\min_{b(X \setminus F)=0} F$. The latter is used as a definition of support by Dow and Werlang (1994) and Eichberger and Kelsey (1994). One problem with this is that it may give a non-unique support and there are also interpretational problems, cf. Section 6 below. Our definition gives a unique support for each lower probability measure, which is characterized by,

LEMMA 2.1.² For any $b \in \mathcal{B}$,

$$(2.3) \quad \text{supp } b = \bigcup_{\pi \in \text{core}(b)} \text{supp } \pi = \bigcup_{E \in \text{supp } m_b} E.$$

Product measure. If $\pi = (\pi_1, \dots, \pi_n)$ is a combination of (independent) probability measures on finite sets X_1, \dots, X_n , then π induces a unique product measure $\pi_1 \otimes \dots \otimes \pi_n$ on the product $X = \times_{i=1}^n X_i$. Things are not as easy with a combination of lower probability measures, $b = (b_1, \dots, b_n)$. There is, however, a unique lower probability measure on X attaining the value $b_1(E_1) \cdot \dots \cdot b_n(E_n)$ for Cartesian sets $E_1 \times \dots \times E_n$, and fulfilling the requirement that any other lower probability measure with this property is larger, see Hendon *et al.* (1996). We will use this lower probability measure as the product. The product $b_1 \otimes \dots \otimes b_n$ will, with a slight abuse of notation, be identified with b , and is characterized by its mass function m_b . One of the results of Hendon *et al.* (1996) is that for this definition of the product b , the mass function m_b is given by the mass functions m_i of the marginal lower probability measures b_i as follows,

$$(2.4) \quad m_b(E) = \begin{cases} m_1(E_1) \cdot \dots \cdot m_n(E_n), & \text{if } E = E_1 \times \dots \times E_n \\ 0, & \text{otherwise} \end{cases}.$$

Note that if all the lower probability measures involved are additive, then the product is just the ordinary product measure.

Preference representations. Assume that an individual holds a preference relation \succsim on \mathcal{B} . We will assume that \succsim satisfies the von Neumann–Morgenstern axioms with respect to lower probabilities, implying that \succsim can be represented by an affine function U . Thus,

$$(2.5) \quad U(b) = U\left(\sum_{E \in \mathcal{P}(X)} m_b(E)b_E\right) = \sum_{E \in \mathcal{P}(X)} m_b(E)U(b_E).$$

Define $u : X \rightarrow \mathbb{R}$, by $u(x) = U(b_{\{x\}})$. For an additive probability measure π , formula (2.5) then reads $U(\pi) = \sum_{x \in X} \pi(x)u(x)$. The expected utility hypothesis for choices among risky alternatives is thus a special case of (2.5), and the function u is the von Neumann–Morgenstern utility function, and reflects the attitude towards risk as usual. The attitude towards ambiguity is reflected in the evaluation of elementary lower probability measures b_E , for $\#E > 1$. It is natural to make the restriction,

$$(2.6) \quad \min_{x \in E} u(x) \leq U(b_E) \leq \max_{x \in E} u(x).$$

Under (2.5), (2.6) is equivalent to a requirement that any lower probability measure b is just as good as some probability measure π in $\text{core}(b)$. Further investigations are performed by Jaffray (1989) and Hendon *et al.* (1994). Both get (2.5), and only differ with respect to the evaluation of elementary lower probability measures.

Jaffray (1989) obtains a representation where each elementary lower probability measure is evaluated as a convex combination of its best and worst outcome: There is $\alpha : X \times X \rightarrow [0, 1]$ such that if $\underline{x}_E \in \arg \min_{x \in E} u(x)$, and $\bar{x}_E \in \arg \max_{x \in E} u(x)$ then,

$$(2.7) \quad U(b_E) = \alpha(\underline{x}_E, \bar{x}_E)u(\underline{x}_E) + (1 - \alpha(\underline{x}_E, \bar{x}_E))u(\bar{x}_E),$$

where α satisfies the restriction that $U(b_E) \leq U(b_F)$ if $\underline{x}_E \precsim \underline{x}_F$ and $\bar{x}_E \precsim \bar{x}_F$. Inserting (2.7) into (2.5) gives utility associated to each lower probability measure b expressed in terms of von Neumann–Morgenstern utilities. The higher $\alpha(x, y)$, the more pessimistic is the evaluation of a set with x and y as extreme elements. A particular example is when $\alpha(\cdot, \cdot)$ is constant. This is the evaluation advocated by Hurwicz (1951) in the presence of ambiguity.

Hendon *et al.* (1994) obtain a representation where elementary lower probability measures are evaluated Bayes-consistently: There is a conditional probability system $\mu : X \times X \rightarrow [0, 1]$ such that,

$$(2.8) \quad U(b_E) = \sum_{x \in E} u(x)\mu(x|E),$$

which can again be inserted into (2.5). If $\mu(x|\{x, y\}) \leq \frac{1}{2}$, whenever $x \succsim y$, μ is called pessimistic. A particular case is $\mu(\cdot|X)$ being uniform. In this case the preference relation is called ambiguity neutral.

Both representations fulfill (2.5) and (2.6). Both have extreme pessimism or optimism, where b_E is exactly as good as \underline{x}_E or \bar{x}_E , respectively, as special cases. But this is almost the only intersection of the two, see Hendon (1995). In the expected utility calculations in this paper we only assume (2.5) and (2.6) unless explicitly stated otherwise.

3. NASH EQUILIBRIUM IN LOWER PROBABILITIES

Now we turn to games and will let expectations take the form of lower probability measures. It is then necessary to make assumptions on the players' preferences with respect to lower probability measures over strategy combinations. Assume for the game considered that the outcome depends on the strategic choices of the players as given by the function $f : S \rightarrow X$, from the set of strategy combinations to the set of outcomes, which could be monetary payments. From a lower probability measure b on the set of strategy combinations one can derive a lower probability measure b^f on the set of outcomes by $b^f(E) = b(f^{-1}(E))$ for all $E \subseteq X$. To assume that preferences over the b s only depend on the b^f s thus derived would be parallel to the approach normally taken in game theory, and it would ensure that it is in principle possible to elicit the players' preferences independently of the game.

In the context of lower probabilities, however, such an assumption is much stronger than in the context of (additive) probability measures. Consider the following game form, where the outcomes are monetary payments and assume that preferences only depend on the (derived lower probabilities over) amounts of money won:

		Player 2	
		<i>L</i>	<i>R</i>
	<i>U</i>	\$0,\$0	\$0,\$0
Player 1	<i>M</i>	\$0,\$7	\$0,\$0
	<i>D</i>	\$0,\$7	\$0,\$7

Assume further that player 2 is totally ambiguous about what player 1 will play, such that the lower probability on each of player 1's pure strategies as well as the lower probability on any proper subset of his strategies is zero. Player 2 has to choose between his pure strategies; strategy *L* induces the lower probability measure $b_{\{\$0,\$7\}}$ on outcomes, i.e., total ambiguity with respect to the two outcomes \$0 and \$7, and so does the pure strategy *R*. Assuming that player 2's preferences only depend on the derived lower probability measure on outcomes would imply that player 2 had to be indifferent between these two strategies. This would be a strong assumption; note that *L* weakly dominates *R*.

We do not necessarily want to impose such a strong assumption and therefore we assume that players have preferences directly on lower probabilities over the set of strategy combinations. This means that the domain of the utility function depends upon the game analyzed, such that preferences need not be neutral with respect to duplication of opponents' strategies.

Jaffray (1989) shows that the assumption that only derived lower probabilities over outcomes matter together with a monotonicity condition implies that the representing utility function has to fulfill (2.7). It follows that our analysis still applies if one wants to assume this; one just has to assume that the representing utility function associating utilities to lower probabilities over strategy combinations has this property.

We consider a finite strategic form game

$$G = (N, (S_i, U_i)),$$

with players $i \in N = \{1, \dots, n\}$, pure strategy sets S_i , and affine utility functions U_i on the set \mathcal{B} of lower probability measures on $S = \times_{i \in N} S_i$. The utility functions are assumed to fulfill (2.5) and (2.6), with S now taking the place of X .

Denote by Δ_i the set of additive probability measures on S_i , by \mathcal{B}_i the set of lower probability measures on S_i , and define $\Delta = \times_i \Delta_i$, and $\mathcal{B} = \times_i \mathcal{B}_i$. Let $\mathcal{P}^c(S) = \mathcal{P}(S_1) \times \cdots \times \mathcal{P}(S_n)$. Identifying $b = (b_1, \dots, b_n)$ with $b_1 \otimes \cdots \otimes b_n$, and using affinity of U_i , (2.5), and the property (2.4) of the product measure, we evaluate $b \in \mathcal{B}$ by

$$\begin{aligned} (3.1) \quad U_i(b) &= \sum_{E \in \mathcal{P}(S)} m_b(E) U_i(b_E) \\ &= \sum_{E_1 \times \cdots \times E_n \in \mathcal{P}^c(S)} m_{b_1}(E_1) \cdots m_{b_n}(E_n) \\ &\quad \times U_i(E_1 \times \cdots \times E_n), \end{aligned}$$

where $U_i(E_1 \times \cdots \times E_n)$ means $U_i(b_{E_1 \times \cdots \times E_n})$. So expressing a player's ambiguous preferences is equivalent to assigning utilities to the Cartesian subsets of S .

The pure best reply correspondence of player i , $PBR_i : \mathcal{B} \implies S_i$, is given by,

$$PBR_i(b) = \arg \max_{s_i \in S_i} U_i(b_{\{s_i\}}, b_{-i}).$$

Let $BR_i(b)$ denote the set of beliefs on player i consistent with some pure best reply being played,

$$\begin{aligned} BR_i(b) &= \{b_i \in \mathcal{B}_i \mid \text{supp } b_i \subseteq PBR_i(b)\} \\ &= \{b_i \in \mathcal{B}_i \mid b_i(PBR_i(b)) = 1\}. \end{aligned}$$

In this sense BR_i is defined analogously to the usual mixed strategy best reply correspondence. But $BR_i(b)$ is not equivalent to $\arg \max_{b_i \in \mathcal{B}_i} U_i(b_i, b_{-i})$. Consider the following game $(\{1, 2\}, (\{U, D\}, \{L, R\}), (U_1, U_2))$, where U_1 and U_2 are given by:

	L	R	LR
U	5,0	0,5	0,0
D	0,5	5,0	0,0
UD	0,0	0,0	0,0

Note that in order to describe the preferences fully, the table has been extended to report the $U_i(E_1 \times E_2)$ s for all Cartesian subsets of $S_1 \times S_2$. The numbers indicate that both players are extremely

ambiguity averse. Consider, for instance, the utility to player 1 of playing U when holding a lower probability b_2 on player 2, with $b_2(L) = b_2(R) = 1/3$, such that $m_2(L) = m_2(R) = m_2(LR) = 1/3$. This gives at least $(1/3) \cdot 5 + (2/3) \cdot 0 = 5/3$, and at most $(2/5) \cdot 5 + (1/3) \cdot 0 = 10/3$. The $1/3$ of probability mass, which is assigned to $\{L, R\}$ can, so to speak, be distributed between L and R . According to the table above, the utility for this individual is $(1/3) \cdot U_1(U, L) + (1/3) \cdot U_1(U, R) + (1/3) \cdot U_1(U, LR) = (1/3) \cdot 5 + (1/3) \cdot 0 + (1/3) \cdot 0 = 5/3$, so he is extremely ambiguity averse, assigning all the free probability mass between L and R to the worst outcome R .

Now consider the best reply of player 1 when he holds the theory b_2 assigning lower probability $1/2$ to both of L and R , an additive probability measure. Playing U and D both give utility $5/2$, so $PBR_1(b_2) = \{U, D\}$. Then, $BR_1(b_2)$ which is the set of theories on 1, which are consistent with 1 playing a best reply against b_2 , includes the theory $b_{\{U,D\}}$, describing total ambiguity about player 1's choice. But in itself $b_{\{U,D\}}$ does not maximize $U_1(\cdot, b_2)$, since $U_1(b_{\{U,D\}}, b_2) = 0$. Let $b = b_{\{U,D\}} \otimes b_2$. Then $m_b(\{U, D\}) \times \{L\} = m_b(\{U, D\}) \times \{R\} = 1/2$, and $m_b(\{U, D\}) \times \{L, R\} = 0$, and the conclusion follows using (3.1) and the figures of the table.

Since we only need to know PBR_i in order to calculate BR_i , it is not necessary to know how a player evaluates a Cartesian subset of S , where he does not play a pure strategy. Sufficient information for calculating best replies in the example is thus,

	L	R	LR
U	5,0	0,5	0,-
D	0,5	5,0	0,-
UD	-,0	-,0	

LEMMA 3.1. *The following statements are equivalent:*

- i. $b_i \in BR_i(b)$.
- ii. $\text{core}(b_i) \subseteq BR_i(b)$.
- iii. $b_i \geq b_{PBR_i(b)}$.

iv. If $m_{b_i}(E_i) > 0$, then $E_i \subseteq PBR_i(b)$.

Furthermore, BR_i has a closed graph.

We now generalize Nash equilibrium in the interpretation of a best reply consistent theory by allowing the theory to take the form of a combination of lower probability measures.

DEFINITION . A (Nash) equilibrium in lower probabilities (LPE) b^* , is a combination of lower probabilities where the theory on every player is consistent with choosing best replies,

$$\forall i \in N : b_i^* \in BR_i(b_i^*, b_{-i}^*).$$

Defining $BR : \mathcal{B} \implies \mathcal{B}$, by $BR(b) = BR_1(b) \times \dots \times BR_n(b)$, we obtain the simple equilibrium condition $b^* \in BR(b^*)$, i.e., b^* is a fixed point of BR .

The set of LPE may be larger than the set of mixed strategy equilibria. Consider as an example G_1 from the introduction and assume that both players are extremely ambiguity averse, as described by the extension,

	<i>L</i>	<i>R</i>	<i>LR</i>
<i>U</i>	7,0	0,7	0,-
<i>M</i>	0,7	7,0	0,-
<i>D</i>	3,6	3,6	3,-
<i>UM</i>	-,0	-,0	
<i>UD</i>	-,0	-,6	
<i>MD</i>	-,6	-,0	
<i>UMD</i>	-,0	-,0	

One LPE is the mixed equilibrium $b_1(U) = b_1(M) = 1/2$, and $b_2(L) = b_2(R) = 1/2$. But there are others. The set of LPE is,

$$\begin{aligned} & \{(b_1, b_2) \in \mathcal{B} \mid b_1(U) = b_1(M), b_1(UM) = 1, \\ & \quad \text{and } b_2(L) = b_2(R) \geq 3/7\} \cup \\ & \{(b_1, b_2) \in \mathcal{B} \mid b_1(D) = 1, \text{ and } b_2(L), b_2(R) \leq 3/7\} \cup \\ & \{(b_1, b_2) \in \mathcal{B} \mid b_1(U) = b_1(M), \text{ and } b_2(L) = b_2(R) = 3/7\}. \end{aligned}$$

In particular, any (b_1, b_2) satisfying $b_1(D) = 1$, and $b_2(L), b_2(R) \leq 3/7$ is an equilibrium, yielding utilities (3,6).

It is easily shown that,

PROPOSITION 3.2. *A mixed strategy equilibrium is an equilibrium in lower probabilities.*

An immediate corollary of this and the existence theorem for mixed strategy equilibrium is,

PROPOSITION 3.3. *There exists an equilibrium in lower probabilities.*

Given the definition of pure best replies, it is also possible to generalize the rationalizability concept of Bernheim (1984) and Pearce (1984). Define for each game the set of strategies which are rationalizable with lower probabilities by recursively removing pure strategies which are never best replies to lower probabilities over remaining strategies for the opponents. If there are only finitely many strategies for each player this recursive process will be completed in a finite number of steps and any strategy which is in the core of an *LPE* is rationalizable. Similarly, one could generalize the best response property of Pearce (1984), and define a collection of subsets of strategies to have the lower probability best response property, when all strategies in the collection are best replies to some lower probability over the collection. Then it is straightforward to generalize Pearce's (1984) characterization of rationalizable strategies as the maximal strategy subset combination that has the best response property. Allowing for lower probabilities in the definition of rationalizability implies that the set of rationalizable strategies is enlarged compared to the original definition, just as our equilibrium concept allows for more equilibria than the ordinary Nash equilibrium concept. As for the original concepts, it is possible to find games with strategies which are rationalizable with lower probabilities, but which are never played in *LPE*.

4. PERFECT EQUILIBRIUM IN LOWER PROBABILITIES

Since we have (greatly) extended the set of equilibria by allowing lower probabilities, the need for refinements has not become less. Now we define a variation of (trembling hand) perfect equilibrium. In the usual definition a Nash equilibrium is a perfect equilibrium

if it is robust against small perturbations of the game, in which players have to play each pure strategy with at least some small positive probability. When players' expectations take the form of lower probabilities we should, as a parallel, demand robustness with respect to a situation where there is so much ambiguity that no subset of the pure strategies for any player is excluded, implying that the perturbed strategy should have a completely mixed mass function.

In the sequel, let m_i be short for m_{b_i} . A perturbation is a function $\eta_i : \mathcal{P}(S_i) \rightarrow \mathbb{R}$ fulfilling $\eta_i > 0$ and $\sum_{E \in \mathcal{P}(S_i)} \eta_i(E) < 1$. Define $\mathcal{B}_i(\eta_i) = \{b_i \in \mathcal{B}_i \mid m_i(E_i) \geq \eta_i(E_i) \forall E_i \in \mathcal{P}(S_i)\}$. Let $\eta = (\eta_i)_{i \in N}$. Let $G(\eta)$ be the game where the set of possible theories on each player is constrained to $\mathcal{B}_i(\eta_i)$.

DEFINITION . *A perfect equilibrium in lower probabilities (PLPE) b^* , is an LPE such that there is a sequence (η^t) , with $\eta^t \rightarrow 0$, and a sequence (b^t) , where b^t is an LPE of $G(\eta^t)$ for all t , such that b^t converges to b^* .*

In an equilibrium b of $G(\eta)$ there is maximal belief on the set of pure best replies, i.e., for all $i \in N$,

$$b_i(PBR_i(b)) = 1 - \sum_{E_i \not\subseteq PBR_i(b)} \eta_i(E_i),$$

equivalent to,

$$(4.1) \quad m_i(E_i) = \eta_i(E_i) \text{ for } E_i \not\subseteq PBR_i(b).$$

The theory $b = (b_1, \dots, b_n) \in \mathcal{B}$ is an ϵ -perfect LPE if $m_i \gg 0$ for all $i \in N$, and

$$\text{if } E_i \not\subseteq PBR_i(b), \text{ then } m_i(E_i) \leq \epsilon.$$

We now have the following analog to a well-known theorem on perfect equilibrium.

PROPOSITION 4.1. *The following statements are equivalent:*

- i. b^* is a perfect equilibrium in lower probabilities.
- ii. There is a sequence (b^t) , where b^t is an ϵ^t -perfect LPE for all t , $\epsilon^t \rightarrow 0$, such that b^t converges to b^* .
- iii. There is (b^t) with $m_i^t \gg 0$ for all i and t , and $b^* \in BR(b^t)$ for all t , such that b^t converges to b^* .

Note that all the *LPE* of G_1 where 1 plays D , believing $b_2^*(L)$, $b_2^*(R) \leq 2/5$, are perfect. They are supported by a sequence (b^t) with $m_1^t(E_1) = (1/6) \cdot (1/t)$ for $E_1 \neq \{D\}$, and $m_1^t(D) = 1 - 1/t$ for player 1, and $m_2^t(E_2) = \max\{m_2^*(E_2), 1/t\}$.

Surprisingly it is not true that all (ordinary) perfect equilibria are *PLPE*. Consider a game where,

	L	R
U	4,0	0,4
M	0,4	4,0
D	2,2	2,2

Players are ambiguity averse, with $U_1(U, b_{LR}) = U_1(M, b_{LR}) = u' < 2$. The theory $b = (b_1, b_2)$, where $b_1(U) = b_1(M) = 1/2$, and $b_2(L) = b_2(R) = 1/2$ is a perfect equilibrium but not a *PLPE*. Consider a sequence $(b^t) = (b_1^t, b_2^t)$ converging to b , and with completely mixed mass functions. Note that b_1 is a best reply of player 1 to b_2^t only if $b_2^t(L) = b_2^t(R) = 1/2$: for both U and M to be used, they have to be equally good, which implies $b_2^t(L) = b_2^t(R) = p \leq 1/2$. And $U_1(U, b_2^t) = 4 \cdot p + 0 \cdot p + u' \cdot (1 - 2p) = 2 \cdot 2p + u' \cdot (1 - 2p)$. Since $u' < 2$, this is less than or equal to 2, with equality if and only if $p = 1/2$, i.e., if $m_2^t(LR) = 0$, contradicting that b_2^t has a completely mixed mass function.

PROPOSITION 4.2. *There exists a perfect equilibrium in lower probabilities.*

This does not follow from the existence of an ordinary perfect equilibrium, but is proved directly using Kakutani's fixpoint theorem.

We say that $s_i \in S_i$ is used in b_i , if $s_i \in \text{supp } b_i$, or equivalently, there is $E_i \subseteq S_i$ with $m_i(E_i) > 0$ and $s_i \in E_i$. From *iii* of Proposition 4.1 it is immediate that if s_i is never a best reply against a theory with completely mixed mass function then s_i is never used in *PLPE*.

Define lower probability dominance between pure strategies as follows: s_i (weakly) *LP-dominates* s'_i if, for every $b \in \mathcal{B}$, $U_i(s_i, b_{-i}) \geq U_i(s'_i, b_{-i})$ and for some $b' \in \mathcal{B}$, $U_i(s_i, b'_{-i}) > U_i(s'_i, b'_{-i})$. Just as a perfect equilibrium cannot use a dominated strategy, we have that if b is a *PLPE* then it cannot use *LP*-dominated pure strategies.

5. COOPERATION BY RATIONAL PLAYERS IN PRISONER'S DILEMMA

Consider the game $G_{PD} = (\{1, 2\}, (\{C, D\}, \{C, D\}), (U_1, U_2))$, where the preferences of 1 and 2 are given by,

	C	D	CD
C	3,3	0,4	3,-
D	4,0	1,1	1,-
CD	-,3	-,1	

These preferences are not consistent with the representation of Jaffray (1989), since this would require $U_1(b_C \times b_{CD}) \leq U_1(b_D \times b_{CD})$, but they are consistent with the representation of Hendon *et al.* (1994), and correspond to the conditional probability systems μ_1, μ_2 fulfilling,

$$\begin{aligned}
 (5.1) \quad & \mu_1((C, C) | \{(C, C), (C, D)\}) \\
 & = \mu_1((D, D) | \{(D, C), (D, D)\}) \\
 & = \mu_2((C, C) | \{(C, C), (D, C)\}) \\
 & = \mu_2((D, D) | \{(C, D), (D, D)\}) = 1
 \end{aligned}$$

These preferences are rather special. For player 1, for example the situation that the game ends in one of the boxes (C, C) or (C, D) , not at all knowing which one, is exactly as good as the situation that the game ends in (C, C) , while total ambiguity between (D, C) and (D, D) is exactly as good as (D, D) . As will be demonstrated below we do *not* need such extreme preferences for the phenomenon of cooperation in equilibrium.

Obviously G_{PD} has a unique Nash equilibrium, (D, D) . But (b_1, b_2) , with $b_1(C) = b_1(D) = b_2(C) = b_2(D) = 1/3$ is an *LPE*. Since $U_1(C, b_2) = U_1(D, b_2) = 2$, player 2 can doubt whether 1 plays C or D , and b_1 has support in $\{C, D\}$. Similarly for 2. The equilibrium has ambiguity about which outcome will occur; (C, C) occurs with probability between $1/9$ and $4/9$, and likewise for the other three outcomes. The players both evaluate the equilibrium at 2, which is just as good as, for instance, the lottery $(1/2)(C, C) \oplus (1/2)(D, D)$. We will say that an equilibrium b of G_{PD} establishes (some) cooperation if, for both players i , the utility obtained from playing a best reply against b is strictly higher than the utility from (D, D) .

For the preferences given above there are other equilibria with ambiguity. The highest possible degree of cooperation is obtained in the equilibrium (b_1, b_2) , with $b_1(C) = b_2(C) = 2/3$, and $b_1(D) = b_2(D) = 0$. Both players get utility 3. The probability of (C, C) is in $[4/9, 1]$, (C, D) and (D, C) are played with a probability in $[0, 1/3]$, and (D, D) occurs with a probability in $[0, 1/9]$.

The attitudes towards ambiguity embodied in the cpss μ_1 and μ_2 have the consequence that in case of total ambiguity, i acts *as if* he believes that j plays C if i himself plays C , and that j plays D if i plays D . Note, however, that total ambiguity, $b_1(C) = b_1(D) = b_2(C) = b_2(D) = 0$, is not an *LPE*. Against total ambiguity about 2, the unique best reply of 1 is C , and then it is not possible to assign lower probability 0 to the event of 1 playing C .

In the *LPE* considered above, $b_1(C) = b_1(D) = b_2(C) = b_2(D) = 1/3$, something similar occurs. If player 1 plays C , then the lower probabilities directly associated to the choices L and R of player 2 give him an expected utility of $(1/3) \cdot 3 + (1/3) \cdot 0$, but in addition to this, the ‘free’ probability mass of $1/3$ only associated to player 2 playing *either* L or R gives him an expected utility of $(1/3) \cdot 3$. So, when player 1 plays C , the free probability mass ‘goes to’ the good outcome (C, C) ; while if player 1 plays D , it ‘goes to’ the bad outcome (D, D) . It is important to note that the story is *not* that player 1 *believes* that if he plays C , so probably will player 2; and if he plays D , so probably will player 2. It is player 1’s attitude towards ambiguity which makes him behave *as if* this were the case.

An even stronger result is obtained in the twice repeated Prisoner’s Dilemma. Consider the following theory on each player:

- i: C in period 1. If both played C in period 1 then
 $b_i(C) = 2/3$ and $b_i(D) = 0$ in period 2. Otherwise
 $b_i(D) = 1$ in period 2.

We see that a deviation from C in period 1 can immediately gain utility 1 for the deviator but costs 2 in the second period. So this is *LPE* in all subgames. The outcome path is $((C, C), ((2/3), 0), (2/3, 0))$ with an average utility of 3. The construction of the strategies supporting this outcome is trivial. The crucial point is to get cooperation in equilibrium in the one-shot game.

Numerous explanations have been given for players choosing cooperation in a one-shot Prisoner's Dilemma without binding agreements. Perhaps the players believe that the game is repeated infinitely many times, as in the Folk Theorem. Perhaps the players believe that utilities are substantially mis-specified. Perhaps the players believe that the game is mis-specified in other ways; e.g., they believe that it is possible to commit oneself to strategies like 'cooperate if and only if the opponent cooperates'. See Binmore (1992) for a survey and discussion of such arguments.

To us it is a more attractive explanation – and one that is consistent with common knowledge of rationality – that a player, when there is ambiguity, may behave *as if* he believes that his opponent will 'think as himself' and be in a cooperative mood if and only if he is so himself. This argument was suggested by Rapoport (1966), and supported by Hofstadter (1983) and in a way by Howard (1988), too. Usually it is rejected in game theory since rationality must imply that a strictly dominated strategy is never played. If defection is preferred to cooperation when the opponent is defecting, *and* defection is preferred to cooperation when the opponent is cooperating, then defection must also be preferred to cooperation when the opponent's strategy is unknown; see the discussion in Binmore (1992). The reader may recognize a Savage-type independence axiom at work in this argument. It is exactly this axiom that is weakened in decision theories deriving non-additive subjective probabilities, and cooperation in the Prisoner's Dilemma is possible when the players' expectations are non-additive.

An experimental investigation of cooperation in the one-shot Prisoner's Dilemma by Shafir and Tversky (1992) supports the notion that cooperation may indeed be due to ambiguity. They compare the rates of cooperation in a conventional one-shot Prisoner's Dilemma to the rates of cooperation in the same game, where player 2 is informed about the move of player 1 before player 2 himself moves. They find that it is more likely for player 2 to choose *C* when he does not know the choice of player 1 than when he does, even when he is told that player 1 played *C*. Shafir and Tversky therefore attribute some of the cooperation observed "to people's reluctance to consider all the outcomes, or to their reluctance to formulate a clear preference in the presence of uncertainty about those outcomes" (1992, p.

457). This is as in our model, where the choice of C by player 2 can be explained only if there is ambiguity for player 2 about the move of player 1. If player 2 is told the move of player 1, then there is no ambiguity and a rational player 2 must play D .

Note that we have *not* shown that cooperation is possible regardless of the attitude towards ambiguity, but only that there are such attitudes — the μ_1 and μ_2 of (5.1) — that yield cooperation. As noted, μ_1 and μ_2 are very special; in particular they do *not* reflect ambiguity aversion, i.e., the property that for all x, y , $\mu_i(x|\{x, y\}) \leq 1/2$ if i prefers x over y . It *is*, however, possible to have cooperation in an *LPE*, even with ambiguity-averse preferences.³ The following proposition give exact conditions on the preferences under which cooperation in Prisoner's Dilemma is possible.

PROPOSITION 5.1. *Cooperation is possible in an LPE of a Prisoner's Dilemma, a 2×2 game G_{PD} where*

$$(5.2) \quad D_i C_j \succ_i C_i C_j \succ_i D_i D_j \succ_i C_i D_j,$$

if and only if

$$(5.3) \quad U_i(C_i, b_{S_j}) \geq U_i(D_i, b_{S_j}).$$

It is possible in a PLPE if and only if (5.3) holds with strict inequality. Furthermore, a particular cps makes cooperation possible for any G_{PD} if and only if it satisfies (5.1).

It is, of course, possible to make assumptions on preferences over lower probabilities which rule out cooperation in Prisoner's Dilemma. As a prominent example it is easy to see that the utility representation advocated by Jaffray (1989) and given in Section 2 above, can never fulfill the conditions of Proposition 5.1. In particular this shows that the max-min rule (or Choquet integration) is not compatible with cooperation.

6. RELATED LITERATURE

Generalizations of game theory that allow players to hold ambiguous beliefs have also been suggested by Dow and Werlang (1994), Klibanoff (1993), and Lo (1995). In contrast to our work, they all apply some version of Choquet expected utility (CEU). In CEU an

agent's preference in an uncertain environment is represented by a capacity and a utility function on outcomes, such that uncertain acts are ranked by their Choquet expected utility with respect to the capacity.⁴ CEU is used explicitly in the model of Dow and Werlang, which is also applied in Eichberger and Kelsey (1995), whereas Klibanoff and Lo model ambiguity as a set of conceivable probability measures, where ambiguous prospects are evaluated by the minimal conceivable expected payoff. This is similar to using CEU; see Gilboa and Schmeidler (1989).

The use of CEU is questionable, however. In CEU there is no separation of the decision maker's theory of the world and his attitude towards ambiguity; for a further discussion see Hendon *et al.* (1994) and Sarin and Wakker (1995). As such it is impossible to describe the players in terms of preferences, including attitudes towards risk and ambiguity, and then investigate which (possibly ambiguous) theories are best-reply-consistent. Consider a CEU decision maker, player 1, whose capacity on player 2 shifts from b_2 to b'_2 . It is impossible to say whether b_2 and b'_2 differ because 1 has changed his theory on 2, or because he has a new attitude towards ambiguity, or both. So if capacities do not represent pure belief the interpretation of an equilibrium in capacities is unclear. Furthermore, if players have different attitudes towards ambiguity, the case with more than two players cannot be handled. If there is ambiguity about the action taken by player 3, the capacities that player 1 and 2 hold on 3 would have to be different, if 1 and 2 had the same assessment of the ambiguity concerning player 3's move. This destroys the possibility of testing for common expectations in equilibrium, which should be possible for a generalization of Nash equilibrium. In conclusion, the CEU-approach of Dow and Werlang (1994), Klibanoff (1993) and Lo (1995) can be defended, if one argues that the capacities involved are interpreted as representing pure, but ambiguous beliefs, but then one must accept the assumption that players are always extremely ambiguity averse (pessimistic).

The papers mentioned differ in their definitions of equilibrium. Lo's (1995) model is most in line with our work, since he demands that players assign unit probability to opponents playing best replies. This is not the case in Klibanoff (1993), where players only have to consider it possible that other players play a best reply. Dow's

and Werlang's (1994) equilibrium concept only demands that the players' assessments of the ambiguity is such that it is possible that there is probability 1 of best replies. However, their definition of the support of a belief function allows for the lower probability of the set of best replies to be less than 1 in equilibrium – it can even be zero. Total ambiguity about everything, for example, is always an equilibrium according to Dow's and Werlang's definition.

Finally, it should be noted that, in contrast to our model, the models of Klibanoff (1993) and Lo (1995) allow players actually to choose mixed strategies and thereby to hedge against ambiguity. It would be possible for player 1 in the game G_1 of Section 1 to play U with probability $1/2$ and M with probability $1/2$, and thus obtain a non-ambiguous expected payoff of $7/2$. This is not in accordance with the interpretation of mixed strategy equilibria suggested here, and in fact we find it hard to justify. When a player has taken a random draw and is about to move, there will still be ambiguity. Consider a player 1 who wants to play $(1/2)U \oplus (1/2)M$ as suggested above, and starts out by tossing a fair coin, finding that U should be played. If there is total ambiguity about the move of player 2 and player 1 is ambiguity averse, then player 1 prefers D over U , but following the mixed strategy demands that player 1 should choose U rather than D . In other words, when the pure strategies U and M are not themselves best replies it is not (*ex post*) credible for player 1 to adopt strategy $(1/2)U \oplus (1/2)M$. If it were possible to commit to some mixed strategies, e.g., by delegating the random draw and move to an agent, then these mixed strategies should be included in the pure strategy set.

7. CONCLUDING REMARKS

In this paper we have formulated a concept of Nash equilibrium in which the players are allowed to be ambiguous about the strategic choices of the other players. The ambiguity has been modeled using lower probability measures. Exact capacities, or even just sets of conceivable probability measures, would perhaps be preferable as models of ambiguity. But then one would need a model of players' preferences with respect to such ambiguous objects. At the moment this just does not seem tractable.⁵ In particular it is hard to use sets of

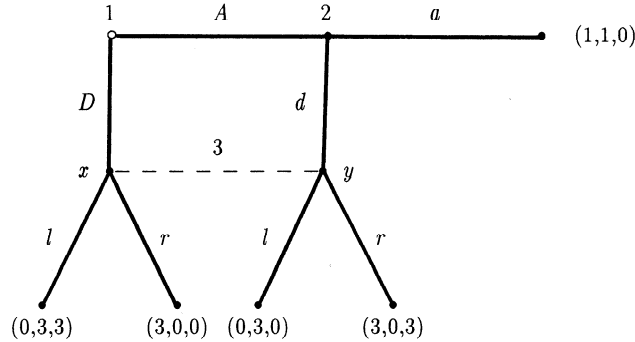


Fig. 1. An extensive form game with three players.

probability measures as primitives in a von Neumann–Morgenstern-like decision theory, since they do not lend themselves to applications of the mixture set theorem.

We have applied the solution concept to the Prisoner’s Dilemma and given an explanation of cooperation that hinges on ambiguity and the players’ attitudes towards ambiguity.

For applications it is of interest to note that the set of *LPE* of a game is equivalent to the set of Nash equilibria in a game where the players’ strategy sets are the subsets of the original strategy sets, such that a mixed strategy corresponds to a mass function, see Appendix B. By this it is possible to calculate the set of *LPE* for a particular game using standard algorithms. A similar result applies for *PLPE*.

Other interesting applications may be examples where players have cautious strategies at their disposal. Consider the extensive form game of Figure 1, where the players are assumed to be extremely ambiguity averse.

In this game player 1 goes *D* if he believes that player 3 goes *R* with probability above $1/3$. And player 2 goes *d* if he believes that player 3 goes *l* with probability above $1/3$. So, if the common theory on 3 is a probability measure, i.e., $b_l + b_r = 1$, at least one of 1 and 2 must go down. Allowing the theory on 3 to be a lower probability measure and assuming players 1 and 2 to be ambiguity averse, we can get (A, a) as an equilibrium outcome. One such equilibrium is $(A, a, b_{\{l,r\}})$, which even can be shown to be perfect, in the (agent) strategic form of the game.

Since this example is robust it shows the existence of potential applications of *LPE* to extensive form games. Furthermore, the example captures a point raised in more traditional game theory. It should be possible to have (A, a) in equilibrium since, if (A, a) is played, player 3's information set is off path, and it should then be allowed that players 1 and 2 have *different* theories on player 3, which may explain both A and a . This is captured by the concept of a self-confining equilibrium, Fudenberg and Levine (1993). We see here that it is *not* required that players 1 and 2 hold different beliefs on 3 for (A, a) to be played in equilibrium, all that is needed is that they hold sufficiently imprecise beliefs and are sufficiently ambiguity averse, cf. Greenberg (1995) for a similar point.

A problem with *LPE* in extensive form games is how to refine it. One could of course apply *PLPE* to the agent strategic form of a game, but it is not clear that this is reasonable. One problem is the definition of a Bayes' rule for lower probability measures, which is also the reason why there is no straightforward way of generally defining concepts of sequential or perfect Bayesian equilibrium in lower probabilities.

APPENDIX A: PROOFS

Proof of Lemma 2.1. Assume that $x \in \text{supp } b$. If $x \notin \bigcup_{\pi \in \text{core}(b)} \text{supp } \pi$ then $\pi(X \setminus \{x\}) = 1$ for all $\pi \in \text{core}(b)$, i.e., $b(X \setminus \{x\}) = \min_{\pi \in \text{core}(b)} \pi(X \setminus \{x\}) = 1$, contradicting that $x \in \text{supp } b$.

Assume that there is x such that $x \in \bigcup_{\pi \in \text{core}(b)} \text{supp } \pi$, and $x \notin \bigcup_{E \in \text{supp } m_b} E$. The correspondence $\text{core}(\cdot) : \mathcal{B} \implies \Delta$ is easily shown to be affine, so $\text{core}(b) = \sum_{E \in \text{supp } m_b} m_b(E) \cdot \text{core}(b_E)$. If, for some set E , $x \notin E$, then $\pi(x) = 0$ for all $\pi \in \text{core}(b_E)$, and since $x \notin \bigcup_{E \subseteq \text{supp } m_b} E$, $\pi(x) = 0$ for all $\pi \in \text{core}(b)$, contradicting $x \in \bigcup_{\pi \in \text{core}(b)} \text{supp } \pi$.

Assume that $x \in \bigcup_{E \in \text{supp } m_b} E$, i.e., there is E_x with $x \in E_x$ and $m_b(E_x) > 0$. Then $b(X \setminus \{x\}) = \sum_{E: x \notin E} m_b(E) < 1$, i.e., if $b(F) = 1$ then $x \in F$. ■

Proof of Lemma 3.1. $i \implies ii$: If $\pi_i \in \text{core}(b_i)$ then by Lemma 2.1, $\text{supp } \pi_i \subseteq \text{supp } b_i \subseteq \text{PBR}_i(b)$, i.e., $\pi_i \in \text{BR}_i(b)$.

$ii \implies iii$: Since $b_i \geq 0$ we only have to show that $b_i(E) = 1$ if $PBR_i(b) \subseteq E$. But $b_i(E) = \min_{\text{core}(b_i)} \pi_i(E) = 1$ since $\pi_i(E) = 1$ for all $\pi_i \in \text{core}(b)$ by ii .

$iii \implies i$: If $b_i \geq b_{PBR_i(b)}$ then $\text{supp } b_i \subseteq \text{supp } b_{PBR_i(b)} = PBR_i(b)$.

$i \iff iv$ is immediate by Lemma 2.1.

Let (b^t, b_i^t) be a sequence in $\mathcal{B} \times \mathcal{B}_i$ converging to (b, b_i) with $b_i^t \in BR_i(b^t)$ for all t . We have to show that if $m_i(E_i) > 0$ then $E_i \subseteq PBR_i(b)$. By convergence of m_i^t to m_i we must have $m_i^t(E_i) > 0$ infinitely often, i.e., $E_i \subseteq PBR_i(b^t)$ infinitely often, and by continuity of U_i , $E_i \subseteq PBR_i(b)$. \blacksquare

Proof of Proposition 4.1. $i \implies ii$: Let (b^t) be as in ii . Define $\epsilon(\eta) = \max_{i \in N} \max_{E_i \subseteq S_i} \eta_i(E_i)$. Then b^t is an ϵ -perfect LPE. Since $\epsilon(\eta^t) \rightarrow 0$, b^* fulfills ii .

$ii \implies iii$: Let (b^t) be as in ii . Consider E_i with $m_i^*(E_i) > 0$. We will show that from some $T_{E_i} \in \mathbb{N}$, $E_i \subseteq PBR_i(b^t)$. Assume not. Then, infinitely often, $m_i^t(E_i) \leq \epsilon^t$, and since $\epsilon^t \rightarrow 0$ and m^t is convergent, $\lim_{t \rightarrow \infty} m_i^t(E_i) = 0$, contradicting $m_i^*(E_i) > 0$. So, for $t \geq \max_{i \in N} \max_{E_i \subseteq S_i} T_{E_i}$, $b^* \in BR_i(b^t)$.

$iii \implies i$: Let (b^t) be as in iii . Let (ϵ^t) be a sequence with $\epsilon^t > 0$ and $\epsilon^t \rightarrow 0$. Define $\eta^t : \mathcal{P}(S_i) \rightarrow \mathbb{R}$ by, for all $i \in N$ and $E_i \subseteq S_i$,

$$\eta_i^t(E_i) = \begin{cases} m^t(E_i), & \text{if } m^*(E_i) = 0 \\ \epsilon^t, & \text{otherwise.} \end{cases}$$

Then $\eta_i^t \rightarrow 0$, and for all t large enough, $\sum_{E_i} \eta_i^t(E_i) < 1$. Further, for t large enough, $b^t \in \mathcal{B}(\eta^t)$ and by (4.1), b^t is an LPE of $G(\eta^t)$. \blacksquare

Proof of Proposition 4.2. Consider a perturbation η^t . Define $\mathcal{B}(\eta^t) = \times_i \mathcal{B}_i(\eta_i^t)$. The correspondence $BR : \mathcal{B}(\eta^t) \implies \mathcal{B}(\eta^t)$ is upper hemi-continuous (u.h.c.): We only need to check that BR_i is u.h.c. Let (b^s, b_i^s) be a sequence in $\mathcal{B}(\eta^t) \times \mathcal{B}_i(\eta^t)$ converging to (b, b_i) with $b_i^s \in BR_i(b^s)$ for all s . We have to show that $b_i \in BR_i(b)$, i.e., by (4.1) that if $m_i(E_i) > \eta_i(E_i)$ then $E_i \subseteq PBR_i(b)$. By convergence of m_i^s to m_i we must have $E_i \subseteq PBR_i(b^s)$ infinitely often, and by continuity of U_i , $E_i \subseteq PBR_i(b)$.

Since $\mathcal{B}(\eta^t)$ is compact and convex, there is a fixed point of BR by Kakutani's fixed point theorem, i.e., there is an equilibrium

$b^t \in LPE(G(\eta^t))$. By compactness of \mathcal{B} , (b^t) has a convergent subsequence. ■

Proof of Proposition 5.1. ‘Only if’. Assume that some player i strictly prefers D over total ambiguity. Then C is never a best reply, and the only LPE is (D, D) .

‘If’. When (5.3) holds, $U_1(C, (0, 0)) \geq U_1(D, (0, 0))$ and $U_1(C, (0, 1)) < U_1(D, (0, 1))$, so by affinity there is $y \in [0, 1[$ such that $U_1(C, (0, y)) = U_1(D, (0, y))$, with $y = 0$ if and only if $U_1(C, (0, 0)) = U_1(D, (0, 0))$. Similarly there is an x such that $U_2((0, x), C) = U_2((0, x), D)$. So $((0, x), (0, y))$ is an LPE .

For the characterization of $PLPE$, note that C is weakly dominated by D in G_{PD} if (5.3) holds with equality for a player. Consequently C is not used in any $PLPE$ of G_{PD} . If (5.3) holds with strict inequality, $((0, x), (0, y))$ is shown to be a $PLPE$ by construction of a sequence as in Proposition 4.1.iii.

For any u_1, u_2 satisfying (5.2) there is a continuum of cps 's such that (5.3) holds. In particular, if μ_1 and μ_2 satisfy (5.1) then (5.3) is satisfied with inequality for any u_1, u_2 satisfying (5.2). On the other hand, if (5.1) is not satisfied then there are u_1, u_2 satisfying (5.2) such that (5.3) is not satisfied. Assume, for example that $\mu_1((C, C) | \{(C, C), (C, D)\}) = p < 1$. Then consider u_1 such that $u_1(D, C) > u_1(C, C) = 1$, $u_1(D, D) = 0$, and $u_1(C, D) < -p(1 - p)$. Obviously (5.1) is satisfied, but $U_1(D, b_{S_2}) \geq u_1(D, D) = 0$, and $U_1(C, b_{S_2}) = p \cdot u_1(C, C) + (1 - p) \cdot u_1(C, D) < p - p = 0$, violating (5.3). Similarly, for any other violation of (5.1) there are examples where (5.3) is violated. ■

APPENDIX B: DEFINITIONS USING THE MIXED SUBSET EXTENSION

In order to investigate the relation to ordinary (mixed strategy) equilibrium, we seek a transformed game in which the mixed strategy equilibria correspond to the equilibria in lower probabilities of the original game.

Define the mixed subset extension \mathcal{G} of G as the game

$$\mathcal{G} = \left(N, \left(\prod_i, U_i \right) \right),$$

where $\prod_i = \Delta(\mathcal{S}_i)$, $\mathcal{S}_i = \mathcal{P}(S_i)$, and U'_i is to be defined later. So a pure strategy E_i in \mathcal{G} corresponds to a non-empty subset of the pure strategies in G . Consequently, a mixed strategy π_i in \mathcal{G} is a mass function on S_i and thus corresponds to a belief function on S_i . Now define the utility functions by,

$$U'_i(\pi_1, \dots, \pi_n) = \sum_{E_i \in \mathcal{S}_i} \pi(E_i) \cdot U'_i(E_i, \pi_{-i}),$$

with,

$$(A.1) \quad U'_i(E_i, \pi_{-i}) = \min_{s_i \in E_i} U_i(s_i, b_{\pi_{-i}}),$$

where b_π is the belief function corresponding to the mass function π . Notice that \mathcal{G} is not an extension of the game $(N, (\mathcal{S}_i, U_i))$, i.e., we do not necessarily have $U'_i(E) = U_i(b_E)$. But, for any $s_i \in S_i$, we do have $U'_i(s_i, \pi_{-i}) = U_i(s_i, b_{\pi_{-i}})$.

Let PBR'_i denote pure best replies (i.e., mixed strategies with support on a singleton) in \mathcal{G} , and let BR'_i be ordinary best replies, $BR'_i(\pi) = \arg \max_{\pi'_i} U'_i(\pi'_i, \pi_{-i})$. Note, from linearity, that we have the usual characterization of best replies; $\pi_i \in BR'_i(\pi)$ if and only if $E_i \in PBR'_i(\pi)$ for all $E_i \in \text{supp } \pi_i$.

Consider a player $i \in N$ and a strategy combination π in \mathcal{G} . For any $E_i \in \mathcal{S}_i$, and by definition (A.1), $E_i \subseteq PBR_i(b_\pi)$ if and only if $E_i \in PBR'_i(\pi)$, i.e.,

$$(A.2) \quad PBR'_i(\pi) = \mathcal{P}(PBR_i(b_\pi)).$$

Note that (A.2) will hold for any definition of $U'_i(E_i, \pi_{-i})$ as long as we have $U'_i(E_i, \pi_{-i}) < \max_{s_i \in E_i} U_i(s_i, b_{\pi_{-i}})$ whenever $\max_{s_i \in E_i} U_i(s_i, b_{\pi_{-i}}) > \min_{s_i \in E_i} U_i(s_i, b_{\pi_{-i}})$; the essential feature of (A.1).

By (A.2),

$$\begin{aligned} b_i^* \in BR_i(b) &\Leftrightarrow \forall E_i \in \text{supp } m_i^* : E_i \subseteq PBR_i(b) \\ &\Leftrightarrow \forall E_i \in \text{supp } m_i^* : E_i \in PBR'_i(m_b) \Leftrightarrow m_i^* \in BR'_i(m_b). \end{aligned}$$

We can now easily show:

PROPOSITION A.1. *b is an LPE of G , if and only if m_b is a Nash equilibrium of \mathcal{G} .*

Proof. Follows directly from the above. ■

PROPOSITION A.2. *b is a PLPE of G , if and only if m_b is a perfect equilibrium of \mathcal{G} .*

Proof. First note that (b^t) has a completely mixed mass function if and only if (m_{b^t}) is a sequence of completely mixed strategies in \mathcal{G} . Further, $b \in BR(b^t)$ if and only if $m_b \in BR'(m_{b^t})$, and finally $b^t \rightarrow b$ as $t \rightarrow \infty$ if and only if $m_{b^t} \rightarrow m_b$ as $t \rightarrow \infty$. So, m_b is a perfect equilibrium of \mathcal{G} if and only if b is a PLPE of G . ■

These results suggest that one property of an appropriate definition of LP-domination between belief functions on player i would be that b_i LP-dominates b'_i if and only if m_i dominates m'_i in \mathcal{G} .

ACKNOWLEDGMENT

We are grateful to Birgit Grodal and two anonymous referees of this journal for comments on earlier versions of this paper.

NOTES

¹ A capacity is a mapping $v : 2^X \rightarrow [0, 1]$, fulfilling $v(\emptyset) = 0$, $v(X) = 1$, and $E \subseteq F \implies v(E) \leq v(F)$ (monotonicity). It is exact if for all E , $v(E) = \min\{\pi(E) | \pi \in \text{core}(v)\}$, where the core is defined as for belief functions, cf. Schmeidler (1972).

² The proof of this lemma and all succeeding results are to be found in Appendix A.

³ An example is the cps μ_1 given by $\mu_1(C, C) = 0.32$, $\mu_1(C, D) = 0.34$, $\mu_1(D, C) = 0.01$, and $\mu_1(D, D) = 0.33$ with μ_2 defined similarly. Here player 1 is ambiguity averse since μ_1 decreases in the order $DC - CC - DD - CD$, where μ_1 increases, but $\mu_1((C, C) | \{(C, D), (C, D)\}) = 0.32 / (0.32 + 0.34)$ is very close to but below $1/2$, while $\mu_1((D, C) | \{(D, C), (D, D)\}) = 0.01 / (0.01 + 0.33)$ is very close to 0, maintaining the qualitative feature of (5.1) that player 1 is more optimistic when he plays C than when he plays D . With these preferences one can calculate that $b_i(C) = 411/1533$ and $b_i(D) = 0$ for $i = 1, 2$ constitute an LPE.

⁴ If a capacity on a set X is 2-monotone (or convex), the Choquet integral of a function $u : X \rightarrow \mathbb{R}$ is simply $\min_{\pi \in \text{core}(v)} \sum_{x \in X} u(x)\pi(x)$. For a general definition of the Choquet integral and theories of Choquet expected utility, see Gilboa (1987), Schmeidler (1989), and Sarin and Wakker (1992).

⁵ Unless one is ready to assume that all players are extremely pessimistic, as explained in Section 6.

REFERENCES

- Bernheim, B.D.: 1984, 'Rationalizable Strategic Behavior', *Econometrica* **52**(4), 1007–1028.
- Binmore, K.: 1992, 'Suppose Everybody Behaved Like That', MS, Economics Department, University of Michigan, Ann Arbor.
- Dow, J. and Werlang, C.: 1994, Nash Equilibrium under Knightian Uncertainty: Breaking Down Backward Induction', *Journal of Economic Theory* **64**(2), 305–324.
- Eichberger, J. and Kelsey, D.: 1995, 'Non-additive Belief and Game Theory', Mimeo, University of Saarbrücken.
- Ellsberg, D.: 1961, 'Risk, Ambiguity and the Savage Axioms', *Quarterly Journal of Economics* **75**, 643–669.
- Fudenberg, D. and Levine, D.: 1993, 'Self-Confirming Equilibrium', *Econometrica* **61**(3), 523–547.
- Gilboa, I.: 1987, 'Expected Utility with Purely Subjective Non-additive Probabilities', *Journal of Mathematical Economics* **16**(1), 65–88.
- Gilboa, I. and Schmeidler, D.: 1989, 'Maximin Expected Utility with Non-unique Prior', *Journal of Mathematical Economics* **18**(2), 141–153.
- Greenberg, J.: 1995, 'Stable (Incomplete) Contracts in Dynamic Games', mimeo, McGill University.
- Harsanyi, J.: 1973, 'Games with Randomly Distributed Payoffs: A New Rationale for Mixed Strategy Equilibrium Points', *International Journal of Game Theory* **3**, 211–225.
- Hendon, E.: 1995, 'Properties of Various Representations of Preferences on Lower Probabilities', Ch. 8 in *Fictitious Play in Games and Lower Probabilities in Decision Theories*, Ph.D. Dissertation, University of Copenhagen.
- Hendon, E., Jacobsen, H.J., Sloth, B. and Tranæs, T.: 1994, 'Expected Utility with Lower Probabilities', *Journal of Risk and Uncertainty* **8**(2), 197–216.
- Hendon, E., Jacobsen, H.J., Sloth, B. and Tranæs, T.: 1996, 'The Product of Capacities and Belief Functions', *Mathematical Social Sciences*, **32**, 95–108.
- Hofstadter, D.R.: 1983, 'Metamagical Themes', *Scientific American* **248**, 14–20.
- Howard, J.: 1988, 'Cooperation in the Prisoner's Dilemma', *Theory and Decision* **24**, 203–213.
- Hurwicz, L.: 1951, 'Optimality Criteria for Decision Making under Ignorance', Cowles Commission Discussion Paper Statistics, No. 370.
- Jaffray, J.-Y.: 1989, 'Linear Utility Theory for Belief Functions', *Operations Research Letters* **8**, 107–112.
- Klibanoff, P.: 1993, 'Uncertainty, Decision, and Normal-form Games', Mimeo, Stanford University.
- Lo, K.C.: 1995, 'Equilibrium in Beliefs under Uncertainty', Mimeo, University of Toronto.
- Pearce, D.G.: 1984, 'Rationalizable Strategic Behaviour and the Problem of Perfection', *Econometrica* **52**(4), 1029–1050.
- Rapoport, A.: 1966, *Two-Person Game Theory*, Ann Arbor, Michigan: University of Michigan Press.
- Rubinstein, A.: 1991, 'Comments on the Interpretation of Game Theory', *Econometrica* **59**(4), 909–924.

- Shafir D. and Tversky, A.: 1992, 'Thinking through Uncertainty: Nonconsequential Reasoning and Choice', *Cognitive Psychology* **24**, 449–474.
- Sarin, R.K. and Wakker, P.: 1995, 'On the Interpretation of Likelihood in Choquet Expected Utility', Mimeo.
- Schmeidler, D.: 1972, 'Cores of Exact Games, I', *Journal of Mathematical Analysis and Applications* **40**, 214–225.
- Schmeidler, D.: 1989, 'Subjective Probability and Expected Utility without Additivity', *Econometrica* **57**(3), 571–587.
- Shafer, G.: 1976, *A Mathematical Theory of Evidence*, Princeton: Princeton University Press.

*Institute of Economics
University of Copenhagen
Stuadiestræde 6
DS-1455 Copenhagen K.
Denmark*