American Economic Journal: Microeconomics 2009, 1:1, 1–16
http://www.aeaweb.org/articles.php?doi=10.1257/mic.1.1.1

Assessing Strategic Risk†

By R. J. Aumann and J. H. Dreze*

In recent decades, subjective probabilities have been increasingly applied to an adversary's choices in strategic games (SGs). In games against nature (GANs), the subjective probability of a state can be elicited from lotteries yielding utility 1 if that state obtains, 0 otherwise. But in SGs, making such a lottery available changes the game, and so the players' incentives. Here, we propose a definition of subjective probabilities in SGs that uses actually available strategies only. The definition applies also to GANs where the decision maker's options are restricted. The probabilities that emerge need not be unique, but expected utilities are unique. (*JEL* D81)

Half a century ago, when decision theory and game theory were young, it was common to perceive a dichotomy between (a) *games against nature,* in which the "adversary" is a neutral "nature," and (b) *strategic games,* in which the adversary is an interested party or parties. Games against nature were analyzed using several criteria, most prominent being the maximization of subjective expected utility—i.e., expected utility when the probabilities assigned to nature's moves are "subjective" or "personal," as in Leonard J. Savage (1954);[1] whereas strategic games were analyzed by minimax or, more generally, strategic equilibrium in the sense of John F. Nash (1951). No need was seen to reconcile or even relate the approaches, perceived as proceeding from distinct conceptual foundations.

In the ensuing years, the dichotomy gradually disappeared. It was recognized that games against nature and strategic games are in principle quite similar, and can—perhaps should—be treated similarly. Specifically, a player in a strategic game should be able to form subjective probabilities over the strategies of the other players, and his own strategy choice should yield him maximal expected utility with respect to these subjective probabilities.

There is, however, a difficulty with applying the notion of subjective probability to strategic games. In games against nature, subjective probabilities are constructed from the decision maker's preferences among gambles that "stake prizes" on nature's possible choices. In strategic games, this translates to staking prizes on

* Aumann: Department of Mathematics and Center for Rationality, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel (e-mail: raumann@math.huji.ac.il); Dreze: CORE, Université Catholique de Louvain, 34, Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium (e-mail: dreze@core.ucl.ac.be). The authors are grateful to Sergiu Hart, Martin Meier, Jean-Francois Mertens, Bernard Walliser, and Shmuel Zamir for important input. We alone remain responsible for the contents.

† To comment on this article in the online discussion forum visit the articles page at: http://www.aeaweb.org/articles.php?doi=10.1257/mic.1.1.1

[1] Related ideas appeared in Frank P. Ramsey (1931) and Bruno de Finetti (1937), and are taken up in Jacques H. Dreze (1961) and Frank J. Anscombe and Robert J. Aumann (1963).

strategy choices of the adversary. But that may change the incentives of the players, and in particular, the propensity to play this or that strategy. This invalidates the whole process.

The crux of the difficulty is that games of strategy are delicate objects: any modification of the definition of a game, however innocuous at first sight, may change the nature of the game and the behavior of the players. Accordingly, we must take the game as given, also when analyzing a player's preferences.

This paper starts from the premise that a player (the *protagonist*) has well-defined preferences over the outcomes of the game, over the pure strategies open to her, and over lotteries among these elements. We impose on these preferences consistency requirements identical to those used for games against nature. We show that these preferences imply existence of a utility function on the outcomes, and a probability distribution on the strategies of the other players, such that the preferences over lotteries (and hence over mixed strategies) are consistent with expected utility maximization. Though probabilities need not be unique (Section VB), expected utilities are unique.

In order to interpret that result, we show that it also obtains when analyzing a game against nature where the decision maker must choose one from a given set of acts,[2] and preferences are elicited only for these acts, their consequences, and lotteries among these elements.

Thus, expected utility maximization emerges under the same premises and with the same characterization in strategic games and in games against nature. In that sense, the dichotomy vanishes altogether.

Section I describes our framework informally. Section II sets forth our main result informally. Section III is devoted to mathematical preliminaries, Section IV to the formal statement of our result, Section V to discussion, Sections VI and VII to proofs, and Section VIII to the literature.

I. A Common Framework

It will be useful to use the same terminology for games against nature (GANs) and for strategic games (SGs). In either case, the "adversary"—be it nature or an interested party or parties—has several alternatives, called *strategies of the adversary.* The decision maker—henceforth *protagonist*—also has several alternatives, called *strategies of the protagonist.* Together, the strategies of the adversary and of the protagonist determine the outcome of the game. Thus, each of the protagonist's strategies may be thought of as a function from the adversary's strategies to the possible outcomes: an "act" in the terminology of Savage (1954), a "horse lottery" in that of Anscombe and Aumann (1963) (henceforth A-A).

A game (either GAN or SG) is thus defined by the set $C = \{c\}$ of outcomes (consequences), the set $R = \{r\}$ of strategies available to the protagonist (acts), the set

[2] For instance, the acts defining a specific decision problem—like choosing an insurance policy, or a portfolio of assets, and so on.

$S = \{s\}$ of strategies available to the adversary (states), and[3] an *outcome function* h, which associates a consequence $h(r, s)$ with each strategy pair (r, s).

Following standard practice in decision theory, we rely on the primitive concept of "preference." Preferences are applied to outcomes c as well as to strategies r. To obtain cardinal preferences, we introduce *lotteries,* defined as objective probability distributions. A *mixed consequence* is a lottery on consequences, an element γ of the set $\Delta(C)$ of lotteries on C. A *mixed strategy* is an element ρ of the set $\Delta(R)$ of lotteries on R. In order to calibrate utilities for the elements of C and R to the same scale, we introduce *hybrid lotteries,* i.e., elements λ of the set $\Delta(R \cup C) =: \Delta$.

Operationally, such a lottery λ results either in (a) the outright selection of a specified pure outcome of the game, or in (b) the game being played, with the protagonist choosing a specified pure strategy. More specifically, alternative (b) results in awarding to the protagonist the outcome associated by the game with a specified strategy of the protagonist, combined with the strategy actually chosen by the adversary when playing the game. For each strategy s of the adversary, λ yields a mixed consequence λ_s in a natural way: if λ chose a consequence, then λ_s chooses the same consequence; and if λ chose a pure strategy r, then λ_s chooses the outcome of the game when the protagonist chooses r and the adversary chooses s. Note that all mixed consequences are in Δ, as are all mixed strategies; thus the preferences on Δ apply also to mixed consequences and to mixed strategies.

If a hybrid lottery λ results in the game $G := (C, R, S, h)$ being played, the definition of G is fully respected, and it will be played as such, irrespective of what consequence c may have been specified by λ as a mutually exclusive alternative to G.

II. Main Result

The following two assumptions are made:

N-M: The preference order satisfies the usual assumptions of von Neumann-Morgenstern utility theory; and

MONOTONICITY: If one hybrid lottery λ always yields a mixed consequence preferred to that yielded by another one λ', no matter what the adversary does, then λ is preferred to λ'; likewise for weak preference.[4]

We then have the following:

MAIN THEOREM (verbal statement): *There exists a function on the consequences—unique up to positive linear transformations*[5]*—called a* utility *function, and a probability distribution p on the adversary's strategies, such that*

[3] Throughout, terms being defined are italicized.
[4] Preference or indifference.
[5] Multiplication by a positive constant and addition of an arbitrary constant.

(*i*) *one hybrid lottery is preferred to another if and only if its expected utility is greater; and*

(*ii*) *the expected utility of a hybrid lottery is the same for all p satisfying (i).*

Condition (i) says that the utility u and the probabilities p represent the preferences numerically. Condition (ii) says that though the subjective probabilities are not unique, they are "payoff equivalent," in that for each of the alternatives available to the protagonist, all yield the same payoff. The theorem applies to games of strategy as well as to games against nature, through suitable interpretation of the common framework.

III. Formal Treatment: Preliminaries

The set of all probability distributions[6] on a finite set A is denoted $\Delta(A)$. Note that if $\alpha, \alpha' \in \Delta(A)$ and $t \in (0,1)$, then also $t\alpha + (1-t)\alpha' \in \Delta(A)$. Abusing our notation, we write α and a interchangeably if α assigns probability 1 to a; that is, we do not distinguish between a and a lottery that chooses a with certainty. No confusion should result.

A *preference order* $\gtrsim$ on $\Delta(A)$ is a transitive,[7] reflexive,[8] and complete[9] binary relation on $\Delta(A)$. If $\alpha \gtrsim \beta$ and $\beta \gtrsim \alpha$, write $\alpha \sim \beta$ and say that α is *indifferent* to β. If $\alpha \gtrsim \beta$ and $\alpha \nsim \beta$, write $\alpha \succ \beta$ and say that α is *preferred* to β. An *N-M utility* for $\gtrsim$ is a real-valued function u on $\Delta(A)$ such that

$$\alpha \gtrsim \alpha' \text{ iff } u(\alpha) \geq u(\alpha'), \text{ and} \tag{1}$$

$$u(t\alpha + (1-t)\alpha') = tu(\alpha) + (1-t)u(\alpha'). \tag{2}$$

Various mutually equivalent axiom systems for N-M utility theory are available (John von Neumann and Oskar Morgenstern 1944, R. Duncan Luce and Howard Raiffa 1957, and others). We say that a preference order *satisfies the axioms of von Neumann-Morgenstern utility theory*—and call it an *N-M preference order*—if it satisfies any one of those systems.

PROPOSITION A: *An N-M preference order on* $\Delta(A)$ *has an N-M utility.*

[6] Nonnegative real-valued functions whose values sum to 1.
[7] $\alpha \gtrsim \beta$ and $\beta \gtrsim \gamma$ imply $\alpha \gtrsim \gamma$.
[8] $\alpha \gtrsim \alpha$.
[9] $\alpha \gtrsim \beta$ or $\beta \gtrsim \alpha$.

IV. Formal Treatment: The Main Theorem

The viewpoint taken here is that of a single player, the *protagonist,* also called *Rowena;* it is her[10] subjective probabilities for the strategy choices of the other players that we will define. Also, the preferences appearing below are hers, as are the utilities. It is convenient to combine all the other players into a single one, called *Colin;* we will see that no loss of generality is involved.

A *game G* consists of

- a finite set R with members r (Rowena's *pure strategies*),
- a finite set S with members s (Colin's *pure strategies*),
- a finite set C with members c (*pure consequences*), and
- a function $h : R \times S \to C$ (the *outcome function*).[11]

Members $\rho, \rho', \ldots$ of $\Delta(R)$ are called *mixed strategies of Rowena;* members $\gamma, \gamma', \ldots$ of $\Delta(C)$ are called *mixed consequences;* members $\lambda, \lambda', \ldots$ of $\Delta(R \cup C)$ (henceforth simply Δ) are called *hybrid lotteries* (or simply *lotteries*). If $\lambda \in \Delta$, set $\lambda = t\rho + (1-t)\gamma$, where $\rho \in \Delta(R)$, $\gamma \in \Delta(C)$, and $t \in [0,1]$. For each pure strategy s of Colin, let $\rho_s \in \Delta(C)$ be the mixed consequence that results when Rowena plays ρ and Colin plays s, and let $\lambda_s := t\rho_s + (1-t)\gamma$ be the mixed consequence that results when Rowena uses the lottery λ and Colin plays s. Call a preference order $\gtrsim$ on Δ *monotonic* if $\lambda \gtrsim \lambda'$ whenever $\lambda_s \gtrsim \lambda'_s$ for all s, and $\lambda \succ \lambda'$ whenever $\lambda_s \succ \lambda'_s$ for all s.

Now, let $G = (R, S, C, h)$ be a game, $\gtrsim$ a monotonic N-M preference order on Δ; so in particular, $\gtrsim | \Delta(C)$ is an N-M preference order on $\Delta(C)$, so has an N-M utility u, unique up to positive linear transformations.

MAIN THEOREM (formal statement): *There exists a probability distribution p on S, such that for any hybrid lotteries λ, λ',*

$$(3) \qquad \lambda \gtrsim \lambda' \text{ if and only if } \sum_{s \in S} p_s u(\lambda_s) \geq \sum_{s \in S} p_s u(\lambda'_s);$$

and, if p^* is another such probability distribution, then for any hybrid lottery λ,

$$(4) \qquad \sum_{s \in S} p_s u(\lambda_s) = \sum_{s \in S} p^*_s u(\lambda_s).$$

For hybrid lotteries λ, define

$$(5) \qquad u(\lambda) := \sum_{s \in S} p_s u(\lambda_s);$$

because of (4), this does not depend on p, so (3) becomes

[10] The protagonist is female. The other players are of indeterminate gender; we refer to them as "he," to distinguish them from the protagonist.

[11] Without loss of generality, we could take $C = R \times S$ and let h be the identity; but nothing would be gained thereby, the notation would become more cumbersome, and the ideas less transparent.

$$\lambda \gtrsim \lambda' \text{ if and only if } u(\lambda) \geq u(\lambda'). \tag{6}$$

In words, Rowena evaluates lotteries by their expected utility, which is uniquely defined on Δ (up to positive linear transformations).

Note that the Main Theorem applies to GANs as well as to SGs.

V. Discussion

A. *GANs and SGs*

In a game against nature, nature is oblivious to the protagonist and her options; therefore, even if her options change, her probabilities for nature's choices should not. In a strategic game, the adversary takes the protagonist's options into account when choosing his strategy; therefore, if the protagonist's options change, her probabilities for the adversary's choices may well change.

In GANs, the standard approach (Savage 1954, A-A 1963) to defining the protagonist's probability for a particular state s of nature uses a strategy r_s of the protagonist that yields her utility 1 if s occurs, 0 otherwise. Her probability for s is then defined as her utility $u(r_s)$ for this strategy; in words, the number p such that she would as soon have a dollar with objective[12] probability p as a dollar if nature chooses s. If r_s is not available in the given GAN, just add it; as noted above, the protagonist's probability for s should not be affected.

In SGs, this does not work, because adding r_s to the protagonist's options *does* affect her probabilities for s—again, as noted above.

The approach described in the preceding sections overcomes the difficulty by restricting attention to those strategies that are actually available to the protagonist in the given game; adding r_s is forbidden. While the method was developed for SGs, formally it applies equally well to GANs; specifically, to GANs in which the set of strategies available to the protagonist is—for whatever reason—restricted. Thus the formal distinction between GANs and SGs vanishes completely.

But conceptually, an important distinction between the two does remain. In GANs, even when the protagonist's strategies are in fact restricted, she can *imagine* the strategy r_s, and so evaluate it. But in SGs, adding r_s changes the adversary's view of the game; therefore imagining r_s cannot lead to a coherent definition[13] of the protagonist's subjective probabilities.

B. *Nonuniqueness and Payoff Equivalence*

The subjective probabilities in the Main Theorem need not be unique. A simple example is a two-person game whose matrix has two identical columns, in which case the total subjective probability assigned to both columns can be divided between them in an arbitrary way. They are, however, *payoff equivalent,* in the sense that for

[12] "Objective" probabilities are associated with coin tosses, roulette spins, and the like.
[13] See the item entitled "*Ignorance*" in Section VD.

a given utility function on the consequences, all subjective probability distributions satisfying the Main Theorem yield the same expected utility for each hybrid lottery. That is condition (4) in the formal statement of the Main Theorem.

Thus, consider a three-way election for president of a certain country, with candidates *A*, *B*, and *C*. If the protagonist must choose between getting utility 1 with certainty and a bet that yields utility 2 if *A* wins and 0 otherwise, then our procedure will uniquely determine only her probability p_A for *A* winning. For the probabilities p_B and p_C that *B* or *C* will win, we can say only that they sum to $1 - p_A$, but nothing about their individual values; this is the nonuniqueness. The payoff equivalence says that the individual values of p_B and p_C don't matter to the protagonist when making her choice; she should choose the bet if and only if $p_A > 1/2$.

C. *Hybrid Lotteries*

Hybrid lotteries may be interpreted as strategies in extensive games of a kind often seen in real life. Specifically, each hybrid lottery λ may be seen as yielding a mixed strategy ρ with probability *t*, and a mixed outcome γ with the complementary probability $1 - t$. Denote by G^{λ} the extensive game in which nature chooses, with respective probabilities *t* and $1 - t$, whether *G* is to be played or whether the outcome is to be γ; then the lottery λ is equivalent (for the protagonist Rowena) to playing ρ in G^{λ}. One would not expect the adversary Colin to play G^{λ} differently[14] from *G*, so the protagonist's preferences between the different lotteries λ accurately reflect her estimate of how the adversary will play *G*.

Some readers have asked whether considering hybrid lotteries is like considering the original game to which one has added acts yielding Rowena a constant, independent of Colin's choice. The answer is no. Rowena cannot use constant acts that are not in *R* when playing the game; Colin knows this, and, indeed it is commonly known.

D. *Some Dead Ends*

"Admissible" Preferences.—One might have thought it sufficient to work with the space $\Delta(C) \cup \Delta(R)$ consisting only of mixed consequences and mixed strategies, rather than the much larger space $\Delta(C \cup R)$ of hybrid lotteries. Preferences on $\Delta(C) \cup \Delta(R)$ induce preferences on each of $\Delta(C)$ and $\Delta(R)$, and a "matching" of $\Delta(C)$ with $\Delta(R)$. Call the preferences on $\Delta(C) \cup \Delta(R)$ *admissible* if each of the induced preference orders is N-M, the one on $\Delta(R)$ is monotonic, and there is also monotonicity as "between" $\Delta(C)$ and $\Delta(R)$; i.e., if Rowena (weakly) prefers a mixed consequence γ to all outcomes of a mixed strategy ρ, no matter what Colin does, then she (weakly) prefers γ to ρ, and similarly in the opposite direction. In utility terms, we get N-M utility functions *u* on $\Delta(C)$ and $\Delta(R)$ that are *calibrated to the same scale,* so that it is meaningful to compare the utilities of mixed consequences and mixed

[14] For each player *i*, the matrix of G^{λ} is obtained from that of *G* by multiplying the whole matrix by the constant *t* and adding the constant $(1-t)u^i(\gamma)$, where u^i is *i*'s utility.

strategies, and appropriate to use the same notation—u—for both. Admissibility is about all one can ask for without going to hybrid lotteries; but it is not enough to yield utilities and subjective probabilities representing the preferences. For example, consider Game 1 (see display), consequences c being denoted by their utilities $u(c)$. Suppose $u(T) = 3/4$ and $u(B) = 1/2$. This induces an admissible preference order[15] on $\Delta(C) \cup \Delta(R)$. If q were the subjective probability of Colin's playing left, condition (5) in Section IV would yield $3/4 = u(T) = q \cdot 1 + (1 - q) \cdot 0 = q$, and $1/2 = u(B) = q \cdot 2/3 + (1 - q) \cdot 1/3$, so $q = 1/2$, a contradiction.

	L	R
T	1	0
B	2/3	1/3

Game 1

Separate Preferences on Strategies and Consequences.—Another potential option is to forgo the "matching" between $\Delta(C)$ and $\Delta(R)$ (which "drives" the above example), and to proceed from separate N-M preference orders on $\Delta(C)$ and $\Delta(R)$, obeying monotonicity on $\Delta(R)$. With this, one does get utilities and subjective probabilities representing the preferences. Indeed, we have the following:

AUXILIARY THEOREM: *Let $\trianglerighteq$ and $\gtrsim$ be N-M preference orders on $\Delta(C)$ and $\Delta(R)$ respectively, with $\gtrsim$ monotonic; i.e., $\lambda \gtrsim \lambda'$ whenever $\lambda_s \trianglerighteq \lambda'_s$ for all s, and $\lambda > \lambda'$ whenever $\lambda_s \triangleright \lambda'_s$ for all s. Then there exists a utility function u on $\Delta(C)$, unique up to positive linear transformations, and a probability distribution p on S, such that for λ, λ' in $\Delta(R)$,*

$$\lambda \gtrsim \lambda' \text{ if and only if } \sum_{s \in S} p_s u(\lambda_s) \geq \sum_{s \in S} p_s u(\lambda'_s).$$

But here, payoff equivalence (the analogue of condition (4) in Section IV), is *not* guaranteed. Indeed, consider Game 2 (see display), where, as before, consequences are denoted by their utilities. The relation $T \gtrsim B$ fully determines a preference order on $\Delta(R)$, which is represented by $p = (p_L, p_R) = (t, 1 - t)$ for any t in the half-open interval $(1/2, 1]$. The corresponding utilities are $u(T) = t$ and $u(B) = 1 - t$, which are of course different for different t.

[15] Monotonicity on $\Delta(R)$ is vacuously fulfilled, since for mixed strategies ρ and ρ', preferences between ρ_L and ρ'_L are opposite to those between ρ_R and ρ'_R. As between $\Delta(C)$ and $\Delta(R)$, let $\rho := \alpha T + (1 - \alpha)B \in \Delta(R)$. Then $\rho \sim 1/2 + 1/4\alpha$, which is always strictly between the consequences $2/3 + 1/3\alpha$ and $1/3 - 1/3\alpha$ that may result when ρ is played. Thus if γ is $\geq$ each ρ_s, then $\gamma \geq 2/3 + 1/3\alpha > 1/2 + 1/4\alpha \sim \rho$, and if each ρ_s is $\geq \gamma$, then $\rho \sim 1/2 + 1/4\alpha > 1/3 - 1/3\alpha \geq \gamma$.

	L	R
T	1	0
B	0	1

Game 2

The failure of payoff equivalence is a serious drawback, as the utilities of strategies are not calibrated to the same scale as those of consequences; the protagonist has no clear idea of what using a particular strategy is worth to her. So if we are to keep the game unchanged, hybrid lotteries remain as the only satisfactory option for defining the protagonist's subjective probabilities.

Side Bets.—The standard approach to defining the protagonist's subjective probabilities in games against nature relies on "side bets." In such a bet she gets, in *addition* to her payoff from the game, an amount δ if the adversary plays a specified strategy s; nothing else is changed. One then defines her probability for s as that number p such that she would as soon opt for the side bet, as for δ with objective probability p.

In strategic games this doesn't work because, as we said in the introduction, side bets may change the game. For example,[16] in the coordination game G_1 below, one may expect the Pareto dominant outcome BR. A side bet on L in the amount of $\delta = 8$ adds 8 to the row player's payoffs in column L, i.e., transforms the game to G_2. If side bets "don't matter," we should expect BR in G_2 as well. By the same token, adding 8 to the column player's payoffs in row T of G_2 should not matter; this yields G_3, commonly known as the "Stag Hunt."[17] But here BR, which is Pareto dominated by TL, is far from compelling; indeed, John C. Harsanyi and Reinhard Selten (1987) select BR in G_1 and TL in G_3.

	L	R
T	1, 1	0, 0
B	0, 0	7, 7

G_1

	L	R
T	9, 1	0, 0
B	8, 0	7, 7

G_2

	L	R
T	9, 9	0, 8
B	8, 0	7, 7

G_3

Side bets leave most equilibrium notions—including that of Nash (1951) and correlated equilibrium (Aumann 1974)—invariant. Nevertheless, they subtly change incentives, as the example shows.

[16] Communicated by Sergiu Hart. To avoid difficulties, assume dollar payoffs and linear utilities.
[17] See Barry O'Neill (1994, 1004–5) for a discussion of this game and some of the literature on it.

Ignorance.—As discussed in Section VA, in GANs the probability of a state s of nature is Rowena's utility for a strategy r_s yielding utility 1 if s occurs, 0 if not. In SGs this does not work because r_s is usually unavailable to Rowena; adding it may change Colin's view of the game, and so Rowena's probability that he chooses s.

One might think it enough to imagine a situation in which r_s really *is* available to Rowena, but *Colin does not know that it is,* so his choices—and Rowena's probabilities for his making those choices—are not affected. But that is not very satisfactory. What does Rowena think about Colin's state of mind? If she considers it possible that he considers it possible that r_s is available to her, then that already changes the game, and we have the same difficulty as before. If not, then she knows that he knows that it is unavailable. But then how can it be available? There is a basic incoherence in situations where something false is "known."

Adding Dominated Strategies.—Though in SGs, adding r_s may in general change Rowena's probabilities for Colin's choices, one might think that that is not so when r_s is *strictly dominated*. If so, we could define Rowena's probabilities as follows: without loss of generality, take all of Rowena's payoffs in the given game to be > 1; otherwise, simply recalibrate the utility function. Now, for each strategy s of Colin, add the strategy r_s of Rowena. She will not use this strategy, as it is strictly dominated; so for practical purposes, the game appears unchanged, and Rowena's utility for r_s should constitute an adequate definition of her probability that Colin chooses s.

But on closer examination, this, too, breaks down. To eliminate the possibility that Rowena will use a strictly dominated strategy, Colin must know that she is rational. So if adding r_s is not to affect Rowena's probabilities of Colin's choices, she must know that he knows that she is rational. For this, we must assume at least second-order knowledge of rationality, which in a general theory of probability assessments in games is unacceptably strong.[18]

In fact, more than second-order knowledge of rationality is needed; nothing less than common knowledge of rationality will do. If Colin does not know that Rowena knows that he knows she is rational, then he might think that she thinks that he thinks she might use r_s; in that case he would choose accordingly, so the game would be essentially affected after all. So he does have to know that, and she must know that he knows it. And so on.

E. *The Utility of Playing a Game*

The utility of playing a game for the protagonist is naturally defined as the utility of her most preferred strategy. That this is well-defined is a consequence of payoff equivalence; for example, if one uses only the Auxiliary Theorem, then one does not get a utility for playing the game.

We purposely use the term "utility"—rather than value—since in game theory "value" means something else. For example, the value of tic-tac-toe is a draw, whereas the utility of playing a round of the game could be, say, a win, depending

[18] For example, it is not assumed in Aumann and Adam M. Brandenburger (1995).

on how the protagonist assesses the situation. In brief, one might say that the utility is an expectation, whereas the value is a *rational* expectation; for further discussion, see Aumann and Dreze (2008, section VIIC).

F. *Beliefs about Beliefs*

The method described here yields the protagonist's probability assessments of what the adversary will *do,* but not of what he *believes.* Belief systems (or hierarchies), which embody players' probability assessments of each other's probability assessments, have played a central role in game theory for over forty years, ever since the pioneering work of Harsanyi (1967–68). The current work does not enable construction of such systems from preferences.

VI. Affine Monotonic Functions

For points x, y in $\mathbb{R}^n$, write $x \gg y$ if $x_i > y_i$ for all i, and write $x \geq y$ if $x_i \geq y_i$ for all i. A real-valued function[19] f from a convex set D in $\mathbb{R}^n$ to $\mathbb{R}$ is called *affine* if $f(tx + (1 - t)y) = tf(x) + (1 - t)f(y)$ for all x, y in D and t in $(0, 1)$. It is called *monotonic* if $x \gg y$ implies $f(x) > f(y)$, and $x \geq y$ implies $f(x) \geq f(y)$, for all x, y in D.

PROPOSITION B: *Let D be a convex subset of $\mathbb{R}^n$, and g an affine monotonic real-valued function on D. Then there exist nonnegative $q_1,\ldots,q_n$, not all of which vanish, and a real q_0, such that $g(x) = q_0 + \sum_{i=1}^{n} q_i x_i$ for all x in D.*

In this section, we prove Proposition B. Readers willing to accept the proposition on faith may proceed to the proof of the main results in the next section.

The origin $(0,\ldots,0)$ of $\mathbb{R}^n$ is denoted $\mathbf{0}$. A *linear subspace* (or simply *subspace)* L of $\mathbb{R}^n$ is a subset of $\mathbb{R}^n$ that, together with any two points x, y in it, and any real number t, contains $x + y$ and tx. A function f on L is *linear* if $f(x + y) = f(x) + f(y)$ and $f(tx) = tf(x)$ for all x, y in L and all real t. Note that a function on L is linear if and only if it is affine. If f is a linear function on L, and t is a constant, then the set $\{x \in L : f(x) \geq t\}$ is a (closed) *half-space* of L, and the set $\{x \in L : f(x) = t\}$ is a *hyperplane* in L; the hyperplane *separates* two convex subsets D and D' of L if $f(x) \geq t$ for all x in D and $f(x) \leq t$ for all x in D'. A *polyhedral* convex set is the intersection of half-spaces. A *linear manifold* is a hyperplane in some subspace of $\mathbb{R}^n$. The *relative interior* of a convex set D in $\mathbb{R}^n$, denoted $ri(D)$, is its interior relative to the smallest linear manifold containing it.

LEMMA 1: *Let D and D' be nonempty convex sets in $\mathbb{R}^n$, with D polyhedral. Then a necessary and sufficient condition for the existence of a hyperplane that separates D from D' and does not include D' is that D does not meet $ri(D')$.*

[19] To avoid trivialities, we assume in the sequel that *sets* are not singletons and *functions* are not constant.

PROOF:

R. Tyrrell Rockafellar (1970, Theorem 20.2).

LEMMA 2: *Any monotonic linear function g on a subspace L of* $\mathbb{R}^n$ *may be extended to a monotonic linear function on all of* $\mathbb{R}^n$.

PROOF:

Define $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \geq \mathbf{0}\}$, and $L\_ := \{x \in L : g(x) \leq 0\}$. Then $ri(L\_) = \{x \in L : g(x) < 0\}$; since g is monotonic, $\mathbb{R}^n_+$ does not meet $ri(L\_)$. As $\mathbb{R}^n_+$ is polyhedral, by Lemma 1 there is a hyperplane H in $\mathbb{R}^n$ that separates $\mathbb{R}^n_+$ from $L\_$ and does not include $L\_$; choose the linear function f on $\mathbb{R}^n$ defining H to be nonnegative on $\mathbb{R}^n_+$, nonpositive on $L\_$. We claim that for each x in L,

$$g(x) = 0 \textit{ if and only if } f(x) = 0. \tag{7}$$

To see this, suppose first that $g(x) = 0$; then $x \in L\_$, so $f(x) \leq 0$. Also $g(-x) = -g(x) = 0$, so $-x \in L\_$, so $f(-x) \leq 0$, so $f(x) \geq 0$, so $f(x) = 0$. In the opposite direction, let $f(x) = 0$, and suppose that $g(x) \neq 0$. Then for any y in L, we have $g(y - (g(y)/g(x))x) = 0$, so $f(y - (g(y)/g(x))x) = 0$, so $f(y) = (g(y)/g(x))f(x) = 0$. So H includes all of L, and in particular $L\_$, contrary to what we stipulated; so (7) is proved. Since f is nonpositive on $L\_$, it follows that f and g always have the same sign.

Now, choose an x in L with $g(x) \neq 0$. Possibly redefining f by multiplication by a positive constant, we may take $f(x) = g(x)$. If y is any member of L, then $f(y - (f(y)/f(x))x) = f(y) - f(y) = 0$, so $g(y - (f(y)/f(x))x) = 0$, so $g(y) = f(y)g(x)/f(x) = f(y)$. So f extends g; and it is monotonic, as it is nonnegative on $\mathbb{R}^n_+$ and is not constant.

PROOF OF PROPOSITION B:

By possibly applying a translation, we may suppose without loss of generality that $\mathbf{0} \in ri(D)$. Let L be the smallest linear manifold that includes D. Since L contains $\mathbf{0}$, it is a linear space, and there is a unique extension of g to a linear function g' on L; as g is monotonic, so is g'. Applying Lemma 2, we obtain an extension f of g' from L to a linear monotonic function on all of $\mathbb{R}^n$. Let $f(x) = \sum_{i=1}^n q_i x_i$ for all x in $\mathbb{R}^n$; all linear functions on $\mathbb{R}^n$ have this form. The q_i cannot all vanish, for then f would be constant; and they are nonnegative, as f is monotonic. This proves[20] the proposition.

VII. Proofs of the Theorems

The idea is to think of the N-M expected utility of a lottery λ as an affine monotonic function of the S-vector $u(\lambda_S)$ of the utilities $u(\lambda_s)$ (Rowena's utility if Colin plays s), and then to apply Proposition B. Subtracting the additive constant q_0, and

[20] The term q_0 is due to the translation at the beginning of the proof.

then multiplying by a positive constant so that the other coefficients sum to 1, constitutes a positive linear transformation; so the result still represents Rowena's preferences, and we can think of the coefficients as Rowena's probabilities for Colin's strategies. When the vectors $u(\lambda_S)$ span the Euclidean space in which they live, then Proposition B follows from every linear monotonic function on $\mathbb{R}^n$ having the form $\sum_{i=1}^{n} q_i x_i$, with $q_i \geq 0$. When they do not span, one applies Lemma 2, which may be of some independent interest; it is not quite straightforward, and we have not found it in the literature.

For the formal proofs, please refer to the formal treatment in Section IV. Set $S := \{s_1, \ldots, s_n\}$. With each lottery λ in Δ, associate the point $u(\lambda_S) := (u(\lambda_{s_1}), \ldots, u(\lambda_{s_n}))$ in $\mathbb{R}^n$. Let D be the set of all the points $u(\lambda_S)$ when λ ranges over Δ; by (2), D is convex. Let w be an N-M utility function for $\gtrsim$, and define f on D by $f(x) := w(\lambda)$ for any λ for which $u(\lambda_S) = x$. That there is such a λ follows from $x \in D$; and by the monotonicity of $\gtrsim$, (1), and (2), f is well-defined as a function of x, and is affine and monotonic. So we may apply Proposition B; by multiplying w by a positive constant, we may take the q_i to be nonnegative and sum to 1. Then setting $p_{s_i} := q_i$ yields (3).

The proof of the Auxiliary Theorem is the same, except that $\Delta\ (= \Delta(R \cup C))$ is replaced by $\Delta(R)$, and u is a utility function for $\trianglerighteq$.

To prove the payoff equivalence (4), set $u_p(\lambda) := \sum_{s \in S} p_s u(\lambda_s)$ for all hybrid lotteries λ. Suppose that p and p^* satisfy (3), and let λ be a hybrid lottery. Preference-wise, λ must be between the most preferred and the least preferred consequence; so there is a mixed consequence ρ with $\lambda \sim \rho$. Then (3) yields $u_p(\lambda) = u_p(\rho) = u(\rho) = u_{p^*}(\rho) = u_{p^*}(\lambda)$, which is (4).

VIII. Literature

Luce and Raiffa (1957, 306) were among the earliest[21] to suggest assigning subjective probabilities to an adversary's choices in a strategic game; they wrote as follows: "The problem of individual decision making under uncertainty can be considered a one-person game against a neutral nature. Some of these ideas can be applied indirectly to individual decision making … where the adversary is not neutral but a true adversary. … One *modus operandi* for the decision maker is to generate an *a priori* probability distribution over the … pure strategies … of his adversary by taking into account both the strategic aspects of the game and … 'psychological' information … about his adversary, and to choose an act which is best against this … distribution." They go on to explore the idea of "side bets" (see the item entitled "*Side Bets*" in Section VD above), noting some difficulties with it, and informally suggesting a possible way around them. No formal model was developed, and no definite conclusion reached.

In the unpublished dissertation of Dreze (1958, 16), one reads "… it is always possible to formalize the decision problem facing a player in a game of strategy as a game against nature, where states of nature are described with reference to the

[21] They cite an earlier paper by Joseph L. Hodges, Jr., and Erich L. Lehmann (1952) who suggest that a player in a two-person zero-sum game might assign subjective probabilities to the eventuality that his adversary will make a "mistake." But this is not really in the spirit of this paper, nor of Luce and Raiffa's suggestion.

opponent's strategies. The usefulness of the theory of games of strategy resides in the fact that it helps the player to estimate the probabilities of the various states of nature so defined."

It appears that Walter Armbruster and Werner Böge (1979) and Böge and Theo Eisele (1979) were the first to construct formal models in which each player directly[22] assigns subjective probabilities to the strategy choices of the others. A relatively early application of this idea is Brandenburger and Eddie Dekel (1987). The representation of the value of a game to a player as a subjectively expected utility is implicit in the work of Robert F. Nau and Kevin F. McCardle (1990).

Joseph B. Kadane and Patrick D. Larkey (1982) wrote that the problem of a player in a game is no different from any other one-person decision problem. In particular, they suggested abandoning altogether all notions of equilibrium. Instead, they proposed simply that each player form, in some unspecified and unrestricted way, a probability distribution over the other players' strategies, and then maximize against that. To form the probabilities, they suggested using disciplines like cognitive psychology rather than decision or game theory.

In the precisely opposite direction, Marco Mariotti (1995, 1108) wrote that "a divorce is required between game theory and individual decision theory… strategic decision principles may be radically different from individual decision-theoretic principles."

T	G'
B	2, 2

Game G

	L'	R'
T'	1, 7	0, 0
B'	0, 0	3, 3

Game G'

	L'	R'
TT'	1, 7	0, 0
TB'	0, 0	3, 3
B	2, 2	2, 2

Game GG'

Game G (see display) is an extensive game: if Rowena chooses T, then G' is played; otherwise, both players get 2. Mariotti argues that in G, a prudent Rowena might well play B, which assures her 2, whereas if she plays T, she might get only 1—her payoff at a reasonable outcome of G' (the Pareto undominated strict Nash equilibrium (T', L')). Then she would also play B in GG', which is simply the strategic form of G. But in GG', we may first eliminate TT', by strong domination; then L', by weak domination,[23] and then B, as $3 > 2$.

The perspective of the current work resolves the difficulty. In the abstract, (T', L') indeed cannot be ruled out in G'. But if Rowena chose T in G, it's unlikely that she would choose T' in G'; presumably Colin realizes this, and Rowena realizes that he

[22] Previously, Aumann (1974) had already used subjective probability in analyzing games; but in that analysis, players use "subjectively mixed strategies"—peg their pure strategy choices on events (like outcomes of horse races) whose probability is not agreed upon—rather than simply assigning a subjective probability to the other players' choices.

[23] Mariotti uses a slightly different argument for this, but it comes to the same thing.

does, so she is likely to assign a high probability to L'. When Rowena plays T in G, she's not merely deciding to play G'; she's deciding to play G' *in a situation where she could have gotten 2 for sure*. That's an altogether different kettle of fish.

Some of the ideas underlying the current work appear already in Mariotti's stimulating paper. Inter alia, that a strategy in a game corresponds to an act in Savage's one-person decision theory; that "only some acts (strategies) are feasible for each player in a given game," and that the players should "rank only the strategies available in that game" (p. 1102).

Finally, Dreze (2005) proposes defining subjective probabilities in strategic games by using only the protagonist's "revealed" preference for the strategy she actually chooses over other strategies available to her.

REFERENCES

Anscombe, Frank J., and Robert J. Aumann. 1963. "A Definition of Subjective Probability." *Annals of Mathematical Statistics*, 34(1): 199–205 (A-A in the text).

Armbruster, Walter, and Werner Böge. 1979. "Bayesian Game Theory." In *Game Theory and Related Topics*, ed. O. Moeschlin and D. Pallaschke, 17–28. Amsterdam: Elsevier.

Aumann, Robert J. 1974. "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics*, 1(1): 67–96.

Aumann, Robert J., and Adam M. Brandenburger. 1995. "Epistemic Conditions for Nash Equilibrium." *Econometrica*, 63(5): 1161–80.

Aumann, Robert J., and Jacques H. Dreze. 2008. "Rational Expectations in Games." *American Economic Review*, 98(1): 72–86.

Böge, Werner, and Theo Eisele. 1979. "On Solutions of Bayesian Games." *International Journal of Game Theory*, 8(4): 193–215.

Brandenburger, Adam M., and Eddie Dekel. 1987. "Rationalizability and Correlated Equilibria." *Econometrica*, 55(6): 1391–1402.

Dreze, Jacques H. 1958. *Individual Decision Making under Partially Controllable Uncertainty*. PhD diss. Columbia University.

Dreze, Jacques H. 1961. "Les fondements logiques de l'utilité cardinale et de la probabilité subjective."[24] In *La Décision*, 73–87. Paris: Centre National de la Recherche Scientifique (CNRS).

Dreze, Jacques H. 1987. *Essays on Economic Decisions under Uncertainty*. Cambridge: Cambridge University Press.

Dreze, Jacques H. 2005. "Nested Identification of Subjective Probabilities." Louvain-la-Neuve, CORE Discussion Paper 2005/90 (revised November 2007).

de Finetti, Bruno. 1937. "La prévision: ses lois logiques, ses sources subjectives." *Annales de l'Institut Henri Poincaré*, 7(1): 1–68.

Harsanyi, John C. 1967–68. "Games of Incomplete Information Played by Bayesian Players, I–III." *Management Science*, 14(3): 159–82, (5) 320–34, (7) 486–502.

Harsanyi, John C., and Reinhard Selten. 1987. *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Kadane, Joseph B., and Patrick D. Larkey. 1982. "Subjective Probability and the Theory of Games." *Management Science*, 28(2): 113–20.

Luce, R. Duncan, and Howard Raiffa. 1957. *Games and Decisions*. New York: John Wiley and Sons.

Hodges, Joseph L., Jr., and Erich L. Lehmann. 1952. "The Uses of Previous Experience in Reaching Statistical Decisions." *Annals of Mathematical Statistics*, 23(3): 396–407.

Mariotti, Marco. 1995. "Is Bayesian Rationality Compatible with Strategic Rationality?" *Economic Journal*, 105(432): 1099–1109.

Nash, John F. 1951. "Non-cooperative Games." *Annals of Mathematics*, 54(2): 286–95.

Nau, Robert F., and Kevin F. McCardle. 1990. "Coherent Behavior in Non-cooperative Games." *Journal of Economic Theory*, 50(2): 424–44.

[24] For an English translation, please see Dreze (1987).

von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior.* Princeton: Princeton University Press.

O'Neill, Barry. 1994. "Game Theory Models of Peace and War." In *Handbook of Game Theory with Economic Applications*. Vol. 2, ed. R. J. Aumann and S. Hart, 999–1053. Amsterdam: Elsevier.

Ramsey, Frank P. 1931. "Truth and Probability." In *The Foundations of Mathematics*, ed. R. B. Braithwaite, 156–98. London: Routledge and Kegan Paul.

Rockafellar, R. Tyrrell. 1970. *Convex Analysis.* Princeton: Princeton University Press.

Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: John Wiley and Sons.