BA 513/STA 234: Ph.D. Seminar on Choice Theory
Professor Robert Nau
Spring Semester 2008

Readings for class #7: Beyond Nash equilibrium—correlated equilibrium, common knowledge, the common prior assumption, and joint coherence (revised February 22, 2008)

Primary readings:

1a. "Correlated equilibrium as an expression of Bayesian rationality" by Robert Aumann, *Econometrica*, 1987

1b. "Coherent behavior in noncooperative games" by Robert Nau and Kevin McCardle, *Journal of Economic Theory*, 1990

1c. "On the geometry of Nash equilibria and correlated equilibria" by Robert Nau, Sabrina Gomez Canovas, and Pierre Hansen, *Int. J. Game Theory*, 2004

Supplementary readings:

2a. "Non-cooperative games" by John Nash, *Annals of Mathematics*, 1951

2b. "Subjectivity and correlation in randomized strategies" by Robert Aumann, *Journal of Mathematical Economics*, 1974

2c. "Agreeing to disagree" by Robert Aumann, *Annals of Statistics*, 1976

2d. "The common prior assumption in economic theory" by Stephen Morris, *Economics and Philosophy*, 1995

Guide to the readings:

Last week we saw that there is a blooming, buzzing confusion of solution concepts for noncooperative games. Different solution concepts emphasize different lines of reasoning that "rational" game players might use, and they often lead to different conclusions about how a game should be played. For example, the players might first try to eliminate some strategies using strong or weak domination arguments; and if weak domination is used, the order of elimination may matter. If the game is in extensive form, they might try to solve it from back to front, using backward induction or subgame perfection as a guide. In other cases they might solve the game from front to back, emphasizing first-mover advantages or opportunities for signaling. Or they might convert an extensive form game to one of several normal forms before

proceeding. They might play randomized instead of pure strategies, with or without correlation. They might (or might not) worry about the possibility that other players will make "mistakes." They might (or might not) believe that two mistakes are less likely than one, or that costly mistakes are less likely than harmless ones. They might try to influence each others' behavior by making threats that would be painful to carry out, or they might regard such threats as lacking credibility. They might look for opportunities to coordinate on a mutually advantageous outcome. They might be content to seek a security level rather than shoot for the moon. And so on. It may seem, as Aumann suggested in his survey paper, that there is really no "correct" solution concept— there are just different "indicators" that one might choose to study. This week I will argue, nevertheless, that there IS a correct Bayesian solution concept for noncooperative games—and it is not Nash equilibrium or any of its refinements. Rather, it is a variant of Aumann's own concept of correlated equilibrium, although somewhat different from what Aumann himself proposed, and it leaves considerable room for subjectivity on the part of the players. It does not place strong *a priori* restrictions on how the players might choose to play a particular kind of game, but it does impose a strong *a posteriori* consistency condition on the strategies they may ultimately play in light of the knowledge and beliefs they have constructed. Furthermore, it dissolves the traditional distinctions that are made between individual, strategic, and competitive rationality: they are all seen to be different manifestations of a single rationality principle, namely *coherence* or *no-arbitrage*. To make the case for this position and trace the history of its development, I will begin by reviewing some issues concerning common knowledge and the common prior assumption.

Common knowledge and common priors

Game theory starts from the assumption the various facts about the game are "common knowledge." In order for a fact to be common knowledge, it is not enough that everyone should know it to be true: everyone must also know *that everyone knows it*, and that everyone knows that everyone knows it, and so on ad infinitum. The distinction between what everyone knows and what everyone-knows-that-everyone-knows is illustrated by the following famous puzzle:

> In an isolated village, there are 40 married couples. According to local custom, if a woman discovers that her husband has been unfaithful, she is required to denounce him in the public square at the precise stroke of noon on the very next day. Notwithstanding this proscription of adultery, every woman in the village is having affairs with the husbands of *all* the other women, while naïvely imagining that her own husband remains faithful to her. Things continue happily in this fashion for a long time. Then one day a missionary arrives. He quickly sizes up the situation and announces to the entire village: "there is a husband among you who has been unfaithful." Now, of course, the mere fact that there is infidelity in the village is not news to anyone, since everyone has been unfaithful with everyone else's spouse. And indeed, on the day after the missionary's announcement, nothing out of the ordinary happens. On the second day after the announcement, nothing unusual happens, and so it goes for more than a month. Then, on the 40th day after the announcement, the women all converge on the square at noon and denounce their husbands!

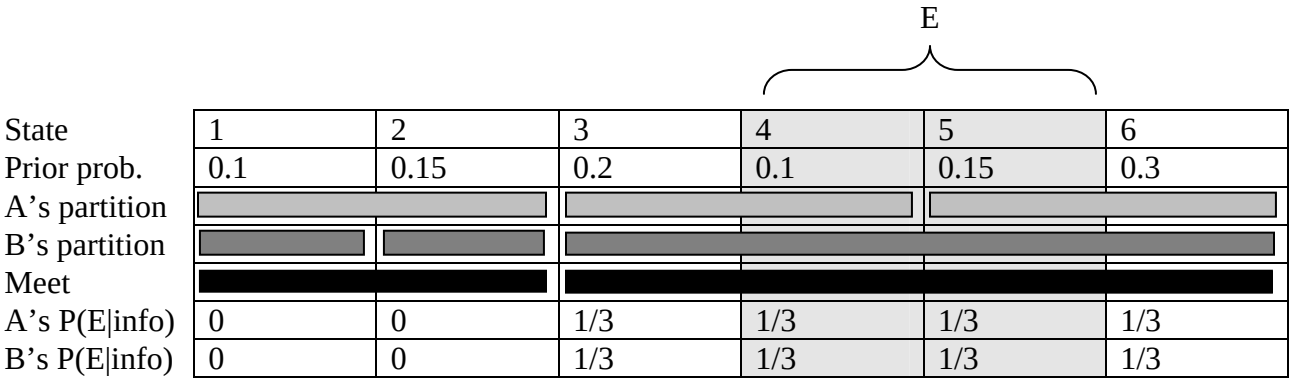
What has happened in this example? Before the missionary's visit, infidelity in the village is *mutual knowledge*—i.e., everyone knows it—but it is not *common knowledge*. The missionary's announcement makes the infidelity *common* knowledge, triggering a chain of reasoning that leads, only after 40 days, to the realization by every woman that her own husband has been unfaithful. The proof is by induction. Suppose that there are only two couples, Smith and Jones, and Mrs. Smith and Mrs. Jones have each had an affair with the other's husband. Consider Mrs. Smith's position after hearing the missionary's announcement. She knows that Mr. Jones has been unfaithful (with her), so she need not immediately suspect her own husband. However, she knows that IF her own husband has been unfaithful, Mrs. Jones would certainly know about it. On the hypothesis that her own husband is faithful, she expects Mrs. Jones to denounce her husband the very next day, knowing that the missionary's announcement could only have referred to him. But if Mrs. Jones fails to denounce her husband the very next day, then Mrs. Smith must realize that Mrs. Jones is in exactly the same position that she is in, and that BOTH husbands have been unfaithful. Hence Mrs. Smith (and also Mrs. Jones) will denounce her husband on the second day after the announcement. Next, suppose there are *three* couples: Smith, Jones, and Brown. Consider Mrs. Brown's position. On the hypothesis that her own husband is faithful, she expects Mrs. Smith and Mrs. Jones to both denounce their husbands on the second day after the missionary's announcement, based on the preceding analysis. If they do not, it can only mean that her own husband is also unfaithful. So she (and also Mrs. Smith and Mrs. Jones) will denounce her husband on the *third* day after the announcement. We can keep adding couples and repeating the same analysis, and each additional couple will add one day to the time required for everyone to fully grasp the situation.

A mathematical formalization of the concept of common knowledge was first given by the philosopher David Lewis in 1966. Then, in 1976, Aumann gave a simple set-theoretic characterization of common knowledge in his famous (4-page!) paper on "agreeing to disagree." Suppose that there is a set of states of the world, exactly one of which will occur, and that each individual has her own *information partition* that determines how much she knows about the state that occurs. In general, a *partition* of a set is a way of dividing it up into mutually exclusive and collectively exhaustive subsets. An *information* partition for an individual is a partition of the set of states of the world such that the individual knows which of the subsets contains the true state of nature, but is unable to discriminate among the states within that subset. (The concept of an information partition is a standard tool in information economics.) For example, suppose that there are six states, and that individual A has the information partition ({1, 2}, {3, 4}, {5, 6}). Each set of states in curly brackets is an element of her information partition—i.e., a set of states among which she is unable to discriminate. Thus, if state 1 or 2 occurs, she knows that one of these two states has occurred, but she doesn't know which one. Similarly, she is unable to discriminate between states 3 and 4, or between 5 and 6. Meanwhile, suppose that individual B has the information partition ({1}, {2}, {3, 4, 5, 6}). Then if state 1 or state 2 occurs, B knows exactly what has happened, while she is unable to discriminate among states 3 through 6.

The individuals' information partitions determine their private knowledge as a function of the state that occurs. But when is the occurrence of a state, or some subset of states, *common knowledge*? Aumann provides a simple answer to this question. An event is common knowledge if it includes an element of the *meet* of the information partitions that includes the true state. The meet of the different individuals' information partitions is the *finest common*

coarsening—i.e., the finest information partition that is coarser than any of the individual partitions. In the preceding example, the meet of A's and B's information partitions is ({1, 2}, {3, 4, 5, 6}). Hence, if state 1 or 2 occurs, then it is common knowledge that one of these two states occurred, although only B knows for sure which one. Similarly, if state 3, 4, 5, or 6 occurs, it is common knowledge that one of those four states has occurred, and A knows further whether it is 3 or 4 or whether it is 5 or 6, while B does not.

Aumann applies this definition of common knowledge to the following problem: suppose that two (or more) individuals start with a **common prior distribution** over states of nature, but they subsequently receive different information as determined by their respective information partitions. (Why should they have a common prior distribution? Good question! We will return to it later…) Each individual will update the common prior by applying Bayes' theorem, in order to obtain a personal posterior distribution. Moreover, since the information partitions are common knowledge, every agent knows what every other agent's posterior distribution will be (hypothetically) in every state of nature. Now, suppose it happens that the agents' posterior probabilities for some event are common knowledge in the state that actually occurs. What this means is that every agent's posterior probability for that event must be *constant* within the element of the meet of their information partitions that includes the true state. In the preceding example, suppose that the common prior distribution is (0.1, 0.15, 0.2, 0.1, 0.15, 0.3), and suppose that the event of interest is E = {4, 5}. (In other words, E is true if either state 4 or state 5 occurs.) Suppose that state 3 or 4 occurs. Then A knows (only) that either state 3 or 4 has occurred, in which case E is true if and only if state 4 has occurred. Her posterior probability for E is therefore the same as her posterior probability for state 4 at this point, namely 0.1/(0.2+0.1) = 1/3. On the other hand, suppose that state 5 or 6 occurs. Then A knows (only) that state 5 or 6 has occurred, in which case E is true if and only if state 5 has occurred. Her posterior probability for E is therefore the same as her posterior probability for state 5 at this point, namely 0.15/(0.15+0.3) = 1/3. So, A's posterior probability for E is the same regardless of whether states 3, 4, 5 or 6 occur. Meanwhile, if one of these four states occurs, B knows only that one of them has occurred, but nothing more. (She knows less than A in this case.) Her posterior probability for E is (0.1+0.15)/(0.2+0.1+0.15+0.3) = 1/3. Now, since {3, 4, 5, 6} is an element of the meet of the agents information partitions, and since each player's posterior probability is constant for all states with this element of the meet, it follows that *the players' posterior probabilities for E are common knowledge if states 3, 4, 5, or 6 occur*. Notice that their posterior probabilities also happen to be the *same* in this case. Aumann's theorem states that things must *always* turn out this way: if the players start from a common prior distribution, and if after receiving private information it happens that their posterior probabilities for some event are common knowledge, then those posterior probabilities must be the same. The agents cannot "agree to disagree"!



The result is mathematically trivial (as Aumann points out) once you see how the model is constructed, but nevertheless it is surprising and profound. It is impossible for private information to lead to divergent beliefs under conditions of common knowledge. Numerous other authors have extended this result and applied it to interactions between agents in markets. The typical result is a "no-trade theorem": agents who start with common prior distributions (as is often assumed in information economics) will never wish to engage in speculative trade based on differences in private information that they subsequently receive. As soon as it becomes common knowledge that they wish to trade, their expectations for the value of the asset in question must become identical. The prototypical result of this kind is in Milgrom and Stokey's 1981 paper on "Information, trade, and common knowledge." Some no-trade theorems do not invoke the common prior assumption, but instead assume merely that the prior allocation of state-contingent wealth is Pareto efficient. This is essentially equivalent to the assumption that agents have common prior *risk neutral* probabilities, which is a more plausible assumption to make and one that we will return to later.

The agreeing-to-disagree result depends on two key assumptions: (i) the information partitions are themselves common knowledge, and (ii) there is a common prior distribution on states of the world. Aumann argues that common knowledge of the information partitions is without loss of generality, because the states can always be defined in such a way as to make this assumption true. That is correct, but in order for it be nontrivial, there must be a considerable amount of agreement in the "small worlds" that different agents construct for themselves and attribute to other agents. Still, it is always possible to achieve a meaningful degree of common knowledge *by construction* by appealing to the existence of public events: one of the important functions of public markets and mass media is to make prices and other economic variables and noteworthy events commonly known. The common prior assumption is more problematic.

Harsanyi first introduced the common prior assumption (CPA) in his 1967 paper that presented the solution concept of *Bayesian Nash equilibrium* for games of incomplete information, in which the requirement that players exactly know each other's utility functions is relaxed by permitting them to be merely uncertain about each other's utilities. The introduction of incomplete information greatly extends the range of strategic situations that can be modeled by game theory—e.g., situations such as auctions, in which the players' uncertainty about each others' values is the central issue. But the increased generality of incomplete-information games comes at a price, namely that additional strong assumptions are needed to keep the uncertainty models tractable. First, it must be assumed that the set of possible utility functions for each player can be reduced to a manageable number of "types," which are themselves common

knowledge (like the information partitions in Aumann's agreeing-to-disagree model). Second, it is necessary to constrain the reciprocal beliefs that may exist with regard to such types. For example, if Alice and Bob are uncertain of each other's types, Alice's beliefs about what Bob believes about her type, given his own type, should be consistent with what Bob really believes, and so on in another infinite regress, otherwise it would be impossible for them to reason accurately about each other's intended behavior. The common prior assumption enforces exactly this sort of consistency. In Harsanyi's framework, there is a commonly-held prior distribution over types from which the actual probability distribution of each player concerning all the others' types can be derived by conditioning on her own type. A Bayesian Nash equilibrium is an equilibrium in which each *type* of each player chooses a pure or independently randomized strategy.

The following is an example of an incomplete-information game in which player 1 is uncertain about the type of player 2, which may be *A* or *B* (Myerson 1985, Nau 1992). Player 1 chooses between Top and Bottom and player 2 (whatever her type) chooses between Left and Right. However, the utility payoffs to both players now also depend on the type of player 2. It is assumed that there is a common prior distribution assigning probabilities of 60% and 40%, respectively, to types *A* and *B* of player 2. These are the probabilities that player 1 assigns to player 2's type, and player 2 knows that player 1 assigns, and player 1 knows that player 2 knows that player 1 assigns, and so on. Meanwhile, player 2 knows her own type with certainty at the instant she makes her move, and player 1 knows that she knows it, and so on. There are 8 possible outcomes of the game, and the payoff matrix is as follows:

Type *A* (60%)

	Left	Right
Top	1, 2	0, 1
Bottom	0, 4	1, 3

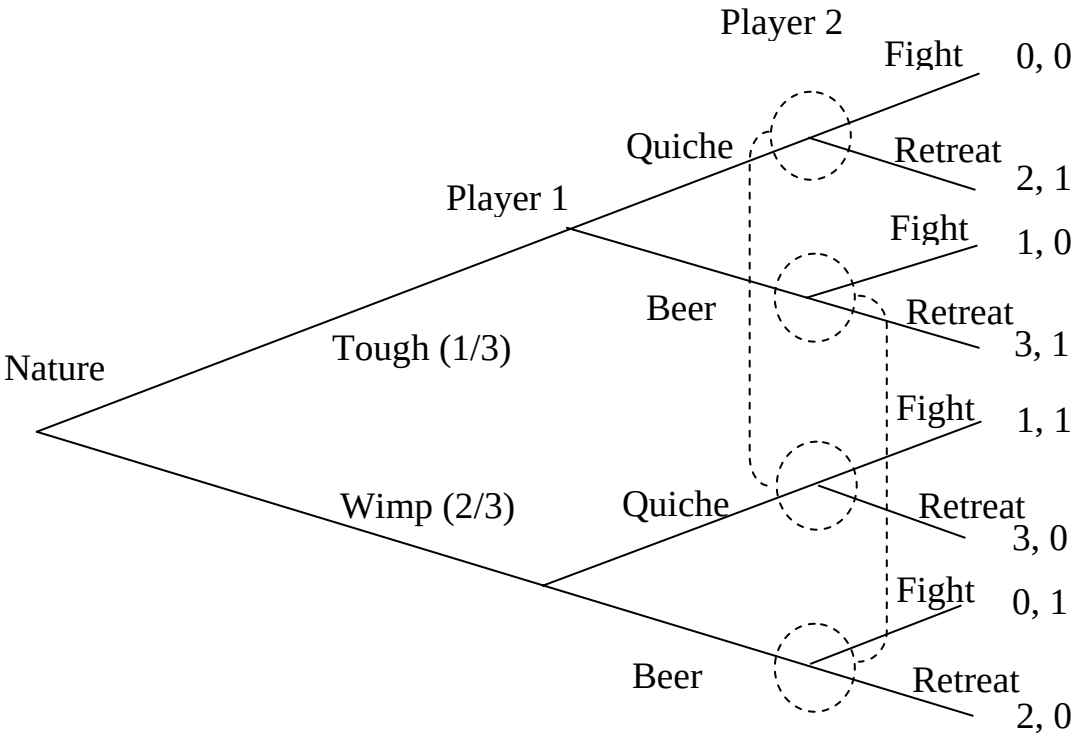
Type *B* (40%)

	Left	Right
Top	1, 3	0, 4
Bottom	0, 1	1, 2

This game has a unique Bayesian equilibrium in which player 1 always plays Top and player 2 plays Left if she is type *A* and Right if she is type *B*. Left is a dominant strategy for type *A* of player 2 and Right is a dominant strategy for type *B*. If the game were played under conditions of complete information, player 2's type would be known to player 1 before her move, and player 1 would play Top if 2's type were *A* (correctly anticipating that 2 would play Left) and 1 will play *Bottom* if 2 were type *B* (correctly anticipating a move of Right). But under incomplete information, player 2 benefits from player 1's uncertainty and is able to obtain a higher payoff when her type is *B*.

Here's another famous example due to David Kreps (also discussed in Binmore's *Fun and Games*). Player 1 walks into a bar and orders either beer or quiche. Player 2 is already there and decides to either fight or retreat from player 1. Player 1 has two types, Tough and Wimp. Tough prefers beer over quiche by 1 utile, while Wimp prefers quiche over beer by 1 utile. Both types of player 1 gain an additional 2 utiles if a fight is avoided. Player 2 is one utile better off by not fighting rather than fighting if player 1 is Tough, and she is one utile better off by fighting rather than not fighting if player 1 is Wimp. Unfortunately, player 2 has incomplete information: she cannot observe player 1's type, but only what he orders. However, the two players share a common prior probability of 1/3 for player 1 being Tough, i.e., player 2 believes that player 1 is

Tough with probability 1/3, and both types of player 1 know this, and player 2 knows that they know it, and so on and so on. The extensive form of the game therefore looks like this:



The dashes indicate the information sets of player 2, i.e., knowledge (only) of whether beer or quiche was ordered. Harsanyi's model translates this extensive form game into a corresponding strategic form in which a strategy of player 1 consists of a choice between quiche and beer for each of his possible types—as though he doesn't yet know his type and must make plans for all of them—and a strategy of player 2 consists of a choice between fight or retreat for each refreshment ordered by player 1. The entries in the strategic-form payoff matrix are *expected* utilities that take into account the common prior probabilities of the types of player 1:

	FF	FR	RF	RR
QQ	2/3, 2/3	2/3, 2/3	8/3, 1/3	8/3, 1/3
QB	0, 2/3	4/3, 0	2/3, 1	2, 1/3
BQ	1, 2/3	5/3, 1	7/3, 0	3, 1/3
BB	1/3, 2/3	7/3, 1/3	1/3, 2/3	7/3, 1/3

Here, the notation QQ means the strategy of choosing Q if Tough and Q if Wimp for player 1, QB means Q if Tough and B if Wimp, etc.; and FF means F if Quiche and F if Beer, FR means F if Quiche and R if Beer, etc. This game has a unique Bayesian equilibrium in which player 1 plays (0, 0, ½, ½) and player 2 plays (½, ½, 0, 0) in the strategic form. That is, the tough guy always has beer, and the wimp has beer with probability ½. Player 2 always fights if player 1 has quiche and fights with probability ½ if player 1 has beer.

The CPA is widely used in information economics because it provides a tractable way to model situations in which agents are differently informed about the true state of nature and yet hold beliefs that are in some sense mutually consistent, but almost no one claims to believe in it. For example, Kreps is quite explicit in rejecting it, and Faruk Gul attacked it in a note in *Econometrica* ("A comment on Aumann's Bayesian view", July 1998), prompting a response from Aumann in the same issue. The standard defense of the CPA is that beliefs ought to be based on information, leading to a thought experiment in which agents are fed precisely the same information—and therefore construct precisely the same beliefs—up to the moment when they observe their private information. But this is patently a fiction: there was never a primordial moment in time when everyone was in exactly the same state of information. Everyone receives different sensory data and filters it through his or her own brain from the moment of birth, and everyone's brain is wired somewhat differently. Why should they form exactly the same beliefs for events that lack objective probabilities? Gul objects to the element of counterfactualism inherent in the thought experiment and presents a simple example of "reasonable" reciprocal beliefs that violate the CPA.

In Gul's example, there are two players who have infinite hierarchies of reciprocal beliefs about each others' types that are superficially consistent with each other but are nevertheless inconsistent with any common prior distribution—in fact, they can be represented by an information model with different priors. Here are the details:

> Player 1 has two types, X and Y
> Player 2 has two types, A and B.
>
> Type A of player 2 believes p(X)=0.4, p(Y)=0.6
> Type B of player 2 believes p(X)=0.5, p(Y)=0.5

…and player 1 knows this, and player 2 knows that player 1 knows it, etc. etc. Meanwhile,

> Player 1 believes p(A) = p(B) = 0.5

…and player 2 knows this, and player 1 knows that player 2 knows it, etc. etc.

The true state of the world happens to be (X, A), i.e., player 1 is really type X, and player 2 assigns probability 0.4 to the event that player 2 is type X. 1. Gul points out that the prior distributions that could potentially represent the beliefs of the two players are only partially determined:

Player 1

	A	B
X	0.5α	0.5α
Y	0.5β	0.5β

Player 2

	A	B
X	0.4δ	0.5γ
Y	0.6δ	0.5γ

for any positive parameters $\alpha+\beta = 1$, $\delta+\gamma=1$. However, there is no *common* prior distribution that represents these beliefs, i.e., no values of α, β, δ, γ for which the two joint distributions

agree, because X and A are independent for player 1 but not independent for player 2. And anyway, there never was a "prior" stage at which player 1 was not type X and player 1 was not type A.

In his response to Gul, Aumann points out that counterfactual reasoning and far-fetched thought experiments are ubiquitous in decision theory (e.g., they were used by Savage and Nash). Aumann presents an axiom that implies the CPA by requiring that if the players could somehow "forget" their private information, they would have a common probability distribution on the state space. This axiom essentially assumes the conclusion, but at least it formalizes the intuitive argument behind the CPA. Aumann's response to Gul's counterexample is "so what?" Anyone who wants to endow the actors in their models with beliefs that are heterogeneous in a way that cannot be explained merely by differential information is free to do so, but to do so is to place the actors in a disequilibrium situation from which it might well be profitable for them to depart as soon as possible. Or as I will show below, *if the agents in Gul's example were willing to bet according to their beliefs, they would immediately be vulnerable to arbitrage*. But the most compelling argument against the CPA, it seems to me, is the intrinsic indeterminacy of subjective probabilities due to incompleteness of beliefs and the difficulties of separating probabilities from utilities. If players cannot uniquely infer each other's true probabilities from preferences, or if they do not have well-defined subjective probabilities to begin with, why should their true probabilities agree? We will see later that this argument too can rebutted, but it will require a reinterpretation of the prior "probabilities" that are commonly held.

Correlated equilibrium

In his 1974 paper on "Subjectivity and correlation in randomized strategies" (included in this week's supplementary readings), Aumann first proposed the concept of a *correlated equilibrium*: a generalization of Nash equilibrium in which randomized strategies are permitted to be correlated. There are two flavors of correlated equilibrium: an "objective" correlated equilibrium is one in which the players agree on the probabilities of any random events to which their strategies may be pegged. A "subjective" correlated equilibrium is one in which they do not necessarily agree on the probabilities. Aumann showed that there are many games in which it is mutually advantageous to the players to use objectively or subjectively correlated strategies. For example, in zero-sum games, the sum of the players' personal *expected* payoffs may be non-zero if they are able to peg their strategies to random events about whose probabilities they disagree, and of course this is precisely why betting markets exist.

An objective correlated equilibrium can be implemented by a mediator who employs a randomization device to generate "recommended" strategies for all the players according to an agreed-upon joint probability distribution. The mediator informs each player of her own recommended strategy generated by the device, but not those of the other players. If the joint distribution is a correlated equilibrium distribution, then by definition it is optimal for the players to adhere to their recommended strategies rather than unilaterally defecting to any other strategies. A correlated equilibrium distribution has the property that when each player updates the common joint distribution based on knowledge of her own recommended strategy, she maximizes her expected payoff by following the recommendation, assuming that all other players play their own recommended strategies, which it is optimal for them to do to. (This is

exactly the same reasoning that supports a Nash equilibrium. The only difference is that here the common joint distribution need not be a product of independent distributions.) It is *not* possible to implement a *subjective* correlated equilibrium in this way. Because of the impossibility of agreeing to disagree under conditions of common knowledge, it is impossible to construct a device whose design is agreed-upon and which produces random strategy recommendations whose probabilities are disagreed-upon. Thus, to implement a subjective correlated equilibrium, it is seemingly necessary to find some "natural" events about whose probabilities the players stubbornly disagree.

In his 1987 paper "Correlated equilibrium as an expression of Bayesian rationality," Aumann argues that objective correlated equilibrium is the natural "Bayesian" solution concept for noncooperative games. The proof of this claim relies on a construction similar to the one used in the agreeing-to-disagree paper. It is assumed that there is a set of states of the world of which the players hold commonly-known private information partitions, and it is assumed that they have a common prior distribution on the states. Each player then chooses a strategy function that maps the elements of her information partition to pure strategies in the game, and it is assumed that every player knows the other players' strategy functions. Aumann shows that if every player is "Bayesian rational in every state of the world" in the sense that it is optimal to adhere to her own strategy function on the assumption that everyone else does likewise, regardless of which state occurs, then the induced probability distribution on outcomes of the game must be an objective correlated equilibrium distribution. The persuasiveness of this argument obviously depends on your attitude toward the common prior assumption. Since many economists and decision theorists are uncomfortable with the CPA, they have (for the most part) not rallied to the standard of correlated equilibrium as the "correct" solution concept for games. Interestingly, many theorists appear to mistakenly believe that Nash equilibrium somehow does not "require" the CPA, even though it is strictly stronger than correlated equilibrium. But Aumann's argument can be interpreted to show that the common prior assumption—in addition to the gratuitous assumption of probabilistic independence—is already implicit in the Nash equilibrium concept.

A very attractive property of correlated equilibria—besides the fact that in some games they yield better payoffs than Nash equilibria—is that they are extremely easy to compute. A correlated equilibrium distribution is defined by a system of linear inequalities that can be solved by linear programming. The simplest case is a 2×2 game. Let the payoff matrix of the game be written as follows, where player 1 chooses the row (Top or Bottom) and player 2 chooses the column (Left or Right):

	Left	Right
Top	a, a'	b, b'
Bottom	c, c'	d, d'

Here a and a' are the utility payoffs to players 1 and 2 when Top-Left is played, and so on. Now let $\boldsymbol{\pi} = (\pi_{TL}, \pi_{TR}, \pi_{BL}, \pi_{BR})$ denote a probability distribution on the four outcomes of the game, and suppose this distribution is used to randomly generate a recommended strategy to each player. Then π_{TL} will be the joint probability that player 1's recommended strategy is Top and player 2's recommended strategy is Left, and similarly for the other outcomes. Assume that each player hears only her *own* recommendation: when the joint recommendation is Top-Left, player 1

knows *only* that her own recommendation is Top, and player 2 knows *only* that her own recommendation is Left. However, if $\boldsymbol{\pi}$ is commonly known, each player is able to use Bayesian updating to compute the probabilities of the recommendations that her opponent has received, given knowledge of her own recommendation. Thus, for example, when player 1 receives a recommendation of Top, she computes the conditional probability of player 2 having received a recommendation of Left to be $\pi_{TL}/(\pi_{TL} + \pi_{TR})$.

In order for $\boldsymbol{\pi}$ to be a correlated equilibrium distribution, it must have the property that, conditional on player 1 receiving a recommendation of Top, she should do at least as well by playing Top as by playing bottom, assuming that player 2 adheres to her own recommended strategy. This means that:

$$\pi_{TL}a + \pi_{TR}b \;\geq\; \pi_{TL}c + \pi_{TR}d$$

The left-hand side of this inequality is the contribution to *a priori* expected utility that player 1 would receive by playing Top whenever Top is recommended, and the right-hand side is the contribution to expected utility she would get by playing Bottom whenever Top was recommended. Equivalently, we can write:

$$\pi_{TL}(a - c) + \pi_{TR}(b - d) \geq 0$$

Similarly, player 1 should have no incentive to play Top when Bottom is recommended, either, which means:

$$\pi_{BL}(c - a) + \pi_{BR}(d - b) \geq 0$$

Continuing in the same fashion, player 2 should have no incentive to play Right when Left is recommended, and vice versa, leading to the additional inequalities:

$$\pi_{TL}(a' - b') + \pi_{BL}(c' - d') \geq 0$$

$$\pi_{TR}(b' - a') + \pi_{BR}(d' - c') \geq 0$$

These four *incentive constraint*s, together with the total-probability and non-negativity constraints, define the set of ALL correlated equilibrium distributions of a generic 2×2 game. If the coefficients in the incentive constraints are arranged in matrix form, they look like this:

	TL	TR	BL	BR
1TB	$a - c$	$b - d$		
1BT			$c - a$	$d - b$
2LR	$a' - b'$		$c' - d'$	
2RL		$b' - a'$		$d' - c'$

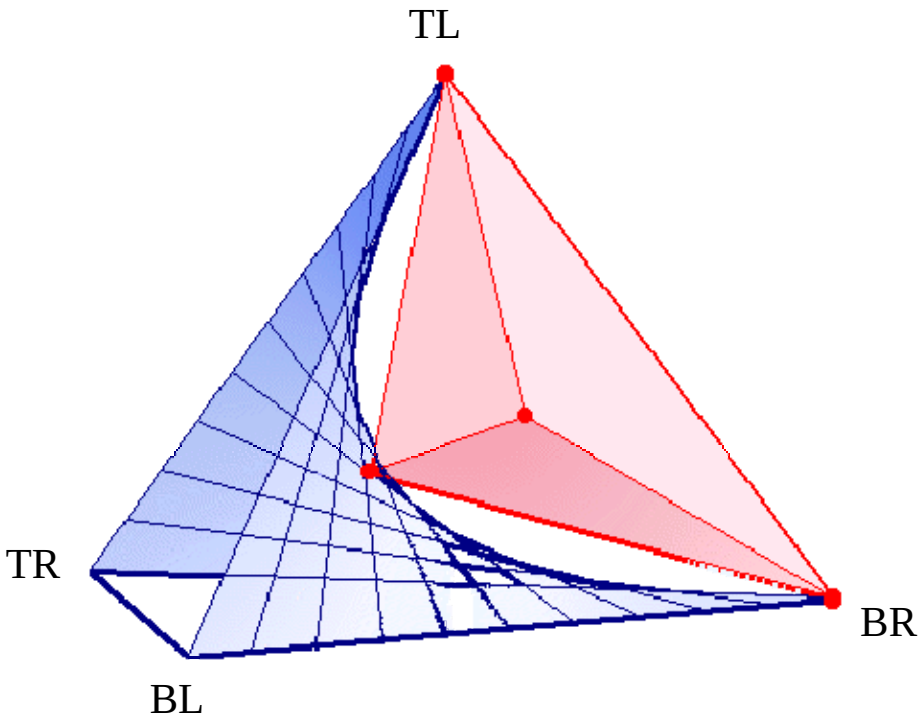
The notation 1TB next to the first row indicates that this row represents the difference in utility for player 1 between adhering to T and defecting to B, given that T has been recommended, and similarly for the other rows. Blank cells are interpreted as zeroes. If we let *G* denote this matrix,

then the system of incentive constraints that define a correlated equilibrium distribution can be written simply as $G\boldsymbol{\pi} \geq \mathbf{0}$. G is the same "rules of the game" matrix introduced in the previous lecture, and we now see that the single linear inequality $G\boldsymbol{\pi} \geq \mathbf{0}$ determines all the rational joint strategies according to Aumann's argument. Voila!

In general, for games of any size, there is an incentive constraint for every combination of a recommended strategy and an alternative strategy for every player. Since it is thus defined by a system of linear inequality constraints, *the set of correlated equilibrium distributions is a convex polytope.* (A polytope is a bounded polyhedron.) If the payoff matrix consists of rational numbers, then the extreme points of the polytope must have rational coordinates, and they are the "corner point solutions" that will be obtained if linear programming is used to solve the system of inequalities while maximizing or minimizing some linear function of the probabilities. A natural objective function to maximize is a weighted sum of the expected utilities of the players: this will yield an equilibrium on the Pareto frontier of the polytope. Now, if one is going to use objective correlated equilibrium as a solution concept, it seems reasonable that one would want to restrict attention to the Pareto efficient equilibria, which are not necessarily Nash equilibria as Aumann shows in some of his examples. However, the proper interpretation of the correlated equilibrium polytope is that it represents the set of all possible rational joint beliefs that might be held by a group of players who have opportunities to communicate but who ultimately will play the game on noncooperative terms. Equivalently, it is the set of beliefs about the outcome of the game that might be held by an outside observer who believes the players to be Bayesian rational. In summary:

> *The inequalities defining the correlated equilibrium polytope are the natural constraints that "mutually expected Bayesian rationality" places on the beliefs that ought to be held with regard to the outcome of a noncooperative game, either by the players themselves or by an outside observer, assuming that (a) the rules of the game (i.e., the utility functions of the players) are already common knowledge, and (b) the players have common prior probabilities with respect to any events to which their strategies may be pegged.*

Since Nash equilibria are a special case of correlated equilibria, the polytope of correlated equilibrium distributions must contain all the Nash equilibrium distributions. Until very recently, it seems that no one had given much though to the question of *where* in the polytope the Nash equilibria might lie, although it is well known that in games with more than two players the set of Nash equilibria can be disconnected, irrational, hard to find, and quite ugly. In my paper (with Sabrina Gomez Canovas and Pierre Hansen at the University of Montreal) called "On the Geometry of Nash Equilibria and Correlated Equilibria" (*Int. J. Game Theory*, August 2004) we show that in any finite nontrivial game with any number of players, the Nash equilibria in general all lie on the *boundary* of the polytope—i.e., on one of its "frontiers," but not always the Pareto efficient frontier. Indeed, Nash equilibria which are completely mixed (i.e., which assign positive probability to every strategy of every player) are usually on the *inefficient* frontier. This analysis leads to a very nice picture of the relationship between Nash and correlated equilibria in 2×2 games, which is reproduced below:



This picture shows the sets of correlated and Nash equilibria of the game "battle of the sexes." (Doesn't it look like a battle of the sexes?) The tetrahedron is the set of *all* probability distributions on joint strategies; the saddle is the set of *independent* distributions; the shaded polytope is the set of *correlated* equilibria; and their three points of intersection are *Nash* equilibria. While this picture refers to a particular 2×2 game, it is *generic* in the sense that any non-degenerate 2×2 game (i.e., a game in which both players have strategies that lead to different payoffs) must have either one or three distinct Nash equilibria. If it has only one Nash equilibrium, then this is also the only correlated equilibrium—i.e., the polytope collapses to a point, which could be anywhere on the saddle. If the game has three distinct Nash equilibria, they must be in the generic positions shown here—two at vertices of the tetrahedron and one in the middle of the saddle—and the correlated equilibrium polytope then typically has five vertices in the configuration shown. There is one other vertex of the correlated equilibrium polytope that you can't see in this picture: it's on the backside.

Here's example of a game that illustrates the potentially strange geometry of Nash equilibria and the difficulty of enumerating them. It is a 2×2×2 game in which player 1 chooses Top or Bottom, player 2 chooses Left or Right, and player 3 chooses Up or Down:

Up

	Left	Right
Top	0, 0 ,2	0, 3 ,0
Bottom	3, 0, 0	0, 0, 0

Down

Top	1, 1, 0	0, 0, 0
Bottom	0, 0, 0	0, 0, 3

The correlated equilibrium polytope is seven-dimensional (i.e. full dimension) and has 8 vertices. Three of the vertices are pure Nash equilibria (TRU, BLU, and BRD). There are also two incompletely mixed Nash equilibria: {TR, ¼ U + ¾ D} and {BL, ¼ U + ¾ D}. Finally, and most curiously, there is a continuum of completely mixed Nash equilibria lying along an open curve (in the middle of a face of the polytope!) connecting the two incompletely mixed equilibria. This is a very simple game (integer payoffs, only 8 outcomes), but its set of Nash equilibria is so weird that Gambit, the public-domain game-solving program, cannot find them all. The number of vertices of the correlated equilibrium polytope is usually quite small relative to the number of outcomes of the game, although there are exceptions. My colleagues and I used a vertex-enumeration algorithm to find all the vertices of the correlated equilibrium polytopes of 250 randomly-generated 4×4 games, as reported in our 2004 paper in *IJGT*. These games had 16 outcomes, but their correlated equilibrium polytopes had 5 or fewer vertices in more than half the cases, and only a single Pareto efficient vertex in over three-quarters of the cases. The unique efficient vertices in these latter games were all Nash equilibria, but not all the Nash equilibria of these games were efficient. However…. in 4 of the 250 examples the polytopes had more than 100,000 vertices. Life can be complicated in more than three dimensions! The take-away from these examples is that except in very simple games, it is a lot to expect of ordinary people to figure out all the rational ways to play.

Coherent decision-making under uncertainty

Aumann's proof that correlated equilibrium is the natural Bayesian solution concept is simple and elegant, but it still requires several strong assumptions that go beyond Bayesian rationality per se. It assumes that Bayesian rationality is common knowledge, and that the utility functions of the players (the "rules of the game") are common knowledge, and that the subjective probabilities of the players are consistent with a common prior. To dig a deeper foundation under these assumptions, let us return to money-based measurements and the no-arbitrage standard of rationality previously introduced in the context of the elicitation and aggregation of coherent beliefs. To set the stage for the solution of a game by these methods, first consider the simpler problem of a *decision under uncertainty*: a situation in which some events are states of nature, all other events are alternatives for a single agent, and the problem is for that agent to rationally choose among those alternatives. Suppose that the event under the control of an agent is whether she carries an umbrella, the event not under her control is whether it rains, and she makes the following statement:

> *Alice:* "I'd carry my umbrella today only if I thought the chance of rain was at least 50%, and I'd leave it at home only if I thought the chance was no more than 50%."

In other words, Alice is willing to accept either or both of the following bets:

Table 1

	Umbrella, Rain	Umbrella, No rain	No umbrella, Rain	No umbrella, No rain
Payoff to Alice for bet #1	\$1	-\$1	\$0	\$0
Payoff to Alice for bet #2	\$0	\$0	-\$1	\$1

Note that these bets depend on Alice's choice (only) as a conditioning event: given that she ends up carrying her umbrella, she will bet *on* the occurrence of rain at 50:50 odds, and given that she ends up not carrying her umbrella, she will bet *against* the occurrence of rain at 50:50 odds. Bets of this kind do *not* reveal any information about Alice's (apparent) *probabilities* for states of nature. Rather, they reveal information about Alice's (apparent) *utilities* for outcomes of events. Assuming that money is equally valuable in all outcomes,[1] it is as if her utility function for outcomes has the following form:

Table 2

	Rain	No Rain
Umbrella	1	-1
No Umbrella	-1	1

To see the correspondence between the acceptable bets in Table 1 and the utility function in Table 2, note that if Alice has constant marginal utility for money, the acceptable bets imply that she will carry her umbrella or not according to whether she thinks the probability of rain is greater or less than 50%, which is exactly the same behavior that is implied by the utility function. Table 2 has the same interpretation as the payoff matrix of a game player: the entries in the cells are utility values. As with any utility functions, the values are not uniquely determined: the origin and scale are arbitrary, and an arbitrary constant may also be added to each column without affecting comparisons of expected utility between alternatives. Thus, for example, an equivalent utility function that perhaps better represents the consequences of getting caught in the rain might be:

Table 3

	Rain	No Rain
Umbrella	0	1
No Umbrella	-2	3

Of course we do not assume, *a priori*, that Alice is an expected-utility maximizer. She may have any reasons whatever for accepting the bets summarized in Table 1. The implied utility functions of Tables 2 & 3 are merely suggestive interpretations of her behavior. But it is worth noting that if Alice *does* think in the fashion of expected-utility analysis, and if she doesn't mind who knows her utilities, then it is *in her own interest* to accept the bets of Table 1. Regardless of the probability that she may assign to the event of rain prior to making her choice, the bets in Table 1 cannot decrease her total expected utility. Indeed, unless she ends up perfectly indifferent between carrying the umbrella or not, they can only strictly increase her total expected utility. The defining quality of the bets in Table 1, under an expected-utility interpretation, is that they amplify whatever differences in expected utility Alice perceives between her two alternatives. For example, if she later concludes that carrying the umbrella has higher expected utility than not carrying it because her probability of rain turns out to be greater than 50%, then she will carry the umbrella and bet #1 will be in force, and bet #1 yields positive marginal utility precisely in the case that her probability of rain is greater than 50%.

[1] More precisely, assume that Alice behaves as if she has *quasilinear utility*, i.e., constant linear utility for money on top of whatever direct utility she derives from the outcome of her situation.

Now suppose that at the same time and place that Alice is making her statement, a second agent is saying the following:

Bob: "I think the chance of rain is at least 75% whether or not you carry your umbrella!"

In other words, Bob is willing to accept either or both of the following bets:

	Umbrella, Rain	Umbrella, No rain	No umbrella, Rain	No umbrella, No rain
Payoff to Bob for bet #1	\$1	-\$3	\$0	\$0
Payoff to Bob for bet #2	\$0	\$0	\$1	-\$3

This statement reveals information about Bob's beliefs: assuming that money is equally valuable to him whether or not it rains, his probability of rain is evidently at least 75%, and furthermore *he does not regard Alice's behavior as informative*. Perhaps Bob is the local weatherman.

We are now in a position to predict (or perhaps prescribe) what Alice should do, namely: *she should carry the umbrella*, because otherwise there is ex post arbitrage. If an observer takes bet #2 with Alice (scaling it up by a factor of two) and bet #2 with Bob, the result is as follows:

	Umbrella, Rain	Umbrella, No rain	No umbrella, Rain	No umbrella, No rain
Payoff to Alice for bet #2 (x2)	\$0	\$0	-\$2	\$2
Payoff to Bob for bet #2	\$0	\$0	\$1	-\$3
Total	\$0	\$0	-\$1	-\$1

Hence, the observer earns a riskless profit in the event that Alice fails to carry her umbrella.

As in the oil-drilling example discussed in class 2, the question might be asked: why should Alice's behavior be constrained in any way by Bob's beliefs? Perhaps she believes that the probability of rain is less than 50% *even after hearing Bob's statement*, in which case, based on the utility function inferred from her own previous testimony, she should not carry the umbrella. But if this is so, she ought to hedge her risks by betting with Bob, and she should keep betting with him and raising the stakes until someone adjusts his or her betting rate to eliminate the arbitrage opportunity in the "no umbrella" event. (And the same for Bob with respect to Alice.) If, instead, each agent hears the other's statement but does not respond to it, they evidently are both certain that Alice will carry the umbrella and the observer will not end up getting a riskless profit. Of course, to an observer, it doesn't matter whether one agent or two is speaking. Bob's statement can just as well be made by Alice in the instant before she decides whether to carry the umbrella, in which case the conclusion is the same: she should take the umbrella.[2]

[2] Interestingly, if Bob's statement about the chance of rain is made *unconditionally*—i.e., if he merely says "I think the chance of rain is at least 75%"—then there is no arbitrage opportunity. Any bet with Bob has a positive payoff for him in the outcome {Umbrella, Rain}, but no bet with Alice has a negative payoff for her in the same outcome that could be used to hedge the observer's risk. In such a scenario, it is possible that Alice's eventual decision as to

Note that decision analysis was carried out in a novel fashion in this example: rather than asking the agent to articulate a probability distribution and utility function and then advising her to choose the alternative with the highest expected utility, we asked the agent—and those around her!—to articulate the *gambles* they were willing to accept and then advised her to choose the alternative that avoided ex post arbitrage. Of course, there is a theorem that says the two approaches are equivalent, with one important exception: the latter method does not require the explicit separation of probability from utility, and it does not assume that the agent's utilities are state-independent. More details are given in my paper "Coherent Decision Analysis with Inseparable Probabilities and Utilities"(*J. Risk and Uncertainty* 1995).

Joint coherence

A famous passage in von Neumann and Morgenstern's book (1944, p. 11) states that "the problem of 2, 3, 4,... bodies" of game theory is qualitatively different from—and harder than—the single-body problem of decision theory:

> "Thus each participant attempts to maximize a function... of which he does not control all variables. This is certainly no maximum problem but a disconcerting mixture of several conflicting maximum problems. Every participant is guided by another principle and neither determines all variables which affect his interest. This kind of problem is nowhere dealt with in classical mathematics."

The next example illustrates that it isn't necessarily so. The passage from a single-agent game against nature to a multiple-agent game of strategy entails no additional modeling assumptions or rationality concepts.[3]

> *Alice:* "I'd carry my umbrella only if I thought there was at least a 50% chance that Bob would dump a bucket of water out the window as I walked by, and I wouldn't carry it only if I thought the chance was less than 50%."

Bob is now cast in the role of rainmaker rather than weatherman, and Alice's implied utility function is the same as before (Tables 2&3), with the event label "Rain" merely replaced by "Dump." But suppose that Bob (unlike nature) has a malicious interest in getting Alice wet, as indicated by the following claim:

> *Bob*: "I would dump a bucket of water out the window as Alice walked by only if I thought there was at least a 50% chance she wasn't carrying her umbrella, and I wouldn't dump it only if I thought otherwise."

---

whether to carry the umbrella will be based on the receipt of private information previously unknown to Bob, and Bob is tacitly admitting this possibility by failing to condition his statement on Alice's behavior.

[3] Aumann (1987) also observes that is is possible to "[do] away with the dichotomy usually perceived between the 'Bayesian' and the 'game-theoretic' view of the world." Our demonstration of this fact is, roughly speaking, the dual of Aumann's, but it does not require the assumption of information partitions on a larger set of events nor the assumption of a common prior distribution, and it also does away with the other perceived dichotomy between strategic and competitive behavior.

In other words, Bob will accept the following bets:

	Umbrella, Dump	Umbrella, No dump	No umbrella, Dump	No umbrella, No dump
Payoff to Bob for bet #1	-$1	$0	$1	$0
Payoff to Bob for bet #2	$0	$1	$0	-$1

Bob is now revealing information about *his* relative utilities for outcomes, not his beliefs. It is as if his utility function is of the form:

	Dump	No dump
Umbrella	-1	1
No umbrella	1	-1

…because (only) someone with a utility function equivalent to this one would accept the bets that Bob has accepted (assuming constant marginal utility for money). Putting Alice's and Bob's apparent utility functions together, we find it is as if they are players in a noncooperative game with the payoff matrix:

	Dump	No dump
Umbrella	1, -1	-1, 1
No umbrella	-1, 1	1, -1

where the numbers in the cells are the utilities for Alice and Bob respectively. This is the game of "matching pennies," and it has a unique Nash equilibrium in which both players randomly choose among their two alternatives with equal probabilities.

What prediction of the outcome of the game can be made by arbitrage arguments? The set of all acceptable gambles is now as follows:

Table 4

	Umbrella, Dump	Umbrella, No dump	No umbrella, Dump	No umbrella, No dump
Payoff to Alice for bet #1	$1	-$1	$0	$0
Payoff to Alice for bet #2	$0	$0	-$1	$1
Payoff to Bob for bet #1	-$1	$0	$1	$0
Payoff to Bob for bet #2	$0	$1	$0	-$1

As it happens, ex post arbitrage is not possible in any outcome, so Alice and Bob may do whatever they please. However, from an observer's perspective, Alice and Bob appear to believe that they are implementing the *Nash equilibrium* solution. That is, they appear to believe that all four outcomes are equally likely. To see this, note that the observer can rescale and add up the gambles in the following way:

	Umbrella, Dump	Umbrella, No dump	No umbrella, Dump	No umbrella, No dump
Payoff to Alice for bet #1 (×4)	\$4	-\$4	\$0	\$0
Payoff to Alice for bet #2 (×2)	\$0	\$0	-\$2	\$2
Payoff to Bob for bet #1 (×1)	-\$1	\$0	\$1	\$0
Payoff to Bob for bet #2 (×3)	\$0	\$3	\$0	-\$3
Total	\$3	-\$1	-\$1	-\$1

The bottom line is equivalent to betting on the outcome {Umbrella, Dump} at odds of 1:3 in favor—i.e., betting as if the probability of this outcome is *at least* 25%. Alternatively, the gambles can be combined this way:

	Umbrella, Dump	Umbrella, No dump	No umbrella, Dump	No umbrella, No dump
Payoff to Alice for bet #1 (×0)	\$0	\$0	\$0	\$0
Payoff to Alice for bet #2 (×2)	\$0	\$0	-\$2	\$2
Payoff to Bob for bet #1 (×3)	-\$3	\$0	\$3	\$0
Payoff to Bob for bet #2 (×1)	\$0	\$1	\$0	-\$1
Total	-\$3	\$1	\$1	\$1

which is equivalent to betting as if the probability of {Umbrella, Dump} is *no more than* 25%. Hence, between them, Alice and Bob appear to believe the probability of {Umbrella, Dump} is *exactly* 25%—and of course by symmetry the same trick can be played with all the other outcomes.

What is remarkable about this example is that neither Alice nor Bob has revealed any information whatever about his or her beliefs: they have merely revealed information about their *utilities* via appropriate gambles. Yet this turns out to be operationally equivalent to asserting beliefs that correspond to a Nash equilibrium. Why did this happen? Of course there is another theorem lurking around, and it is just a generalization of our earlier no-arbitrage theorems to the case of a game of strategy. By the subjective probability theorem, we know that the players behave rationally (avoid ex post arbitrage) if and only if there is a supporting probability distribution that assigns non-negative expected value to every gamble accepted by every player and assigns strictly positive probability to the event that occurs. The supporting probability distribution can be interpreted to represent the commonly-held beliefs of the agents—i.e., the beliefs of a representative agent—notwithstanding the distortions that may be introduced by state-dependent marginal utility for money. Now, if the situation happens to be a game of strategy, and if the players have constant marginal utility for money, and if they accept gambles which reveal their relative utilities for outcomes of the game in the manner illustrated above, then the supporting probability distribution must be an *objective correlated equilibrium* of the game defined by those utilities. The game between Alice and Bob, which is strategically equivalent to matching pennies, happens to have a unique correlated equilibrium, which is also the unique Nash equilibrium. Hence, as soon as the rules of the game are revealed through acceptable bets, the players' apparent beliefs about its outcome are uniquely determined.

The connection between objective correlated equilibrium and no-ex-post-arbitrage was shown in my paper (with Kevin McCardle) on "Coherent Behavior in Noncooperative Games" (*JET* 1990) and extended to incomplete-information games in a later paper ("Joint Coherence in Games of Incomplete Information, *Management Science* 1992) . The outline of the proof will be sketched here. Recall that it is possible find correlated equilibria of a game by solving a simple linear programming problem to obtain a non-negative solution to the system of inequalities $G\pi \geq 0$, where G is the matrix of coefficients of the incentive constraints. (You can do this with Solver in Excel.) The second dialogue between Alice and Bob suggests another interpretation for the matrix G. If the players have constant marginal utility for money—i.e., they are risk neutral toward gambles for money—then each row in the matrix can be interpreted as the payoff vector of a *monetary gamble* that they would be willing to accept. For, suppose that at the instant she is to move, player 1 (e.g., Alice) assigns probabilities π_L and π_R to the events that her opponent will play L and R, respectively. Then she will play Top only if

$$\pi_L (a - c) + \pi_R(b - d) \geq 0.$$

But this same inequality implies that a monetary gamble on Left-vs-Right with payoffs $(a - c)$ and $(b - d)$, respectively, will yield a non-negative increment of expected utility. Hence, *in the event that player 1 chooses to play Top*, she ought to be willing to accept a gamble whose payoffs are payoffs are $(a - c)$ and $(b - d)$ when her opponent plays Left and Right, respectively. In effect, accepting this gamble just amplifies whatever difference in expected utility she perceives between Top and Bottom, in the event that the difference is positive. The first row of the matrix G above can be interpreted as the payoff vector of a *conditional* gamble that yields these two payoffs *given* the event that player 1 plays Top. The gamble is "called off"—i.e., zeroed out—if Bottom is played instead. Hence, if the rules of the game are really common knowledge, and if player 1 has constant marginal utility for money, then it must also be common knowledge that she will accept this gamble. Indeed, she ought to be willing to accept *an arbitrary non-negative multiple* of this gamble. Similarly, the other three rows of the matrix correspond to other conditional gambles that the players ought to be willing to accept in arbitrary non-negative multiples. The matrix shown in Table 4 was constructed in this fashion.

We can now turn this line of reasoning on its head and say that **by accepting monetary gambles that mirror their stakes in the game, the players reveal the matrix *G* that encodes the rules of the game, and thus the rules become common knowledge**. As we have already seen, this matrix determines the set of correlated equilibrium distributions, as well as the set of Nash equilibrium distributions and refinements thereof.

Now consider the perspective of an outside observer who finds the players willing to accept non-negative multiples of the gambles in matrix G. Suppose the observer is naïve in the sense that he has no preconceived beliefs about how the players are going to play the game. He need not even be aware that a "game" is being played. He merely observes two individuals who are willing to accept certain gambles with respect to four events called TL, TR, BL, and BR. A natural question is whether the observer can find any opportunities for *arbitrage*. For example, suppose that some non-negative linear combination of the gambles yields a strictly negative aggregate payoff to the players in some outcome—say, TL—and a zero aggregate payoff in the other

outcomes. Then there is no risk to the observer in making this combination of gambles: he cannot lose, and he will win something if TL subsequently occurs. Now, if TL is observed to occur, the observer will walk away with some of the players' money without having risked any of his own, and he will be entitled to conclude that the players have behaved irrationally. So it seems reasonable to require, as an absolutely minimal standard of rationality, that the players should play a joint strategy in which this kind of arbitrage is impossible. Such strategies may be said to be *jointly coherent*, by extension of de Finetti's term for probability judgments that do not expose an individual to arbitrage.

Which strategies are not jointly coherent? The problem of finding a combination of bets that yields an arbitrage payoff to the observer when a given joint strategy is played is just another linear programming problem. Formally, let $\boldsymbol{x}$ denote a vector whose elements are the multiples of the different gambles (rows of G) that the observer wishes to enforce on the players. Then $\boldsymbol{x} \bullet G$ is the vector of payoffs that the players (in the aggregate) receive from the observer depending on the outcome of the game (i.e., the joint strategy that is played). Suppose there is a vector $\boldsymbol{x}$ such that $\boldsymbol{x} \bullet G \leq \mathbf{0}$ with strict inequality in the j^{th} position. This means the players as a group *never win* money from the observer in any outcome of the game and they *lose* money to him in outcome j. In other words, there is *ex post arbitrage* if outcome j occurs. Now recall the following special case of the separating hyperplane theorem that we saw in class 2:

> **Lemma 2:** For any matrix G, either there exists a non-negative vector x such that $\boldsymbol{x} \bullet G \leq \mathbf{0}$, with strict negativity in the j^{th} position, or else there exists a (not necessarily unique) probability distribution $\boldsymbol{\pi}$, which assigns strictly positive probability to state j, such that $G\boldsymbol{\pi} \geq \mathbf{0}$.

Since $G\boldsymbol{\pi} \geq \mathbf{0}$ is the system of constraints defining a correlated equilibrium distribution, we see that **there is no ex post arbitrage in outcome *j* of the game if and only if outcome *j* has positive probability in a correlated equilibrium**, which means that joint coherence requires the players to act "as if" they have implemented a correlated equilibrium. Voila again! The problem of finding an ex post arbitrage opportunity in outcome j is precisely the *dual* of the problem of finding a correlated equilibrium distribution in which that strategy has positive probability, in the sense of the duality theorem of linear programming. By analogy with the "fundamental theorem of subjective probability" and the "fundamental theorem of asset pricing" (which we will meet later), which state that gambles or asset prices do not lead to arbitrage if and only if they are rationalized by a subjective probability distribution, we can call this result the "fundamental theorem of noncooperative games." Thus, the single axiom of no-arbitrage provides a unified characterization of rational beliefs, rational behavior in games, and rational behavior in markets.

By the way, this duality result also provides the basis for a simple and elementary proof of the existence of correlated equilibria. Of course, the existence of a correlated equilibrium is guaranteed by the existence of a Nash equilibrium. But the existence proof for Nash equilibria depends on a very powerful fixed-point theorem. Since correlated equilibria are much simpler computational objects than Nash equilibria, being defined by systems of linear inequalities, it seems as though it ought to be possible to prove their existence using only methods of linear algebra. This remained an interesting unsolved problem for more than a decade after Aumann first introduced the correlated equilibrium concept. Then in the late '80's, elementary existence

proofs were independently discovered by Sergiu Hart and David Schmeidler and by Kevin McCardle and myself (published in 1989 and 1990, respectively). The Nau/McCardle proof relies on the following simple observation: if a game did not have any correlated equilibria, then by the duality theorem there would be a combination of gambles that the observer could place that would yield a *sure win*—i.e., an arbitrage profit no matter what strategies were played. Life would be very unfair if this were the case: the players would be doomed merely by honestly revealing the tradeoffs that they face in the game! It is easy to show that life cannot be so unfair: no matter what combination of gambles is placed by the observer, the players can always construct randomized strategies that guarantee them an expected payoff of exactly zero, which means there cannot be any guaranteed arbitrage profits for the observer. The appropriate randomized strategies for each player are obtained by iteratively transferring probability from one strategy to another in proportion to the magnitude of the observer's bets, which eventually leads (by a Markov chain argument) to stationary strategies for each player that zero-out the observer's expected gain.

The concept of joint coherence extends naturally to games of incomplete information (as shown in my 1992 paper). The only difference is that in an incomplete information game the players may accept bets that are conditioned on their types as well as on their strategy choices. Let's return to Gul's example. If we translate the players' commonly known beliefs about each others' types into bets they should be willing to accept, then player 1 should accept bets consistent with the statement "if I were type X, I would say the probability of my opponent being A is 50%" and "if I were type Y, I would say the probability of my opponent being A is 50%". Player 2 should accept bets consistent with the statements "if I were type A, I would say the probability of my opponent being X is 40%" and "if I were type B, I would say the probability of my opponent being X is 50%". Presumably neither player will wish to reveal his or her actual type, but these *conditional* bets do not require them to. Now suppose an observer comes along and takes various positive multiples of the bets that are being offered, as follows:

Gul's example with commonly known beliefs backed up by bets

	AX	BX	AY	BY	**Bet multiplier chosen by an observer**
Player 1's bet if X	1	–1			**11**
Player 1's bet if Y			1	–1	**–9**
Player 2's bet if A	3		–2		**–4**
Player 2's bet if B		1		–1	**10**
Total payoff to players	**–1**	**–1**	**–1**	**–1**	

The players lose $1 as a group in every state of the world, so this is an ex ante arbitrage opportunity for the observer: the players are acting irrationally. This actually just another illustration of the fundamental theorem of subjective probability: the existence of an arbitrage opportunity is the dual condition of the absence of a probability distribution that agrees with all of the bets that have been publicly offered, whether are offered by the same person or not.

The correct solution concept for noncooperative games (IMHO)

Does this mean that objective correlated equilibrium is really the "correct" Bayesian solution concept for noncooperative games after all? Not quite! We have made a rather high-handed assumption of constant marginal utility for money, and we have not disposed of the troublesome matter of the common prior assumption, which emerges as a theorem rather than an axiom in the preceding analysis. These two restrictive assumptions are related: if the players really have constant marginal utility for money, then they had better have common beliefs, otherwise they would bet infinite sums of money with each other! So we would like to relax the assumption of constant marginal utility for money, and in so doing to weaken the common prior assumption.

The fundamental theorem of probability and the fundamental theorem of asset pricing both develop an interesting twist when the possibility is admitted that the agents might not have constant marginal utility for money: the supporting probability distributions have to be interpreted as *risk neutral probability distributions*—i.e., products of probabilities and relative marginal utilities for money. The same thing happens here. If the players do not have constant marginal utility for money, then the gambles they ought to accept are distorted by their marginal utilities for money. Let $v_1(a)$ denote player 1's marginal utility for money when her utility payoff is a, and similarly let $v_2(a')$ denote the marginal utility for money of player 2 when her utility payoff is a', and so on. Then the matrix whose rows are the payoff vectors of acceptable monetary gambles is modified as follows:

	TL	TR	BL	BR
1TB	$(a - c)/v_1(a)$	$(b - d)/v_1(b)$		
1BT			$(c - a)/v_1(c)$	$(d - b)/v_1(d)$
2LR	$(a' - b')/v_2(a')$		$(c' - d')/v_2(c')$	
2RL		$(b' - a')/v_2(b')$		$(d' - c')/v_2(d')$

Note that the utility difference in each cell is now divided by the appropriate marginal utility for money in that outcome of the game: this yields the *monetary amounts* whose increments of utility are proportional to the utility differences between outcomes of the game. Let this matrix be denoted by G^*, to distinguish it from the matrix G defined earlier in terms of utility functions alone. The matrix G encodes the rules of the "true" game, while the matrix G^* encodes the rules of the "revealed" game. It is straightforward to show that, if the players are risk averse in the sense that their marginal utility for money decreases with their level of utility, then the revealed game looks like a "fuzzy" version of the true game. The players hedge the bets they are willing to make, compared to the bets they would make if they had constant marginal utility for money, so the correlated equilibrium polytope is a strict superset of the polytope that would be obtained if they had the same original utility payoffs but with constant marginal utility for money.

A distribution $\boldsymbol{\pi}$ satisfying $G^*\boldsymbol{\pi} \geq \mathbf{0}$ is an objective correlated equilibrium distribution of the "revealed" game, and the outcomes to which it assigns positive probability are jointly coherent. But $\boldsymbol{\pi}$ must now be interpreted as a risk neutral probability distribution—which is to say, it may not represent any player's true probability distribution. But the players' "true" distributions must still be in equilibrium—in fact, they must form a *subjective* correlated equilibrium, in which each

player is behaving in a Bayesian rational way according to her own probabilities, but the probabilities of different players need not be mutually consistent. Nevertheless, this is a strong "refinement" of the subjective correlated equilibrium concept, because the *risk neutral distributions* of the different players must be mutually consistent, as they must also be in any arbitrage-free market under uncertainty. So, **the common prior assumption still applies—except that it applies to the players' risk neutral probabilities rather than their true probabilities.** In this form, the CPA is completely harmless and unobjectionable, a natural result of money-backed communication among the players. Thus, in a sense, the theory of noncooperative games reduces to a special case of the theory of markets, rather than the theory of markets reducing to a limiting case of the theory of games. (More details are contained in my working paper on "Arbitrage-Free Equilibria.")

Note that the imponderable mystery of game theory—the *infinite regress* of reciprocal expectations of rationality—has been entirely finessed away. In the example, we said nothing whatever about what Alice believed about what Bob believed about what Alice believed... Yet the infinite regress is implicit in the criterion of no-arbitrage when it is applied to the players as a group. If Alice behaves irrationally on her own (e.g., chooses a dominated alternative), that is an arbitrage opportunity all by itself. If Bob *bets on* Alice to behave irrationally (e.g., chooses an alternative that could only be rational for him if Alice behaved irrationally), then that too is an arbitrage opportunity. If Bob is wrong about Alice, you can collect a little from him, and if he is right, you can collect a lot from Alice and more than cover your loss to Bob. And if Alice bets on Bob to bet on Alice to behave irrationally, that too is an arbitrage opportunity, and so on up the ladder. The simple requirement of no-ex post-arbitrage automatically extrapolates the sequence to infinity—and a little bit more. Meanwhile, the Common Prior Assumption has mutated from a controversial form into an uncontroversial one: the requirement of common prior risk neutral probabilities is merely the condition for a competitive equilibrium in the market for contingent claims on outcomes of the game, and it is the natural result of credible public money-backed communication among the players.

Writing and rewriting the rules of the game

Another by-product of this mode of analysis is that it also provides an answer to the question of how the rules of a game might become common knowledge, or even more importantly, how they might come to be written in the first place. Recall that it is rather difficult in practice to measure anyone's "true" utilities. Indeed, we previously argued on theoretical grounds that it is *impossible*, even in principle, to fully separate utility from probability. Fortunately, the communication process outlined above renders it *unnecessary* to observe true utilities. It suffices for the players to observe the monetary gambles their opponents are willing to accept. These gambles determine the jointly coherent strategies that may be played and the arbitrage-free equilibria that support them. But if the players gamble with each other, this has an interesting side effect: the players can gradually *rewrite the rules of the game* through an accumulation of gambles. Hence, just as we are unable to fully separate utility from probability, so we also are unable to fully separate the revelation of the rules of the game from the writing of those rules.

We previously observed that game theory normally starts from the position that the rules have somehow already been etched in stone, and it then seeks to determine how the players ought to

behave within those fixed rules. This convention of game theory is often (and justly) criticized: why shouldn't the players have some freedom to modify the rules if they don't like the game they are stuck in? Perhaps strategic rationality consists in part of making up good rules by which to play. The approach outlined here incorporates that very idea into the concept of rational behavior in games. Obviously, this approach leaves quite a lot up to the players: they have considerable subjective freedom to believe what they want and to interact with others in materially significant ways during the pre-play communication process. However, to the extent that they end up constructing a game with well-defined rules, they must play it in a way that does not lead to arbitrage, and this requires them to choose joint strategies that are supported by an appropriate kind of equilibrium reasoning.

No-trade revisited

The reinterpretation of the common prior distribution as a risk neutral distribution has yet another interesting side effect: it solves the mystery of the "no-trade" theorems. If you start with the assumption of common prior risk neutral probabilities, you are effectively assuming that *the players have already exhausted any incentives to trade contingent claims on states of the world*. Hence it is not surprising that they should not wish to trade later on when they merely receive additional information through commonly-known information partitions. (With contingent contracts, they could have arranged such trades in advance.) Trade presumably occurs earlier in the process, when the common prior risk neutral distribution is being formed and revealed. Hence, the no-arbitrage standard of rationality does not imply that players with different information will never wish to trade with each other. Rather, it suggests that trade is likely to occur *earlier* as part of the communication process through which the rules of a game become commonly known. The inseparability of probability and utility plays a critical role in this process, obscuring the players' "true" probabilities and rendering it possible for them to sustain heterogeneous beliefs under conditions of common knowledge. (This issue is discussed in more detail in my paper on "The Incoherence of Agreeing to Disagree," *Theory and Decision* 1995)

The bottom line

I have argued that the appropriate standard of rationality in noncooperative games is that of *joint coherence*, and it leads (by our now-familiar duality argument) to the concept of an *arbitrage-free equilibrium*, a refinement of Aumann's concept of subjective correlated equilibrium in which players have *common prior risk neutral probabilities* rather than (on one hand) common prior true probabilities or (on the other hand) entirely unrelated true probabilities. This solution concept is merely the natural extension to the game-theoretic arena of concepts we previously developed for individual decisions under uncertainty and will see again later in the context of asset pricing. Does this mean I think that game theorists ought to drop Nash equilibrium (and Bayesian Nash equilibrium, perfect Bayesian equilibrium, etc.) and use arbitrage-free equilibrium or some other kind of correlated equilibrium instead ? Well, yes and no. I think several conclusions are appropriate.

First, common knowledge of rationality places rather weak constraints on the outcomes of noncooperative games, and you certainly should not rule out *a priori* the possibility that players might consciously or unconsciously use correlated strategies or otherwise hold correlated beliefs in settings where mixed-strategy equilibria are appropriate. If you wish to exclude mixed

strategies, then there is no difference between correlated and Nash equilibrium, and you should just say "equilibrium" rather than "Nash equilibrium." If you don't mind assuming that, as if by magic, the players begin with common knowledge of the "true" rules of the game as well as common prior true probabilities, then the set of *objective* correlated equilibria is the natural solution of the game in the sense that it exactly represents the set of mutually consistent beliefs that might be constructed by rational (narrowly self-interested) players, possibly aided by communication. If you wish to refine this concept, it would seem natural to allow the players to *bargain over the efficient frontier of the set of correlated equilibria* in order to obtain as mutually attractive a solution as possible—although even this dodge will not altogether avoid social dilemmas. Any concern for the efficiency of the solution requires the introduction of "dialog 2" information into the solution—i.e., information about the players' preferences for each others' actions—which strays outside the usual "rules" of noncooperative interaction. However, if the common prior assumption bothers you, you are on much firmer ground in applying it to theoretically-observable risk neutral probabilities than to theoretically-unobservable true subjective probabilities, and that path leads to *subjective* correlated equilibria. If your goal is to make sharp predictions about what will happen in a particular strategic situation, you may need to impose other assumptions in order to shrink the set of possible solutions. The right additional assumptions will depend on the situation being modeled and will not necessarily be the assumptions implicit in a more "refined" solution concept for generic games, so proceed with care. This caveat applies to rational choice models generally: rationality is by itself a very weak hypothesis, as Arrow emphasized in his essay on "Rationality of self and others." You need to be aware when auxiliary hypotheses are snuck in through the back door, and you need to scrutinize their authority and their psychological realism very carefully.

A second conclusion is that there is not a sharp dividing line between "individually rational" and "strategically rational" and "competitively rational" behavior. The boundaries between models of decisions, games, and markets are blurry. On one hand, game players ultimately make their moves as individuals, based on their own subjective views of their own situations, with at best only a finite, low-order regress of the form "I think she thinks I think." On the other hand, games of strategy often are played out in the context of a larger competitive market which determines (a) what the players know, (b) what they *commonly* know, (c) what social norms they respect, (d) what powers of computation are available to them (e) what channels of communication are open to them, (f) what moves are legally or financially or technologically possible, and (g) what their outside options might be. Not incidentally, competitive markets also eliminate the need for people to play games with each other over every little transaction. Thus, it may be unduly confining to think of an agent as being locked in a battle of minds with only one or two other agents, oblivious to everything else that is going on around them. Fortunately, there is one standard of rationality that applies equally to decisions, games, and markets, and it is the standard of no-arbitrage.

A third conclusion is that, when modeling an interactive decision problem in the form of a game, you should ask how and why the rules came to be written and whether it is possible for the players to rewrite them, either through contingent contracts or some other means. That may be where the real action is!