08-005

Correlated Equilibrium and Nash Equilibrium as an Observer's Assessment of the Game

John Hillas
Elon Kohlberg
John Pratt

John Hillas, professor, University of Auckland, New Zealand; Elon Kohlberg, Royal Little Professor of Business Administration, Harvard Business School [ekohlberg@hbs.edu]; John Pratt, William Ziegler Professor of Business Administration, Emeritus, Harvard Business School [jpratt@hbs.edu].

CORRELATED EQUILIBRIUM AND NASH EQUILIBRIUM AS AN OBSERVER'S ASSESSMENT OF THE GAME

JOHN HILLAS, ELON KOHLBERG, AND JOHN PRATT

Abstract. Noncooperative games are examined from the point of view of an outside observer who believes that the players are rational and that they know at least as much as the observer. The observer is assumed to be able to observe many instances of the play of the game; these instances are identical in the sense that the observer cannot distinguish between the settings in which different plays occur. If the observer does not believe that he will be able to offer beneficial advice then he must believe that the players are playing a correlated equilibrium, though he may not initially know which correlated equilibrium. If the observer also believes that, in a certain sense, there is nothing connecting the players in a particular instance of the game then he must believe that the correlated equilibrium they are playing is, in fact, a Nash equilibrium.

Contents

Date: July, 2007.
We are indebted to Jean-François Mertens for very helpful suggestions.

1. Introduction

A classic problem in probability theory concerns an observer making repeated observations of a phenomenon that he believes to be unchanging, for example, the toss of a coin. Now, as the observer makes more observations he learns something about the phenomenon, but eventually he will have learned (almost) everything that it is possible to learn. From that point the phenomenon will appear to the observer to be (approximately) an independent sequence of draws from the same distribution. Thus the observer's prior belief is that the observations are independently and identically distributed according to some unknown "true" probability, about which the observer has some probabilistic assessment.

De Finetti has given this description a precise meaning. The belief that the phenomenon is unchanging is formalised by saying that the observer's assessment is exchangeable. This means that the prior probability of any sequence of realisations is unchanged if the order of the sequence is permuted. De Finetti's Theorem says that an infinite exchangeable sequence can be represented by a mixture of independent identically distributed random variables.

In this paper we take a similar approach to the theory of games. We consider a particular game and imagine a setup where an outside observer makes repeated observations of the play of this game, under conditions which he considers identical. In particular, we think of the plays as taking place in different "rooms," unrelated to one another, with the observer looking into the rooms according to an order that he determines.

Under such a setup, the observer's prior assessment is exchangeable. Therefore, after a large number of observations, the plays of the game will appear to the observer to be independent draws from a fixed distribution on the players' choices.

We now further assume that the observer considers the players to be rational and experienced, and to know at least as much as the observer, so that, no matter how many observations he, the observer, might make he will never be in a position to offer a player beneficial advice. We show that under this assumption the distribution on the players' choices that the observer eventually will believe must be a correlated equilibrium. In particular, it follows that the observer's prior assessment of a single play of the game must be a mixture of correlated equilibria, which itself is a correlated equilibrium.

Continuing in the same spirit, we ask what additional conditions would guarantee that the observer's eventual assessment be not only a correlated equilibrium, but also a Nash equilibrium.

As Aumann (1987) has forcefully argued, it makes no sense to *assume* that the observer's assessment of a one-shot play exhibit mutual independence of the players' choices; even if the players' choices neither can affect one another nor are affected by a common observation which is unknown to the observer, still the observer's assessment may well exhibit correlation. Thus, as in de Finetti's Theorem, independence should be a conclusion rather than an assumption.

The condition we propose is a natural extension of exchangeability. We assume that not only must the assessment of the sequence of the players' choices remain the same if all the players in one room are exchanged with all the players in another room, but also that the assessment must remain the same if only one player from one room is exchanged with the player in the same role in another room. This formalises the idea that there is nothing connecting the players in the same room.

With this stronger condition of separate exchangeability, we show that the observer's assessment must be a mixture of independent and identically distributed random variables, each of which is a Nash equilibrium.

2. Exchangeability

In this section we give a formal definition of exchangeability and the limited results concerning exchangeability that we use.

Definition 1. An infinite sequence of random variables $\{X^t\}_{t=1}^{\infty}$ taking values in some finite set $\mathcal{X}$ is *exchangeable* if for every t and t' we may exchange X^t and $X^{t'}$ leaving the other X's as they were without changing the joint distribution.

Remark 1. The joint distribution remains the same when any finite number of X^t are permuted, since every finite permutation is a composition of pairwise interchanges.

The central result concerning exchangeable processes is the theorem of de Finetti.

Theorem 1 (de Finetti). *Let $\{X^t\}_{t=1}^{\infty}$ be an infinite exchangeable sequence of random variables taking values in $\mathcal{X}$. Then there is a unique measure μ on $\Delta^{\mathcal{X}}$ the space of probability distributions on $\mathcal{X}$ such that* *for any T and any sequence $x^1, \dots, x^T$ with $x^t \in \mathcal{X}$*

$$P(X^1 = x^1, \dots, X^T = x^T) = \int \prod_{t=1}^{T} p(x^t)\, d\mu(p).$$

This says that if the random variables are exchangeable then the sequence is a mixture of independent and identically distributed random variables, that is the probability of a particular finite sequence of observations is the same as it would have been had the sequence been generated by first choosing p according to the distribution μ and then generating the sequence as independent and identically distributed random variables distributed according to p.

An immediate implication of exchangeability is that each of the X^t has the same distribution. Further, this remains true even after we have observed some of the X^t, that is conditional on $X^{t_1}, X^{t_2}, \dots$ all the remaining X's are identically distributed. In fact, the converse is also true and provides another characterisation of exchangeability.

Theorem 2. *If, given any finite subset $X^{t_1}, X^{t_2}, \dots, X^{t_K}$ all the remaining X's are identically distributed, then the sequence $X^1, X^2, X^3, \dots$ is exchangeable.*

Proof. The hypothesis is that, for every finite subset Z and every X^t and $X^{t'}$ not in Z, the conditional distributions of X^t given Z and $X^{t'}$ given Z are the same. Multiplying by the probability of Z shows that

(1) (X^t, Z) and $(X^{t'}, Z)$ have the same joint distribution.

To show that the sequence $X^1, X^2, X^3, \dots$ is exchangeable, we must show that the joint distribution of $X^1, \dots, X^T$ remains the same when any pair of X's are interchanged. Taking the pair to be X^1, X^2, we use the notation $V \sim W$ to mean that $(V, X^3, X^4, \dots, X^T)$ and $(W, X^3, X^4, \dots, X^T)$ have the same joint distribution. The desired conclusion then is $(X^1, X^2) \sim (X^2, X^1)$.

By (1) we show this as follows:

$$(X^1, X^2) \sim (X^{T+1}, X^2) \sim (X^{T+1}, X^1) \sim (X^2, X^1).$$

Specifically, the first $\sim$ follows from $X^t = X^1$, $X^{t+1} = X^{T+1}$, and $Z = (X^2, \dots, X^T)$, and the others are similar. □

The alternative characterisation highlights the distinction between independence and exchangeability. Independence rules out any learning. In contrast, exchangeability allows learning about the process (that is, the distribution of the future Xs), but it does not allow learning that distinguishes between different instances of the process (that is, the observer's distributions on any two future Xs are the same.)

We now give a formal definition of the concept of separate exchangeability[1] that we discussed in the Introduction. The idea is that the set $\mathcal{X}$, the set in which the random variable takes its values has a product structure, $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$ and that as well as being able to exchange X^t and $X^{t'}$ without changing the distribution we can also do this separately for each n.

Definition 2. An infinite sequence of random variables $\{X^t\}_{t=1}^{\infty} = \{(X_1^t, \ldots, X_N^t)\}_{t=1}^{\infty}$ taking values in $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$ is separately exchangeable if, for every n and every t and t', we may exchange X_n^t and $X_n^{t'}$ leaving the other X's as they were without changing the joint distribution.

If the process is separately exchangeable then it is obviously exchangeable. We can exchange X^t and $X^{t'}$ by doing so one coordinate at a time and each time the joint distribution remains unchanged.

In addition, if the process is separately exchangeable then, it is as if there is "nothing connecting" the random variable X_n^t with the random variable $X_{n'}^t$. Just as the fact that there is nothing connecting the outcome of one coin toss with another does not mean that they are independent, so here, we cannot say that X_n^t and $X_{n'}^t$ are independent. However, we do obtain the same kind of conditional independence that de Finetti's Theorem gives us for the outcomes of an exchangeable process.

Theorem 3. *Let $\{(X_1^t, \ldots, X_N^t)\}_{t=1}^{\infty}$ be an infinite separately exchangeable sequence of random variables taking values in $\mathcal{X} = \times_{n=1}^N \mathcal{X}_n$. Then $\{(X_1^t, \ldots, X_N^t)\}_{t=1}^{\infty}$ is exchangeable and the distributions p given in Theorem 1 are product distributions, $p(x) = \prod_n p_n(x_n)$.*

Proof. As we argued above, if the process is separately exchangeable it is exchangeable. So, by de Finetti's Theorem it can be viewed as a mixture of independent and identically distributed random variables. It is easy to see that, conditional on the value of p, the process remains separately exchangeable. We can see this by exchanging X_n^t and $X_n^{t'}$ and then conditioning on the tail event that the time averages converge to a particular p. Since the joint distribution remains unchanged so do the conditional distributions.

Thus, since, conditional on p, (X_n^t, X_{-n}^t) is independent of $(X_n^{t'}, X_{-n}^{t'})$ it follows that, conditional on p, $(X_n^t, X_{-n}^{t'})$ is independent of $(X_n^{t'}, X_{-n}^t)$,

[1] What we are calling separately exchangeable is the same concept that is introduced in Kohlberg and Reny (1997) under the name *coordinate-wise exchangeability*. See pages 299–300 of their paper for a more detailed discussion of this concept.

and in particular, conditional on p, X_n^t is independent of X_{-n}^t. So (with probability 1 under μ) p is a product distribution. □

3. Preliminary result concerning statistical testing

In this section we prove a result that we shall use later in characterising correlated and Nash equilibria. The setup is the following. We observe an infinite sequence $\{X^t\}_{t=1}^{\infty}$ where at each time t the random variable X^t takes on either the value 0 or the value 1. We assume that the process is exchangeable. As we have seen above, this implies that there is some unknown probability p and that, conditional on p, the X^t are independently and identically distributed and that $X^t = 1$ with probability p.

Let $S^t = \sum_{\tau=1}^{t} X^\tau$ and $\bar{p}^t = S^t/t$, that is S^t is the number of times that X^τ has taken on the value 1 up through time t and $\bar{p}^t$ is the proportion of times that X^τ has taken on the value 1 up through time t.

A classical problem of statistical inference is to find a testing rule that will, on the basis of the values of X up to time t and some desired level of precision ε say whether or not to reject the hypothesis that p is some particular value, say p_0.

A standard solution to this problem is to find a value ε_0 such that if $p = p_0$ then the probability that $|\bar{p}^t - p_0| \geq \varepsilon_0$ is less than ε and to accept the hypothesis that $p = p_0$ if $|\bar{p}^t - p_0| < \varepsilon_0$ and to reject it otherwise. The value of ε_0 will typically depend on both the number of observations t and on the hypothesised value p_0. If we wish to reject the hypothesis that $p = p_0$ as often as possible when $p \neq p_0$ subject to the constraint that we reject the hypothesis with at most probability ε when the hypothesis is true then we will choose ε_0 as small as possible subject to this constraint, and, since the distribution of $\bar{p}^t$ depends on both t and p_0, so will this optimal value of ε_0.

In our application of this kind of idea we shall not be so concerned to reject the hypothesis as often as possible when it is wrong, but shall, in other respects, want a bit more from the test.

Firstly, we shall not be testing whether p takes on a particular value, but rather whether it lies in some region. (This is in some ways similar to the problem of finding confidence intervals or regions.)

Secondly we shall not be asking simply for a testing rule that for a particular t rejects with probability at most ε when the hypothesis is true but rather a sequence of rules, one for each t, such that the probability of ever rejecting the hypothesis, when it is true, is at most ε.

To deal with the first aspect we would like to require not simply that the probability that $|\bar{p}^t - p_0| \geq \varepsilon_0$ be small for a particular value of p_0

but that this probability be small for all values of p_0 in the region we are testing. We solve this problem by strengthening, but simplifying, the requirement to say that the probability should be small for all values of p_0. That is, we require ε_0 to be such that, whatever the value of p_0, the probability that $|\bar{p}^t - p_0| \geq \varepsilon_0$ is less than ε.

To deal with the second aspect we shall look at testing rules given by a sequence of ε_t where we reject the hypothesis that $p = p_0$ after making t observations if $|\bar{p}^t - p_0| \geq \varepsilon_t$ and accept it otherwise.

The following theorem tell us that we can find a testing rule that will achieve these objectives.

Theorem 4. *Let $\{X^t\}_{t=1}^{\infty}$ be a sequence of independent and identically distributed random variable taking values in $\{0, 1\}$ with $X^t = 1$ with probability p. Then there is sequence $\varepsilon_t \downarrow 0$ such that for every $\varepsilon > 0$ there is T such that for every $p \in [0, 1]$*

$$P_p(|\bar{p}^t - p| \geq \varepsilon_t \textit{ for some } t > T) \leq \varepsilon. \tag{2}$$

In fact, we prove a sharper result (Theorem 8). Its statement and proof are given in the Appendix.

Remark 2. For fixed p if we set $\varepsilon_t = \varepsilon_0$ for all t then Theorem 4 is one way of stating the strong law of large numbers. For fixed p, but with ε_t as in the theorem, it is a standard extension of the strong law.

Remark 3. Theorem 4 remains true if, rather than taking values in $\{0, 1\}$, X^t takes values in some finite set $\mathcal{X}$ with p a distribution over $\mathcal{X}$.

The theorem allows us to define a testing rule that satisfies the requirements we laid out above. Up to time T always accept the hypothesis. At time $t > T$ reject the hypothesis if $|\bar{p}^t - p_0| \geq \varepsilon_t$ and accept it otherwise.

4. One-player decision problem or game against Nature

We now consider a slightly more general setting. We consider a setting in which there are a finite set of possible states $\mathcal{X}$ and the observer observes a number of instances of a decision-maker choosing among a finite set of possible choices S, having known preferences represented by the expectation of a utility function $u : \mathcal{X} \times S \to \mathbb{R}$. In each instance the observer observes both the realised state and the choice of the decision maker.

We consider the perspective of an outside observer who has the opportunity to observe this situation a large number of times and who regards each observation as being an observation of the same phenomenon. Thus he sees different players, with exactly the same preferences, facing exactly the same problem. The observer does not know exactly how Nature's choice is generated, or how the player makes his choice. We assume that the observer has some exchangeable prior on how the history of both Nature's choices and the players' choices will unfold and may, on the basis of his past observations, suggest modifications to the players' choices. In fact we restrict the observer to modifications of the following form.

Definition 3. A modification to the player's choices is a sequence of functions $m = (m^1, m^2, \dots, m^t, \dots)$ with $m^t : \mathcal{X}^{t-1} \times S \to S$ where $m^t(x^1, \dots, x^{t-1}, s)$ is the observer's recommendation to the player on seeing him about to play s for the tth time and having seen the states $x^1, \dots, x^{t-1}$ the previous $t-1$ times that the player chose s.

Remark 4. There are potentially other modification rules in which the modification the observer proposes after seeing the player choose s for the tth time may depend on the whole history and not just on what we have observed at the times at which he previously chose s. If the player had no more information than the observer then this would be a useful thing for the observer to do. However we do not assume this, and it may be that the player sees things that the observer does not. Moreover we want the modification not to do badly against any exchangeable prior, including ones in which, with positive probability, this is the case. In this situation if the observer uses observations of times at which the player chose something other than s he may be using evidence from situations that the player knows to be irrelevant, and so offering the player bad advice. In any case if we get a result that depends on there being no possibility of an improvement using this kind of modification, the result would also follow from a requirement that there be no possibility of an improvement using a wider class of modifications.

Even though we assume that the player is an expected utility maximiser we introduce the notation of preference since it makes some of the arguments more readable. Given a distribution q on $\mathcal{X}$ we say that the player prefers the choice s to the choice s' given q and write $s \succ_q s'$ if

$$E_q u(X, s) > E_q u(X, s').$$

Definition 4. The modification m is said to be ε-good if for any process obtained as the independent repetition of draws from $\mathcal{X} \times S$ with distribution p and any $s \in S$ having positive probability under p

(1) if s does not maximise $\succ_{p(\cdot|s)}$ the modification m will almost surely choose some s' that does maximise $\succ_{p(\cdot|s)}$ for all but a finite number of the periods in which the player chooses s; and
(2) the modification m will choose some s' such that $s' \succ_{p(\cdot|s)} s$ for all periods in which it recommends a change from s, with probability at least $1-\varepsilon$.

Remark 5. The requirement is that a modification be robust against all exchangeable priors and not just good against the particular prior that the observer holds. We know that an exchangeable process is as if some distribution p on $\mathcal{X} \times S$ is chosen randomly and then independently repeated. Thus if, for any p our modification will work well for the independent repetition of p, it will work well for any exchangeable prior.

Remark 6. Condition (2) of Definition 4 can be thought of as saying two things. If s maximises $\succ_{p(\cdot|s)}$ there is no s' that's better then s and so the condition requires that the probability that the modification ever recommend a change be small. If s does not maximise $\succ_{p(\cdot|s)}$ Condition (1) says that eventually the modification will always recommend a change and Condition (2) says that with high probability this recommended change will always be preferred according to $\succ_{p(\cdot|s)}$.

Definition 5. For a particular exchangeable prior an ε-good modification of the player's choice is said to be an ε-good improvement if there is some time T such that the modification strictly improves the expectation under the prior of the player's expected utility for every period beyond T.

Theorem 5. *If the observer has an exchangeable prior and there exists $\varepsilon > 0$ such that there is no ε-good improvement of the players' choices then the observer's prior must be a mixture of independent repetitions of a random variable with distribution p over $\mathcal{X} \times S$ such that for each $s \in S$ having positive probability under p, s maximises $\succ_{p(\cdot|s)}$.*

Proof. Denote by $\Delta^{\mathcal{X}}$ the set of all distributions over $\mathcal{X}$. Let

$$B_{s's} = \{q \in \Delta^{\mathcal{X}} \mid s' \succ_q s\}.$$

That is, $B_{s's}$ is the set of probability distributions over $\mathcal{X}$ for which s' is a strictly better choice than s. This set is relatively open in $\Delta^{\mathcal{X}}$.

Now, given $\varepsilon > 0$ define the modification as follows: Choose T as in the generalisation of Theorem 4 described in Remark 3.

For $\bar{q} \in \Delta^{\mathcal{X}}$ let

$$S(\bar{q}, \gamma) = \{q \in \Delta^{\mathcal{X}} \mid \|q - \bar{q}\|_\infty \leq \gamma\}.$$

We'll look at modifications m for which m^t depends on the history $x^1, \ldots, x^{t-1}$ only through $\bar{q}^{t-1}$, which gives the empirical proportions of each value in $\mathcal{X}$ for the first $t-1$ times that the player chose s. Consider the modification m where $m^t(\bar{q}^{t-1}, s)$ equals one of the choices that maximises $\succ_{\bar{q}^{t-1}}$ if $t > T$ and $S(\bar{q}^{t-1}, \varepsilon_t) \subset B_{s's}$ for some s', and otherwise recommends no change from s.

Suppose s does not maximise $\succ_{p(\cdot|s)}$. Then $p(\cdot \mid s) \in B_{s's}$ for some s' and, since $B_{s's}$ is open, the strong law of large numbers implies that with probability 1, there is a T_1 such that $\bar{q}^{t-1}$ is in $B_{s's}$ for all $t > T_1$; also, since $\varepsilon_t \downarrow 0$ there is a T_2 such that $S(\bar{q}^{t-1}, \varepsilon_t) \subset B_{s's}$ for all $t > T_2$; and finally there is a T_3 such that any choice that maximises $\succ_{\bar{q}^{t-1}}$ also maximises $\succ_{p(\cdot|s)}$. Thus when t is large enough m will recommend a switch to a choice that maximises $\succ_{p(\cdot|s)}$. This establishes property (1) of Definition 4.

Now, if $m^t(\bar{q}^{t-1}, s) = s' \neq s$ and $s \succsim_{p(\cdot|s)} s'$ then

(i) $t > T$,
(ii) $S(\bar{q}^{t-1}, \varepsilon_t) \subset B_{s's}$, and
(iii) $p(\cdot \mid s) \notin B_{s's}$,

so, from (ii) and (iii), $\|p(\cdot \mid s) - \bar{q}^t\|_\infty \geq \varepsilon_t$. Theorem 4 tells us that

$$P(\|p(\cdot \mid s) - \bar{q}^t\|_\infty \geq \varepsilon_t \text{ for some } t > T) \leq \varepsilon.$$

And this establishes part (2) of Definition 4.

Finally we need to show that if, for the observer's assessment, there is a positive measure of p which put positive probability on s which do not maximise $\succ_{p(\cdot|s)}$, we can find T' such that from T' onward the modification will strictly increase the expected payoff of the player, according to the observer's assessment. But this follows from what we have already shown, since for any $t > T$ there is a positive probability that the history will be such that the modification will recommend a change. □

Remark 7. We have made no attempt in the proof to find the best modification. The modification we define is, in some respects, quite poor. For example, we wait a long time before even recommending against strictly dominated strategies.

Remark 8. Notice that we do not simply hand over the choice of strategy to the observer. The observer "sees" what the player intends to play and then recommends to the player whether to go ahead or to play something else. Consider the following problem.

	Sunny	Raining
T	2	0
M	0	1
B	0	0

Suppose that over a very long series of observations the observer sees (T, Sunny) in 0.49 of the observations, $(M, \text{Raining})$ in 0.49 of the observations, (B, Sunny) in 0.01 of the observations, and $(B, \text{Raining})$ in 0.01 of the observations. Clearly the player is not behaving optimally.

Nevertheless, an outside observer could not, by himself, do better than the player. What the outside observer could do is suggest an improvement to the player. If he saw the player intending to play either T or M he would tell the player to go ahead, while if he saw the player intending to play B he would recommend that he switch to T instead.

Notice that the distribution $p(\cdot \mid s)$ over Nature's choice is $p(\cdot \mid T) = (1, 0)$, $p(\cdot \mid M) = (0, 1)$, and $p(\cdot \mid B) = (1/2, 1/2)$. Now T maximises the player's expected utility against $p(\cdot \mid T)$ and M maximises his expected utility against $p(\cdot \mid M)$. However B does not maximise his expected utility against $p(\cdot \mid B)$, and the observer can offer a profitable modification on seeing B.

5. Correlated Equilibrium

In this section we take an approach to the theory of games similar to the approach we developed in the previous section. We consider an observer of a potentially infinite number of plays of the game. We think of the players as experienced and rational and assume that this is commonly known among the players. We think of the observer as knowing no more, and quite possibly less, than the players, and as knowing this. Under these circumstances, it seems unreasonable for the observer to believe that he can offer any player beneficial advice, that is, recommend an assuredly beneficial departure from the strategy that the player intends to use. We shall show that, when these assumptions are made precise in a natural way, the observer's probabilistic assessment of how the game is played must constitute a correlated equilibrium.

If we take the observer's belief to be that he can never give advice that is beneficial according to his own probabilistic assessment of how the game is being played, then his belief in correlated equilibrium is immediate. Indeed, if an assessment is such that no player has a profitable deviation conditional on his own choice, then the assessment is a correlated equilibrium by definition.

As in the previous section, we explore a less egocentric, more demanding interpretation of the concept of beneficial advice. In particular, we do not assume that the observer will offer advice just because it is good in his own probabilistic assessment. Rather, we assume that he will offer advice only if it would be good in all assessments, that is, in the eyes of every observer with exactly the same knowledge as his. Of course, if he is considering only one play of the game, he could then offer only very limited advice, such as to avoid dominated strategies. We assume instead that the observer can view a large number of plays of the game that are, to him, identical. This does not imply that the plays are independent from his perspective, but only that they are exchangeable, and so the observer's assessment is that the observations are independently, identically distributed according to an unknown "true" distribution p about which the observer has some probabilistic assessment.

Now suppose the observer gives advice based on his previous observations of the game and what he sees a player about to do in the current game. Can he offer advice that would be good regardless of his prior? It would appear that he can offer little more than he could offer in the case of a single game. We therefore relax the requirement just a little and ask: could an observer offer advice that would, whatever the exchangeable prior, with a probability arbitrarily close to 1, never suggest harmful modifications, and with positive probability, suggest improving modifications from some time onward, where "harmful" and "improving" refer to the player's expected payoff under p? We show that if the observer believes that he cannot offer any player such advice, then he must believe that the players are choosing optimally with respect to p at every play, that is, their play is a correlated equilibrium based on the "true" p. In the limit, the observer will learn the "true" correlated equilibrium. At each play along the way, his assessment of the future is a mixture of independent repetitions of correlated equilibria based on his current distribution for p. Since a mixture of correlated equilibria is itself a correlated equilibrium, his assessment at each play is also a correlated equilibrium. His assessment may change as he makes additional observations, but it must always be a correlated equilibrium.

We now flesh out this rather informal analysis and make it more explicit. In the previous section we considered a simple decision-making problem and defined what we meant by an ε-good improvement. We then showed how to construct such an improvement. We now consider the game theoretic problem and show how to define a similar ε-good modification rule in that setting, how to construct such a modification, and when such an ε-good modification rule will be an ε-good improvement. We use the result of the previous section to show that the modification we construct is ε-good, and that if there is no ε-good improvement then the observer must believe that the players are playing a correlated equilibrium, though he may not initially know which correlated equilibrium they are playing.

Consider a game $G = (N, S, u)$ where $N = \{1, \dots, n, \dots, N\}$ is the set of players, $S = S_1 \times \cdots \times S_n \times \cdots \times S_N$ with S_n the set of pure strategies of Player n, and $u = (u_1, \dots, u_n, \dots, u_N)$ with $u_n : S \to \mathbb{R}$ the payoff function of Player n.

Definition 6. A distribution p over S constitutes a *correlated equilibrium* if for each player n it imputes positive marginal probability only to such pure strategies, s_n, as are best responses against the distribution on the others' pure strategies obtained by conditioning on s_n.

Suppose there is some observer who observes an infinite sequence of plays of the game, observing in period t the realised strategy $s^t \in S$. We assume that the observer has some prior over the sequence of plays that he will see and that this prior is exchangeable.

The observer contemplates advising the players. We ask: what must have been the observer's assessment if he did not expect to be able to offer profitable advice. As in the previous section, we ask a bit more of his advice. We require not simply that his advice be good if his prior is correct, but that it not be bad even if his prior is incorrect. Thus we define a modification to a player's strategy as follows.

Definition 7. A modification to Player n's strategy is a sequence of functions $m = (m^1, m^2, \dots, m^t, \dots)$ with $m^t : S^{t-1} \times S_n \to S_n$ where $m^t(s^1, \dots, s^{t-1}, s_n)$ is the observer's recommendation to Player n on seeing him about to play s_n for the tth time and having seen the strategy profiles $s^1, \dots, s^{t-1}$ the previous $t - 1$ times that Player n chose s_n.

Remark 9. The comments we made in Remarks 4 and 8 remain relevant. The observer does not take over from the player. Rather he looks over the player's shoulder, sees what the player is about to play, and then recommends to the player whether to go ahead or to play something else. Further, we could make the modification suggested by the observer when he sees the player about to take a particular choice depend on all the observations rather than only those in which he had previously made that choice, but against some priors this could be quite bad.

That is, the observer observes the past up until the period in which he observes Player n making a particular choice for the tth time. On the basis of what has happened in the $t-1$ times in the past that he has observed Player n making this choice he then decides whether to advise Player n to do something other than his intended action.

Since the observer's prior is exchangeable it means that he has a probability distribution that is the same as if there had been an initial random choice and that the realisations then, conditional on this initial choice, were independently and identically distributed. We do not require robustness against all possible other priors, but only against other exchangeable ones. Thus, after the initial randomisation, the plays will be independently and identically distributed draws over the space S from some distribution p.

Definition 8. A modification to Player n's strategy m is said to be ε-good if for any process obtained as the independent repetition of draws from S with distribution p and any $s \in S_n$ having positive probability under p

(1) if s does not maximise Player n's expected utility against $p(\cdot \mid s)$ the modification m will almost surely choose some strategy that does maximise Player n's expected utility against $p(\cdot \mid s)$ for all but a finite number of the periods in which Player n chooses s; and
(2) the modification m will choose some better strategy against $p(\cdot \mid s)$ for all periods in which it recommends a change from s with probability at least $1-\varepsilon$.

Definition 9. For a particular exchangeable prior an ε-good modification of a player's strategy is said to be an ε-good improvement if there is some time T such that the modification strictly improves the expectation under the prior of that player's expected utility for every period beyond T.

Remark 10. From the perspective of the observer thinking about advising Player n, we are in exactly the situation of the previous section with S_n playing the role of S and S_{-n} the role of $\mathcal{X}$.

Theorem 6. *If the observer has an exchangeable prior and there exists some $\varepsilon > 0$ such that there is no ε-good improvement of any player's strategy then the observer's assessment is a mixture of independent and identically distributed correlated equilibria.*

Proof. The result follows immediately from Theorem 5 on observing that, for each player n, the conditions of Theorem 5 are met with $S = S_n$ and $\mathcal{X} = S_{-n}$. □

Corollary 1. *The observer's assessment of any individual play of the game is a correlated equilibrium.*

Proof. The set of correlated equilibria is convex. (Aumann, 1974) □

6. Nash Equilibrium

We now give a similar characterisation of Nash equilibrium. In many circumstances it makes completely good sense that the outcome of a game will be a correlated equilibrium rather than a Nash equilibrium. Thus we should not expect that the assumptions involved in characterising Nash equilibrium should be compelling. Rather than being compelling assumptions they are the conditions that identify when a correlated equilibrium will be a Nash equilibrium.

Nash equilibria are typically defined in terms of the mixed strategies of the players. Let us denote by Σ_n the set of mixed strategies of Player n, that is the set of all probability distributions over S_n. We give the standard definition of Nash equilibria in a slightly nonstandard way, though we are not the first to give the definition in this form.

Definition 10. A profile of mixed strategies $(\sigma_1, \ldots, \sigma_N)$ in $\Sigma = \Sigma_1 \times \cdots \times \Sigma_N$ is a Nash equilibrium if the distribution over S generated by the independent product of $\sigma_1, \ldots, \sigma_N$ is a correlated equilibrium.

This definition emphasises that a Nash equilibrium may be viewed as a correlated equilibrium in which the players' strategy choices are probabilistically independent. We do not regard correlated equilibria that are not Nash equilibria as necessarily unreasonable. Such equilibria make perfectly good sense when there is, at least implicitly, something connecting the players. For example, two players may have observed some correlated signals. If we wish to rule out such equilibria, we must rule out such connections between the players. In our setting we do this by assuming that the observer views the process as separately exchangeable in the sense that the probability he attaches to an arbitrary finite history is unchanged if the choices of a single player at two different times are switched. An equivalent condition, an analogue in this setting to Theorem 2, is that, given any finite set of plays by all players but one, and any subset of that player's choices, his remaining choices are identically distributed. In this case, as we saw in Theorem 3, not only are the plays independently identically distributed according to an unknown "true" distribution p, but also the players' choices on each play are independent, that is, p is the product of the players' individual distributions. It follows that, if the observer believes that he cannot offer any player beneficial advice and views the process as separately exchangeable, then the "true" correlated equilibrium that the observer eventually learns must be a Nash equilibrium, and his assessment must be a mixture of independent repetitions of Nash equilibria.

Theorem 7. *If the observer has a separately exchangeable prior and there is no ε-good improvement of any player's strategy then the observer's assessment must be a mixture of independent and identically distributed Nash equilibria.*

Proof. The result follows directly from Theorems 3 and 6. □

7. Conclusion

We have examined the concepts of correlated equilibrium and Nash Equilibrium by considering the perspective of an observer of an arbitrarily large number of observations of the game. We argued that it was reasonable to think of the observer's prior on these observations as being exchangeable, and that it then followed that, if the observer did not ever expect to be in a position to offer beneficial advice, his prior would necessarily put weight only on correlated equilibria. If, in addition, the observer's prior was separately exchangeable then his prior would necessarily put weight only on Nash equilibria.

While we find the assumption that the observer's prior be exchangeable is quite compelling our condition for Nash equilibrium, which we term, separate exchangeability, is considerably less so. This is as it should be. There are circumstances in which correlated equilibria that are not independent are quite reasonable. Rather, separate exchangeability is a natural implication of the idea that there is nothing connecting the player in a particular role with the other players in the particular game he is playing. This may or may not, depending on the context, be a reasonable assumption. We show that when this assumption is satisfied the observer will believe that the players are playing a Nash equilibrium, though he may not initially know which one.

In Hillas, Kohlberg, and Pratt (2007) we develop very similar ideas in an explicitly epistemic setting. In that setting it is possible to compare our conditions with those used by Aumann and Brandenburger (1995). Theirs are in some sense minimal, and are to that extent preferable. They do not however contain the conditions of Aumann (1974, 1987), replacing some part of those with the conditions that they use to obtain independence. Our work more clearly separates the conditions that imply correlated equilibrium and the additional ones that imply that the correlated equilibrium exhibits independence across the choices of the different players, and is therefore a Nash equilibrium.

In everything we have done we have simply made assumptions about the view of the observer, and our results give properties of his probabilistic assessment. We have said nothing about what the players actually do. One might, somewhat trivially, say that if we assume also that the observer is correct in his assessment, then the players will actually be playing equilibria.

References

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

Robert J. Aumann: (1987): "Correlated Equilibria as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

Robert J. Aumann and Adam Brandenburger: (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161–1180.

Bruno de Finetti: (1974): *Theory of Probability: A critical introductory treatment*, volume 1, Wiley Interscience Publishers, Chichester.

Bruno de Finetti: (1975): *Theory of Probability: A critical introductory treatment*, volume 2, Wiley Interscience Publishers, Chichester.

William Feller: (1950): *An Introduction to Probability Theory and Its Applications*, volume 1, John Wiley and Sons, New York, first edition.

John Hillas, Elon Kohlberg, and John Pratt: (2007): "An Epistemic Approach to the Relation between Correlated and Nash Equilibria," Unpublished.

Elon Kohlberg and Philip Reny: (1997): "Independence on Relative Probability Spaces and Consistent Assessments in Game Trees," *Journal of Economic Theory*, 75, 280–313.

David B. Peizer and John W. Pratt: (1968): "A Normal Approximation for Binomial, F, Beta, and Other Common, Related Tail Probabilities, I," *Journal of the American Statistical Association*, 63, 1416–1456.

John W. Pratt: (1968): "A Normal Approximation for Binomial, F, Beta, and Other Common, Related Tail Probabilities, II," *Journal of the American Statistical Association*, 63, 1457–1483.

Appendix. A Statistical Result

Here we prove a considerably sharper version of Theorem 4 above.

Theorem 8. *For all $\delta \in (0, \frac{1}{2})$ and $\varepsilon > 0$, there is T such that for all $p \in [0,1]$ and $t > T$,*

$$P_p(|\bar{p}^\tau - p| \geq \tau^{-\delta} \textit{ for some } \tau > t) \leq \varepsilon e^{-t^{1-2\delta-\varepsilon}}.$$

In other words, after some time T, the last τ such that $|\bar{p}^\tau - p| \geq \tau^{-\delta}$ is stochastically smaller for every p than $x^{1/(1-2\delta-\varepsilon)}$ where x has the standard exponential distribution.

Before proving this we prove two preliminary lemmas.

Lemma 1. *For all $\varepsilon > 0$,*

$$P_p(\bar{p}^\tau - p \leq -\varepsilon) \leq \frac{A}{\varepsilon}\left(\frac{pq}{\tau}\right)^{\frac{1}{2}} e^{-\varepsilon^2\tau/2p},$$

where A is a constant, not depending on the other parameters, and $q = 1 - p$.

Proof. By Pratt (1968, equation 7.7), for $0 \leq s \leq \tau p - \frac{1}{2}$,
(3)

$$P_p(\bar{p}^\tau \leq \frac{s}{\tau}) = \frac{1}{-z\sqrt{2\pi}} e^{-\frac{1}{2}z^2 G + R}\left(\frac{\tau p - s - \frac{1}{2}}{\tau p - s + p} \cdot \frac{\tau - s + 1}{\tau - s}\right)\frac{1}{1 + \theta z^{-2}}$$

where $0 \leq \theta \leq 1$, R is a sum of Stirling-formula remainders, $z = (s + \frac{1}{2} - \tau p)(\tau pq)^{-\frac{1}{2}}$ is the usual "half-corrected" standardised deviate, and

$$G = 1 + qg\left(\frac{s + \frac{1}{2}}{\tau p}\right) + pg\left(\frac{\tau - s + \frac{1}{2}}{\tau q}\right) \tag{4}$$

where

$$g(x) = \begin{cases} \frac{1 - x^2 + 2x\log x}{(1-x)^2} & \text{if } x > 0,\ x \neq 1 \\ 1 & \text{if } x = 0 \\ 0 & \text{if } x = 1. \end{cases}$$

The last two factors of equation (3) are easily shown to be at most 1. R has a finite upper bound, say ρ. Peizer and Pratt (1968, Section 10) show that $g(x) = -g(1/x)$ and g is continuous, strictly decreasing, and convex. In particular, therefore, $G \geq 1 - p = q$ in equation (4). For $s/\tau = p - \varepsilon$, we have $-z = \varepsilon(\tau/pq)^{1/2}$ and Lemma 2 follows with $A = \frac{1}{\sqrt{2\pi}} e^\rho$. □

Lemma 2. *For all $\varepsilon > 0$,*

$$P_p(|\bar{p}^\tau - p| \geq \varepsilon) \leq \frac{A}{\varepsilon\sqrt{\tau}} e^{-\varepsilon^2\tau/2},$$

where A is the same constant as in Lemma 1.

Proof. Interchanging p and q on the right hand side in Lemma 2 gives an upper bound for $P_p(\bar{p}^\tau - p \geq \varepsilon)$ and hence

$$P_p(|\bar{p}^\tau - p| \geq \varepsilon) \leq \frac{A}{\varepsilon}\left(\frac{pq}{\tau}\right)^{\frac{1}{2}}\left(e^{-\varepsilon^2\tau/2p} + e^{-\varepsilon^2\tau/2q}\right). \tag{5}$$

Since $pq \leq \frac{1}{4}$ and $p \leq 1$, $q \leq 1$, Lemma 3 follows. □

Proof of Theorem 8. By Lemma 2, with $\varepsilon = \tau^{-\delta}$,

$$\begin{aligned}
P_p(|\bar{p}^\tau - p| \geq \tau^{-\delta} \text{ for some } \tau > t) &\leq \sum_{\tau=t+1}^{\infty} \frac{A}{\tau^{\frac{1}{2}-\delta}} e^{-\frac{1}{2}\tau^{1-2\delta}} \\
&\leq \int_t^\infty A x^{\delta-\frac{1}{2}} e^{-\frac{1}{2}x^{1-2\delta}} \mathrm{d}x \\
&= \int_{\frac{1}{2}t^{1-2\delta}}^\infty \frac{2A}{1-2\delta}(2y)^{\frac{1}{1-2\delta}-\frac{3}{2}} e^{-y}\mathrm{d}y \qquad \begin{cases}\text{by the substitution} \\ y = \frac{1}{2}x^{1-2\delta}\end{cases} \\
&= \frac{A2^a}{1-2\delta}\Gamma(a, \tfrac{1}{2}t^{1-2\delta}) \qquad \begin{cases}\text{where } a = \frac{1}{1-2\delta} - \frac{1}{2} \text{ and} \\ \Gamma(a,x) = \int_x^\infty y^{a-1}e^{-y}\mathrm{d}y \\ \text{is the incomplete Gamma function}\end{cases} \\
&\sim \frac{A2^a}{1-2\delta}\left(\tfrac{1}{2}t^{1-2\delta}\right)^{a-1} e^{-\frac{1}{2}t^{1-2\delta}} \qquad \text{as } t \to \infty \\
&= \frac{2A}{1-2\delta} t^{3\delta-\frac{1}{2}} e^{-\frac{1}{2}t^{1-2\delta}}.
\end{aligned}$$

The asymptotic result used in the second last step is standard. It follows, for example, from Peizer and Pratt (1968, equation 9.5), which also contains other references. Theorem 8 follows. □

Remark 11. The essential features of Lemma 2 are that the right hand side is independent of p and summable in τ. Lemma 1, of course, gives tighter bounds and equation (3) still tighter, but they depend on p. The strong law of large numbers already tells us that $\bar{p}^\tau \to p$ with probability 1 and the law of the iterated logarithm gives the asymptotic magnitude of $\bar{p}^\tau - p$, but Theorem 8, though less tight, provides specific bounds on tail probabilities uniformly in p, which our purposes require and usual proofs do not provide.

Remark 12. $R = R_1(\tau) - R(s) - R(\tau - s)$ where R_1 and R are remainders in the two forms of Stirling's formula, namely,

$$\begin{aligned}\tau! &= \sqrt{2\pi}\tau^{\tau+\frac{1}{2}}e^{-\tau+R_1(\tau)} \\ &= \sqrt{2\pi}(\tau+\tfrac{1}{2})^{\tau+\frac{1}{2}}e^{-\tau-\frac{1}{2}+R(\tau)}.\end{aligned}$$

Feller (1950, pages 44 and 50) shows that $R_1(\tau) \leq \frac{1}{12\tau}$ and $R(\tau) \geq -\frac{1}{24(\tau+\frac{1}{2})}$. Since $R_1(1) = 1 - \frac{1}{2}\log(2\pi) \approx .0811$ and $-R(0) = \frac{1}{2}\log\pi - \frac{1}{2} \approx .0724$ and $-R(1) = \frac{1}{2}\log(3\pi) + \log(\frac{3}{2}) - \frac{3}{2} \approx .0271$, it follows that for $0 \leq s$,

$$R \leq \rho = R_1(1) - R(0) - R(1) = \tfrac{1}{2}\log\left(\frac{3\pi}{2}\right) + \log\frac{3}{2} - 1 \approx .1806,$$

and

$$A = \frac{1}{\sqrt{2\pi}}e^{\rho} = \frac{3\sqrt{3}}{4e} \approx .4779.$$

Remark 13. Using the fact that either p or q is at most a half in the proof of Lemma 3 would replace the right hand side by

$$\frac{A}{2\varepsilon\sqrt{\tau}}\left(e^{-\varepsilon^2\tau/2} + e^{-\varepsilon^2\tau}\right)$$

which is asymptotically smaller by a factor of a half. This improvement would not affect the exponential rate in Theorem 8, however.