CORRELATED EQUILIBRIUM AND BEHAVIORAL CONFORMITY

by

Edward Cartwright and Myrna Wooders

Working Paper No. 05-W26

November 2005

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Correlated equilibrium and behavioral conformity*

Edward Cartwright
Department of Economics
Keynes College, University of Kent,
Canterbury CT2 7NP, UK
E.J.Cartwright@kent.ac.uk

Myrna Wooders
Department of Economics
University of Warwick
Coventry CV4 7AL, UK
M.Wooders@warwick.ac.uk
http://www.warwick.ac.uk/fac/soc/Economics/wooders/

This version: November 2004

Abstract

Is conformity amongst similar individuals consistent with self-interested behavior? We consider a model of incomplete information in which each player receives a signal, interpreted as an allocation to a role, and can make his action choice conditional on his role. Our main result demonstrates that 'near to' any correlated equilibrium is an approximate correlated equilibrium 'with conformity' – that is, an equilibrium where all 'similar players' play the same strategy, have the same probability of being allocated to each role, and receive approximately the same payoff; in short, similar players 'behave in an identical way' and are treated nearly equally. To measure 'similarity' amongst players we introduce the notions of approximate substitutes and a (δ, Q)-class games – a game with Q classes of players where all players in the same class are δ-substitutes for each other.

*This paper is a major revision of Cartwright and Wooders (2003a). The main results of that paper were presented at the the 2002 General Equilibrium Conference held in Athens in May 2002 and at Northwestern University in August 2002. We thank the participants for their interest and comments.

1 Introduction

The concept of correlated equilibrium was introduced in Aumann (1974,1987). In contrast to Nash equilibrium, correlated equilibrium allows strategies of players to be statistically dependent. More precisely, before playing a game, a player receives a signal on which he can condition his choice of action. If the signals are independent across players then a correlated equilibrium is a Nash equilibrium of the original game. Since the signals may, however, be correlated across individuals the set of correlated equilibria is generally larger than the set of Nash equilibria. Correlated equilibria have many appealing properties, as discussed by Aumann (1974, 1987). See also Hart (2005) for a discussion of recent work on how adaptive learning leads to correlated equilibrium play.

An interpretation of correlated equilibrium is to imagine a 'referee' or 'device' distributing instructions to players on what action to play. The probability distribution with which instructions are allocated is common knowledge. A correlated equilibrium is a probability distribution over instructions such that it is in each player's interest to obey the instructions. One appealing aspect of this interpretation of correlated equilibrium is that each player uses a pure strategy. In particular, each player has a pure strategy mapping instructions into actions of the form 'if told to play action x then play action x'. Randomization or mixing results from the distribution of instructions and not from the actions of players themselves. A second aspect of correlated equilibrium is that all players use the same pure strategy and 'obey' the instructions they are given. This suggests that one could connect the concept of correlated equilibrium with conformity to a norm of behavior; every players 'fit into the role' that the device or referee allocates to him. The objective of this paper is to begin to explore the extent to which correlated equilibrium can be connected with the concept of behavioral conformity.

Taking a normal form game as given, we assume that roles are allocated to players by some device. The set of roles is equivalent to the set of actions and so the allocation of a role can be seen as an instruction to play a particular action. A correlated equilibrium is a probability distribution over the allocation of roles such that each player's best response is to 'take on his role'. Alternatively, in 'the game with roles' (where nature allocates roles) it is a Nash equilibrium to play the action corresponding to role allocated. As suggested above, the notion of a correlated equilibrium already builds in some notion of conformity. In general, however, we may question the level of actual conformity if the probability distribution over roles is asymmetric. That is, if different players have different probabilities of being allocated to each role we may begin to question the extent to which players could be said to conform or 'behave in a similar way'. If the distribution over roles is symmetric then it is easier to argue that there is conformity in behavior; each player has the same probability of being allocated each role and each player behaves in the same way once allocated a role: 'similar players behave identically'.

Our main result shows that ‘near to’ any correlated equilibrium is a ‘symmetric’ correlated equilibrium The equilibrium is symmetric in the sense that ‘players who are similar’ are treated identically by the device allocating roles. A symmetric correlated equilibrium has the additional property that similar players receive similar expected payoffs. It is therefore also an equilibrium that one could argue is ‘fair’. It has been widely observed that individuals are motivated by conditions of fairness where fairness equates to equality of outcome or opportunity amongst ‘similar people’.

To illustrate consider the familiar 2-person example of Aumann (1973):

	A	B
A	$6,6$	$2,7$
B	$7,2$	$0,0$

This game has two pure strategy Nash equilibria (A, B) and (B, A). Neither of these equilibria could be thought as demonstrating conformity or fairness. There exists a correlated equilibrium where the device allocates roles (A, A), (A, B) and (B, A) with probability $\frac{1}{3}$ each. This probability distribution is symmetric and results in expected payoffs of 5 for both players.

It is important for our results to measure the similarity between players. We do this by defining the concepts of δ-substitutes and a (δ, Q)-class game. If two players i and j are δ-substitutes and they swap strategies then (i) i's payoff would be within δ of the payoff j previously received and (ii) the payoff of any other player l would change by at most δ. If $\delta = 0$ then we can think of players as being identical in terms of the game. A game is (δ, Q)-class game if the player set can be partitioned into Q subsets or classes where any two players in the same class are δ-substitutes. Any game can be classified as a (δ, Q)-class game for any Q and some δ. It is thus a useful tool to measure the similarity of players in an arbitrary game.

A second example illustrates why it is important to measure the similarity of players. Consider the 3-person game where, in this example and the next two, player 1 picks a row, player 2 picks a column and player 3 picks a matrix.:

A

	A	B	C
A	$0,0,5$	$0,0,0$	$0,0,0$
B	$0,0,0$	$0,0,0$	$0,0,0$
C	$0,0,0$	$0,0,0$	$0,0,0$

B

	A	B	C
A	$2,2,3$	$0,0,0$	$0,0,0$
B	$0,0,0$	$0,0,0$	$0,0,0$
C	$0,0,0$	$0,0,0$	$2,2,3$

C

	A	B	C
A	$0,0,0$	$0,0,0$	$0,0,0$
B	$0,0,0$	$0,0,0$	$0,0,0$
C	$0,0,0$	$0,0,0$	$0,0,5$

There is no strict Nash equilibrium in pure strategies in which all players play the same strategy and get the same payoff.[1] One correlated equilibrium is to allocate roles $(A, A, B), (C, C, B)$ with equal probability; essentially players 1 and 2 toss a coin to decide top left or bottom right but player 3 does not get to see the result of this coin toss. Is this equilibrium symmetric and fair? If we compare all three players then it is clearly neither fair (player 3 gets a higher payoff) or symmetric (player 3 is allocated role B and players 1 and 2 are never allocated role B). The equilibrium is, however, both fair and symmetric if we compare 'similar players'. It can be seen that players 1 and 2 are 0-substitutes. Player 3, by contrast, is 'dissimilar' to players 1 and 2 (e.g. if all play A their payoffs differ by 5). Given players 1 and 2 are treated equally, the equilibrium can be seen as fair and symmetric.

One interesting aspect of the above example of correlated equilibrium is that there is no correlation amongst classes of players. That is, if we think of players 1 and 2 as a class, then they jointly correlate their actions but they do not correlate their actions with those of player 3 (who belongs to a different class). We refer to this a correlated equilibrium with class independence. We may interpret correlation of roles (or coordination of the allocation of roles) as viable within classes but not between classes. But, in general, it turns out that there need not exist a correlated equilibrium that is symmetric and satisfies class independence 'near to' *every* correlated equilibrium. We demonstrate, however, that in any game there does exist a correlated equilibrium satisfying symmetry and class independence. This is most easily illustrated by considering a 2-player game where the 2 players are not 0-substitutes. It may, though, be more instructive to consider the following 3-player example:

A

	A	B	C
A	$5, 5, 4$	$0, 0, 0$	$0, 0, 0$
B	$0, 0, 0$	$0, 0, 0$	$0, 0, 0$
C	$0, 0, 0$	$0, 0, 0$	$0, 0, 0$

B

	A	B	C
A	$2, 2, 3$	$0, 0, 0$	$0, 0, 0$
B	$0, 0, 0$	$0, 0, 0$	$0, 0, 0$
C	$0, 0, 0$	$0, 0, 0$	$2, 2, 3$

C

	A	B	C
A	$0, 0, 0$	$0, 0, 0$	$0, 0, 0$
B	$0, 0, 0$	$0, 0, 0$	$0, 0, 0$
C	$0, 0, 0$	$0, 0, 0$	$5, 5, 4$

As in the previous example, players 1 and 2 are 0-substitutes so can be thought as belong to the same class. Player 3 is in a class by himself. There is a correlated equilibrium in which roles (A, A, A) and (C, C, C) are allocated with equal probability.

[1]The example can easily be modified so there is no (not necessarily strict) Nash equilibrium with these properties.

This is symmetric but does not satisfy class independence given that player 3's role is correlated with that of players 1 and 2. The correlated equilibrium where roles (A, A, B) and (C, C, B) are allocated with equal probability is symmetric and does satisfy class independence.

Subjective correlated equilibrium is a generalization of correlated equilibrium: Having observed a signal (or allocation to a role) a player forms beliefs about the signals others have received (the true state of the world) and behaves relative to these beliefs. Correlated equilibrium is typically defined to correspond to the case where all players have a common prior and so, ex-ante, have the same beliefs about the state of the world. In this paper we shall make the stronger statement that correlated equilibrium corresponds to the case where all players have a common prior and this prior is consistent with the actual distribution with which roles are allocated. Subjective correlated equilibrium is a generalization whereby different players may have different priors, or in our framework, may have priors not consistent with the true mechanism allocating roles.

In a model of bounded rationality it seems of interest to model the case where each player 'believes similar players will behave in the same way'. That is, if a player views two other players as similar he may expect them to behave in a similar way. This appears a reasonable rule of thumb that a player may adopt and it can be equated with expecting similar players to have the same probability of being allocated to roles. We demonstrate (Theorem 3) that 'near to' any correlated equilibrium p is a subjective correlated equilibrium β where similar players are expected to have the same probability of being allocated roles. Further, β is expected to be fair given player beliefs. Note that the equilibrium p may or may not be symmetric and may or may not be fair. Thus, it may be an equilibrium for players to behave according to the rule of thumb 'similar players will behave in the same way' even if they will not in reality do so. Players may also perceive an outcome as fair even if it is not in fact the case. We can illustrate with this three player example:

A

	A	B	C
A	$0,0,0$	$0,0,0$	$0,0,0$
B	$0,0,0$	$0,0,0$	$0,0,0$
C	$0,0,0$	$0,0,0$	$0,0,0$

B

	A	B	C
A	$0,0,0$	$0,0,0$	$2,6,2$
B	$0,0,0$	$0,0,0$	$0,0,0$
C	$2,6,2$	$0,0,0$	$0,0,0$

C

	A	B	C
A	$0,0,0$	$0,0,0$	$0,0,0$
B	$0,0,0$	$0,0,0$	$0,0,0$
C	$0,0,0$	$0,0,0$	$0,0,0$

The distribution $p(A, C, B) = 1$ is a correlated equilibrium. Given that players 1 and 2 are 0-substitutes p is neither fair or symmetric. Suppose all players had the same subjective beliefs $\beta(A, C, B) = 0.5$ and $\beta(C, A, B) = 0.5$. This is a subjective correlated equilibrium that is fair and symmetric (and we shall argue 'close to' p). This example, raises the question of whether it is reasonable for players to have erroneous beliefs about their own allocation to roles. In this example, it may appear reasonable for player 3 to have beliefs β but less reasonable for players 1 and 2 to have beliefs β. If, however, all players have 'correct beliefs' about the distribution of types in their own class then we still obtain a correlated equilibrium. For example, if the beliefs of player's 1 and 2 are $\beta_1(A, C, B) = \beta_2(A, C, B) = 1$ while player 3 has 'naive beliefs' $\beta_3(A, C, B) = 0.5$ and $\beta_3(C, A, B) = 0.5$ we still obtain a subjective correlated equilibrium. In general, we see that it can be an equilibrium to behave and expect similar players to behave in an identical way even if they do not do so in reality.

We proceed as follows: Section introduces the model, Section 3 defines a (δ, Q) class game while Section 4 presents a preliminary result. Section 5 contains the main result. Some extensions are considered in Section 6 and the paper concludes in Section 7. An appendix contains the remaining proofs.

2 A game with roles

A *game* Γ is given by a triple $(N, A, \{u_i\}_{i \in N})$ consisting of a finite player set $N = \{1, .., n\}$, a finite set of *actions* $A = \{1, ..., K\}$, and a set of payoff functions $\{u_i\}_{i \in N}$. An *action profile* consists of a vector $\overline{a} = (\overline{a}_1, ..., \overline{a}_n)$ where $\overline{a}_i \in A$ denotes the *action of player* i. For each $i \in N$ the payoff function u_i maps A^N into the real line.

Take as given a game $\Gamma = (N, A, \{u_i\}_{i \in N})$ and let $R = \{1, ..., K\}$ denote a set of *roles* (where we note R is identical to A). A *role profile* consists of a vector $\overline{r} = \{\overline{r}_1, ..., \overline{r}_n\}$ where $\overline{r}_i \in R$ is the role of player i. The set of role profiles is R^N. Note that R^N is identical to A^N. A *probability distribution over role profiles* is a function p where $p(\overline{r})$ denotes the probability of role profile $\overline{r}$. Let $P = \Delta(R^N)$ denote the set of probability distributions over role profiles. We shall denote by p_i the marginal distribution of p where $p_i(k)$ denotes the probability that player i is allocated role k. Formally, $p_i(k) = \sum_{\overline{r}:\overline{r}_i = k} \overline{r}$.

Once roles are randomly allocated to players (according to distribution p) each player chooses an action. A player can make his action choice conditional on his role. He is not, however, informed of the roles of the complementary player set. A player's payoff does not depend directly on his role or the roles of other players (although it may do indirectly through the choice of action that a distribution of roles induces). We shall assume for the present that distribution p is common knowledge and players have consistent beliefs with respect to the distribution. We relax this assumption in Section 6.

Given game Γ and probability distribution $p \in P$ a game with roles is denoted

Γ^p. For each player $i \in N$ a *strategy* in game Γ^p is a function s_i mapping the set of roles R to the set of actions A. In interpretation $s_i(k)$ is the pure action performed by player i if of role k. Let S denote the set of strategies and let $\overline{s}^* \in S^N$ denote the *strategy profile* for which $s_i^*(k) = k$ for all $i \in N$. That is, each player plays pure action k if allocated role k.

We define payoff function $U_i : P \to \mathbb{R}$ for each player $i \in N$ where

$$U_i(p) := \sum_{\overline{a} \in A^N} p(\overline{a}) u_i(\overline{a}).$$

It can be observed that $U_i(p)$ denotes the expected payoff of player i if roles are allocated according to p and players behave according to $\overline{s}^*$. We say that p is a *correlated ε-equilibrium of game* Γ if and only if

$$U_i(p) \geq \sum_{\overline{a} \in A^N} p(\overline{a}) u_i(s_i(\overline{a}_i), \overline{a}_{-i}) - \varepsilon$$

for all $s_i \in S$. Thus, no player would wish to deviate from the role he is allocated. If p is a correlated ε-equilibrium of game Γ then we can equivalently say that $\overline{s}^*$ is a Nash ε-equilibrium of game Γ^p (where approximate Nash equilibrium is defined in the standard way).

3 Approximate substitutes

Given a game $\Gamma = (N, A, \{u_i\}_{i \in N})$ we consider partitioning the player set N into groups with the property that any two players in the same group can be viewed as approximate substitutes for each other.[2] This requires us to formulate a metric by which to compare players. We consider two different ways of measuring the distance between players. Informally, we say that two players i and j are interaction substitutes if i and j are seen as similar by those with whom they interact; so, if the actions of i and j are interchanged, then the payoffs to other players are only slightly affected. In contrast, we say that players i and j are individual substitutes if they have similar payoff functions. Combining both measures together, we refer to players i and j as approximate substitutes if they are both interaction and individual substitutes.

Approximate substitutes: Let $j, l \in N$ be any two players and $\overline{a}^1, \overline{a}^2 \in A^N$ be any two pure action profiles where (1) $\overline{a}_i^1 = \overline{a}_i^2$ for all $i \neq j, l$, (2) $\overline{a}_j^1 = \overline{a}_l^2$ and (3) $\overline{a}_l^1 = \overline{a}_j^2$. We say that j and l are *δ-interaction substitutes* if

$$\left| u_i(\overline{a}^1) - u_i(\overline{a}^2) \right| \leq \frac{\delta}{n} \tag{1}$$

[2]We remark that the idea of defining approximate substitutes for a given cooperative game, without any reference to an underlying topological space of player types, appears in several papers due to Kovalenkov and Wooders, cf. their 2003 *Journal of Economic Theory* paper and references therein.

for any player $i \in N$, $i \neq j, l$. We say that j and l are *δ-individual substitutes* if

$$\left|u_j(\overline{a}^1) - u_l(\overline{a}^2)\right| \leq \delta. \tag{2}$$

We say that j and l are *δ-substitutes* (or informally approximate substitutes) if they are both δ-interaction substitutes and δ-individual substitutes.

A partition $\{N_1, ..., N_Q\}$ of player set N is a *δ-substitute partition* if for any set N_q in the partition and any two players j and l with $j, l \in N_q$, j and l are δ-substitutes. Given $\delta \geq 0$ and an integer Q, a game Γ is a (δ, Q) *class game* if there is a *δ-substitute partition* of N into Q subsets.

Some trivial observations are: First, the partition into singletons $\{\{1\}, ..., \{n\}\}$ is a 0-substitute partition. Essentially, each player is a 0-substitute for themselves. Also, for any game Γ and any $Q \leq N$, for some finite $\delta \geq 0$ there exists a δ-substitute partition. Finally, if Γ is a (δ, Q) class game then Γ is a (δ', Q') class game for any $\delta' \geq \delta$ and and $Q' \geq Q$. In general, the closer the approximation (the smaller is δ) the larger the number of classes required for Γ to be a (δ, Q) class game.

4 Permutation of action profiles

Take as given a (δ, Q)-class game Γ with a partition into classes $\{N_1, ..., N_Q\}$. For any pure action profile $\overline{a}$ let $h(\overline{a}, k, q)$ denote the number of players in class N_q who play pure action k. Given a pure action profile $\overline{a}$ we say that action profile $\overline{a}'$ is a *permutation of* $\overline{a}$ if $h(\overline{a}, k, q) = h(\overline{a}', k, q)$. That is, players within the class have exchanged strategies amongst each other. For any action profile $\overline{a}$ let $\mathcal{P}(\overline{a})$ denote the set of action profiles that are permutations of $\overline{a}$. Given that we can talk of role and action profiles interchangeably we shall also talk of role profile $\overline{r}'$ as a permutation of $\overline{r}$ if $\overline{r}' \in \mathcal{P}(\overline{r})$

We now state an important preliminary result, the proof of which is contained in an Appendix.

Lemma 1: Let Γ be a (δ, Q)-class game, i and j any two players who are δ-substitutes and $\overline{a}$ any pure action profile. If $\overline{a}' \in \mathcal{P}(\overline{a})$ and $\overline{a}'_j = \overline{a}_i$ then

$$|u_j(\overline{a}') - u_i(\overline{a})| \leq \delta + \frac{n-2}{n}\delta.$$

It is worth emphasizing that Lemma 1 is not immediate from the definition of a δ-substitute partition. In particular, when defining a δ-substitute partition consideration is only given to the effect on payoffs when two players exchange strategies. Lemma 1 treats a permutation whereby all players may change strategy and where there need not be a 'simple exchange' of strategies. It is possible, for instance, that $\overline{a}'_i \neq \overline{a}_j$.

5 Conformity

We define a symmetry condition on the probability distribution over roles. Take as given a game Γ and a partition $\Pi = \{N_1, ..., N_Q\}$ of the player set into classes.

Within class anonymity: A probability distribution over roles p satisfies *within class anonymity (WCA)* if the distribution p treats players from the same class identically. Formally, given any two role profiles $\overline{r}$ and $\overline{r}'$, if $\overline{r}' \in \mathcal{P}(\overline{r})$ then:

$$p(\overline{r}) = p_j(\overline{r}').$$

One might think of within class anonymity as an 'equal opportunity' condition within classes; it implies, for instance that if $i, j \in N_q$ then $p_i(k) = p_j(k)$ for all $k \in A$.

A correlated equilibrium p where p satisfies WCA could be interpreted as an equilibrium where similar players behave identically: they play the same strategy and have the same probability of being allocated each role. Informally, the following Theorem states that near to any correlated equilibrium is an approximate correlated equilibrium satisfying WCA and where players in the same class receive approximately the same payoff.

Theorem 1: Let Γ be a (δ, Q)-class game, let $\Pi = \{N_1, ..., N_Q\}$ be a a partition of the player set into classes and let p^* be a correlated equilibrium of Γ. Then there exists a correlated ε equilibrium p' of Γ satisfying WCA, where $\varepsilon \leq 4\delta$ and

$$\left| \frac{1}{|N_q|} \sum_{j \in N_q} U_j(p^*) - U_i(p') \right| \leq 2\delta \tag{3}$$

for all $i \in N_q$ and all $N_q \in \Pi$.

Proof: Recall that $p^*(\overline{a})$ denotes the probability of role profile $\overline{a}$. Let p' denote a function mapping R^N into the unit interval $[0, 1]$ where

$$p'(\overline{a}) = \frac{\sum_{\overline{a}' \in \mathcal{P}(\overline{a})} p^*(\overline{a}')}{|\mathcal{P}(\overline{a})|} \tag{4}$$

for all $\overline{a} \in R^N$. We conjecture that p' satisfies the desired conditions. This requires us to check four things: (1) p' is a probability distribution over roles, (2) p' satisfies WCA, (3) p' is a correlated equilibrium and (4) (3) holds. We verify each in turn. First, note that $\overline{a} \in \mathcal{P}(\overline{a})$. Also, if $\overline{a}' \in \mathcal{P}(\overline{a})$ then $\mathcal{P}(\overline{a}') = \mathcal{P}(\overline{a})$ for all $\overline{a}', \overline{a} \in \mathcal{A}^N$. Thus, the set of action profiles $\mathcal{A}^N$ can be partitioned into a finite set of sets of actions profiles $\Psi^1, \Psi^2, ..., \Psi^L$ where $\overline{a}, \overline{a}' \in \ominus^l$ if and only if $\overline{a} \in \mathcal{P}(\overline{a}')$.

(1) By construction, $p'(\overline{a}) = p'(\overline{a}')$ for any $\overline{a}, \overline{a}' \in \Psi^l$ and

$$\sum_{\overline{a}\in\Psi^l} p'(\overline{a}) = \left|\mathcal{P}(\overline{a}^l)\right| p'(\overline{a}^l) = \sum_{\overline{a}\in\Psi^l} p^*(\overline{a}) \tag{5}$$

where $\overline{a}^l \in \Psi^l$. Thus, p' is a probability distribution over role profiles.

(2) Observe that,

$$p_i'(k) = \sum_{\psi^l} \sum_{\overline{a}\in\Psi^l:\overline{a}_i=k} p'(\overline{a}) = \sum_{\Psi^l} \left|\left\{\overline{a} \in \Psi^l : \overline{a}_i = k\right\}\right| p'(\overline{a}^l). \tag{6}$$

If $i, j \in N_q$ for some N_q then by definition $\left|\left\{\overline{a} \in \Psi^l : \overline{a}_i = k\right\}\right| = \left|\left\{\overline{a} \in \Psi^l : \overline{a}_j = k\right\}\right|$ and thus, by (6) $p_i'(k) = p_j'(k)$ and WCA is satisfied.

(3) Given that p^* is a correlated equilibrium

$$\sum_{\overline{a}\in\mathcal{A}^N:\overline{a}_i=k} p^*(\overline{a})u_i(k, \overline{a}_{-i}) \geq \sum_{\overline{a}\in\mathcal{A}^N:\overline{a}_i=k} p^*(\overline{a})u_i(k', \overline{a}_{-i}) \tag{7}$$

for all $i \in N$ and $k' \in \mathcal{A}$. For any $i \in N_q$ and $k \in \mathcal{A}$, if $p_i'(k) > 0$ then by construction there exists some $j \in N_q$ such that $p_j^*(k) > 0$. If $i, j \in N_q$ and $\overline{a}'$ is a permutation of $\overline{a}$ and $\overline{a}_j' = \overline{a}_i$ then, from Lemma 1,

$$|u_j(\overline{a}') - u_i(\overline{a})| \leq 2\delta. \tag{8}$$

From (5) and (8) we get

$$\sum_{\overline{a}\in\mathcal{A}^N:\overline{a}_i=k} p'(\overline{a})u_i(k, \overline{a}_{-i}) \geq \sum_{\overline{a}\in\mathcal{A}^N:\overline{a}_i=k} p'(\overline{a})u_i(k', \overline{a}_{-i}) - 4\delta$$

for all $i \in N$ and $k' \in \mathcal{A}$. Thus, p' is a correlated 4δ equilibrium.

(4) By Lemma 1, for any N_q and Ψ^l and $\overline{a}, \overline{a}' \in \Psi^l$

$$\left|\sum_{j\in N_q} u_j(\overline{a}) - \sum_{j\in N_q} u_j(\overline{a}')\right| \leq 2\delta \left|N_q\right|. \tag{9}$$

For each N_q and Ψ^l pick some $\overline{a} \in \Psi^l$ and let

$$u^{jl} = \sum_{j\in N_q} u_j(\overline{a}).$$

Now, by (9),

$$\begin{aligned}\left|\sum_{j\in N_q} U_j(p') - \sum_{j\in N_q} U_j(p^*)\right| &= \left|\sum_{\Psi^l}\sum_{j\in N_q}\sum_{\overline{a}\in\Psi^l} u_j(\overline{a})[p'(\overline{a}) - p^*(\overline{a})]\right| \\ &\leq \left|\sum_{\Psi^l} u^{jl}\sum_{\overline{a}\in\Psi^l}[p'(\overline{a}) - p^*(\overline{a})]\right| + |N_q|\, 2\delta \\ &= |N_q|\, 2\delta.\end{aligned}$$

From (8) and the construction of p' we have $|U_i(p') - U_j(p')| \leq 2\delta$ for all $i, j \in N_q$. Thus, if $i \in N_q$,

$$\left|\frac{1}{|N_q|}\sum_{j\in N_q} U_j(p^*) - U_i(p')\right| \leq 4\delta. \tag{10}$$

This completes the proof.■

An immediate corollary of Theorem 1 is that near to any correlated equilibrium p^* is a correlated equilibrium p' satisfying WCA where $|U_i(p') - U_j(p')| \leq 4\delta$ for all i, j belonging to the same class. Also note that the set of correlated ε-equilibrium satisfying WCA is convex.

6 Remarks and extensions

6.1 Class independence

As discussed in the introduction it is of interest to question when there exists a correlated equilibrium for which is there no correlation of roles between classes. Take as given a game Γ and a partition $\Pi = \{N_1, ..., N_Q\}$ of the player set into classes.

Class independence: Let i and j be any two players belonging to different classes. Let $p_i(k|r_j)$ denote the probability player i has role k given that player j has role r_j. A probability distribution over roles satisfies *class independence (CI)* if $p_i(k) = p_i(k|r_j)$ for all $k, r_j \in R$.

The following result demonstrates the existence of an approximate correlated equilibrium satisfying both CI and WCA.

Theorem 2: Let Γ be a (δ, Q)-class game and let $\Pi = \{N_1, ..., N_Q\}$ a partition of the player set into classes. Then there exists a correlated ε equilibrium p' of Γ satisfying WCA and CI where $\varepsilon \leq 4\delta$

Proof: By the standard existence theorems game Γ has a Nash equilibrium $\overline{\alpha}^* \in \Delta(A)^N$. Denote by $\overline{\alpha}_i^*(k)$ the probability that player i plays action k. Let p^* be a probability distribution over roles where roles are stochastically independent across players and $p_i^*(k) = \overline{\alpha}_i^*(k)$ for all $i \in N$. Clearly p^* is a correlated equilibrium (and satisfies CI). Using the notation from the Proof of Theorem 1, let p' be a probability distribution over roles where,

$$p'(\overline{a}) = \frac{\sum_{\overline{a}' \in \mathcal{P}(\overline{a})} p^*(\overline{a}')}{|\mathcal{P}(\overline{a})|} \tag{11}$$

for all $\overline{a} \in R^N$. From the Proof of Theorem 1 we see that p' is a correlated ε equilibrium satisfying WCA (for some $\varepsilon < 4\delta$). It is clear that p' also satisfies CI. [Note that p' need not be 'fully' independent across players. For example (with three pure actions) $\overline{\alpha}^*$ may have the property that $\overline{\alpha}_1^* = (0.5, 0.5, 0), \overline{\alpha}_2^* = (0, 0, 1)$; suppose that players 1 and 2 constitute a class; p' will be such that $p_1'(3) = 0.5$ but $p_1'(3|r_2 = 1) = 1$.]■

6.2 Subjective Beliefs

We have assumed to this point that all agents have objective beliefs with respect to some known probability distribution p. A more general possibility is that each player has their own subjective beliefs about the probability distribution over roles. The concept of subjective correlated equilibrium was defined by Aumann (1974) and has been treated and refined elsewhere (e.g. Brandenberger and Dekel 1987). Subjective beliefs are natural in thinking about correlated equilibrium: a player receives a signal and interprets the signal as an indication of what others will do; it is not implausible that a player interprets a signal 'incorrectly' or that different players 'interpret signals differently'; this can be modelled with subjective beliefs. It is well known that once subjective beliefs are allowed it becomes difficult to tie down the set of correlated equilibria (Aumann 1974, 1987, Brandenburger and Dekel 1987). We thus focus on a particular form of 'naive belief'; namely we suppose that a player expects δ-substitutes to 'behave in an identical way' or more precisely to have the same probability of being allocated each role. Before defining this precisely we need some definitions.

Let $\beta^i \in P$ denote the beliefs of player i. Thus, $\beta^i(r)$ denotes the probability that player i places on the role profile being r. We say that set of beliefs $\{\beta^i\}_{i \in N}$ are a *subjective correlated ε equilibrium* if

$$\sum_{\overline{a} \in A^N} \beta^i(\overline{a}) u_i(\overline{a}) \geq \sum_{\overline{a} \in A^N} \beta^i(\overline{a}) u_i(s_i(\overline{a}_i), \overline{a}_{-i}) - \varepsilon$$

for each $i \in N$ and $s_i \in S$. This revises the definition of a correlated equilibrium by using the subjective beliefs of i.

Let Γ be a (δ, Q) class game and $\Pi = \{N_1, ..., N_Q\}$ be a partition of the player set into classes. We say that beliefs have a *common prior determined by p (CP(p))* if .

$$\beta^i(\overline{r}) = \frac{1}{|\mathcal{P}(\overline{r})|} \sum_{\overline{r} \in \mathcal{P}(\overline{r})} p(\overline{r}) \tag{12}$$

for all i and $\overline{r}$. Note that if beliefs satisfy CP then $\beta^i(\overline{r}) = \beta^i(\overline{r}')$ for any $\overline{r}' \in \mathcal{P}(\overline{r})$ and any $\overline{r}$. Thus, players expect players in the same class to have the same probability of being allocated each role; it is expected that players in the same class will behave identically. Beliefs may satisfy *CP*(p) but not be 'consistent' with the distribution p through which roles are actually allocated in game with roles Γ^p; in particular if $p_j(k) \neq p_l(k)$ for some $j, l \in N_q$ then the subjective beliefs are not consistent with p (i.e. not objectively correct) because players in the same class will not, in fact, behave identically.

It can be observed that if beliefs satisfy *CP* a player may have erroneous beliefs about his own probability of being allocated each role. We define the concept of a quasi-common prior determined by p where this is no longer the class. Recall that $h(\overline{r}, k, q)$ denotes the number of players in class N_q who have role k. Given a class N_q and role profile $\overline{r}$, let $\mathcal{P}_q(\overline{r})$ denote the set of role profiles where $\overline{r}' \in \mathcal{P}_q(\overline{r})$ if and only if $h(\overline{r}', k, q') = h(\overline{r}, k, q')$ for all $q' \neq q$ and $\overline{r}'_i = \overline{r}_i$ for all $i \in N_q$. Thus, all players within N_q have the same role but the role of players in other classes can be permuted. We say that beliefs have a *quasi-common prior determined by p (QCP(p))* if

$$\beta^i(\overline{r}) = \frac{1}{|\mathcal{P}_q(\overline{r})|} \sum_{\overline{r} \in \mathcal{P}_q(\overline{r})} p(\overline{r})$$

for all $i \in N_q$, all N_q and all $\overline{r}$. We observe that if beliefs satisfy *QCP(p)* then a player has 'correct beliefs' about the probability of role allocations within his own class but potentially erroneous beliefs about the role allocations in other groups. Each player still expects players in other classes to behave identically.

The following result demonstrates that given any correlated equilibrium p it is an approximate subjective correlated equilibrium if players have beliefs *CP*(p) or *QCP(p)*. Thus, it is an equilibrium for players to expect similar players to behave in an identical fashion even if they will not do so in reality.

Theorem 3: Let Γ be a (δ, Q)-class game, $\Pi = \{N_1, ..., N_Q\}$ a partition of the player set into classes and p a correlated equilibrium of game Γ. If beliefs satisfy *CP*(p) or *QCP(p)* then $\{\beta^i\}_{i \in N}$ is a subjective correlated ε equilibrium where $\varepsilon < 4\delta$.

Proof: We need to show that

$$\sum_{\overline{a} \in \mathcal{A}^N} \beta^i(\overline{a}) u_i(\overline{a}) \geq \sum_{\overline{a} \in \mathcal{A}^N} \beta^i(\overline{a}) u_i(s_i(\overline{a}_i), \overline{a}_{-i}) - \varepsilon - 4\delta \tag{13}$$

for all $i \in N$ and $s_i \in S$. Given that p is a correlated ε equilibrium

$$\sum_{\overline{a} \in \mathcal{A}^N} p(\overline{a}) u_i(\overline{a}) \geq \sum_{\overline{a} \in \mathcal{A}^N} p(\overline{a}) u_i(s_i(\overline{a}_i), \overline{a}_{-i}) - \varepsilon \tag{14}$$

for all $i \in N$ and $s_i \in S$. If $\beta^i(\overline{a}) > 0$ and $i \in N_q$ then from (12) there exists $j \in N_q$ such that $p_j(\overline{a}_i) > 0$. It then follows from Lemma 1 and (14) that (13) holds.■

Theorem 3 demonstrates that 'near to' any correlated equilibrium is a subjective correlated equilibrium where beliefs are satisfy CP. The converse need not hold as illustrated by returning to the 2-person game:

	A	B
A	$6,6$	$2,7$
B	$7,2$	$0,0$

Both players are 0-substitutes. Suppose $p(A,A) = \frac{1}{2}$ and $p(A,B) = \frac{1}{2}$. This distribution is not a correlated equilibrium. If beliefs satisfy CP(p) we get $\beta^1(A,A) = \frac{1}{3}, \beta^1(A,B) = \frac{1}{3}$ and $\beta^1(B,A) = \frac{1}{3}$ (with the same for player 2). This is a subjective correlated equilibrium.

With beliefs CP(p) a player has naive beliefs about the distribution of roles within his class while with QCP(p) a player has correct beliefs about the distribution of roles within his class. An intermediate possibility is that a player has 'correct beliefs' only about his own allocation to roles. Suppose that a player then expects players in his own class to behave in the same way as himself. Formally we say beliefs β^i are *introspectively consistent* if $\beta^i(\overline{r}) = 0$ whenever $p_i(\overline{r}_i) = 0$ and $\beta^i(\overline{r}) = \beta^i(\overline{r}')$ for any $\overline{r}' \in \mathcal{P}(\overline{r})$ and any $\overline{r}$. Here the close relationship between subjective and 'objective' correlated equilibrium may be broken as we can illustrate with the previous example. The distribution $p(A,B) = 1$ is a correlated equilibrium. If beliefs are introspectively consistent we get $\beta^1(A,A) = 1$ and $\beta^2(B,B) = 1$ which is not a subjective correlated equilibrium. Conversely if we consider a two person coordination game we see that the distribution $p(A,B) = 1$ is not a correlated equilibrium while beliefs that are class indiscriminate and introspectively consistent are a subjective correlated equilibrium.

7 Conclusions

A number of questions remain for future research:

- Where does the 'exogenous' probability distribution over roles come from? It may be interesting to consider an extended game where players get to propose p or choose a p. This might motivate the notion of class independence as may be difficult to get consensus on correlation across classes.

- In a model of learning, or adaptive learning (Hart 2005), does play resemble the type of correlated equilibria we have been looking at here? Suppose there is imitation or social influence/concerns for fairness, does play converge to a particular correlated equilibria?

- What happens if the p can evolve or if we have subjective beliefs and players update their beliefs somehow about p? If it is expected that similar people do similar things does it evolve that the correlated equilibrium is for similar people to do similar things or is it beliefs that change?

8 Appendix

Proof of Lemma 1: Given that $\overline{a}'$ is a permutation of $\overline{a}$ there exists a one-to-one (not necessarily unique) function β mapping N to N where $\beta(l) \in N_q$ if $l \in N_q$ and $\overline{a}'_l = \overline{a}_{\beta(l)}$ for all $l \in N$. That is, player l plays the action under profile $\overline{a}'$ that $\beta(l)$ plays under profile $\overline{a}$.

We construct a series of action profiles $\overline{a}^1, \overline{a}^2, ..., \overline{a}^n$ and functions $\beta^2, \beta^3, ..., \beta^n$ using the following iterative procedure:

- $\overline{a}^1_1 = \overline{a}_{\beta(1)}$ and $\overline{a}^1_{\beta(1)} = \overline{a}_1$ while $\overline{a}^1_z = \overline{a}_z$ for all other $z \in N$
- $\beta^2(\beta^{-1}(1)) = \beta(1)$ while $\beta^2(z) = \beta(z)$ for all other z.
- $\overline{a}^l_l = \overline{a}^{l-1}_{\beta^l(l)}$ and $\overline{a}^l_{\beta^l(l)} = \overline{a}^{l-1}_l$ while $\overline{a}^l_z = \overline{a}^{l-1}_z$ for all other z.
- $\beta^{l+1}(\beta^{l^{-1}}(l)) = \beta^l(l)$ while $\beta^{l+1}(z) = \beta^l(z)$ for all other z.

First note that $\overline{a}^1$ is a permutation of $\overline{a}'$ and $\overline{a}^l$ is a permutation of $\overline{a}^{l-1}$; this is immediate given that the only change in the action profile is an exchange of actions of players l and $\beta^l(l)$.

Next, observe there for any $l \geq 1$ there can be no $z \in N$ for whom $\beta^{l+1}(z) = l$.[3] This follows from the construction of β^{l+1} and that β is a one-to-one mapping. Thus, $\beta^{l+1}(l+1) \neq l$ and consequently there can be no $z \in N$ for whom $\beta^{l+2}(z) = l$. Continuing this argument we see that there can be no $y \geq l+1$ or $z \in N$ such that $\beta^y(z) = l$. It follows that if $\overline{a}^l_l = \overline{a}'_l$ then $\overline{a}^n_l = \overline{a}'_l$.

We next conjecture that $\overline{a}^n_l = \overline{a}'_l$ for all $l \in N$. Clearly, $\overline{a}^1_1 = \overline{a}'_1$ and so [using the argument of the proceeding paragraph] $\overline{a}^n_1 = \overline{a}'_1$. To complete the proof of the conjecture we show, by induction, that $\overline{a}^{l-1}_{\beta^l(z)} = \overline{a}_{\beta(z)}$ for all $2 \leq l \leq n$ and all $z \geq l$. Let $l = 2$. As β is a one-to-one mapping there exists a unique player $z^* = \beta^{-1}(1)$. For any $z \neq \beta^{-1}(1)$ we have $\beta^2(z) = \beta(z)$ and so clearly $\overline{a}^1_{\beta^2(z)} = \overline{a}_{\beta(z)}$. Now, for $z^* = \beta^{-1}(1)$ we have $\beta^2(z^*) = \beta(1)$. Thus,

$$\overline{a}^1_{\beta^2(z^*)} = \overline{a}^1_{\beta(1)} = \overline{a}_1 = \overline{a}_{\beta(z^*)}.$$

[3] The one exception is if $\beta(l) = l$.

This completes the $l = 2$ case. Now suppose the conjecture holds for $l > 2$. Thus, $\overline{a}^{l-2}_{\beta^{l-1}(z)} = \overline{a}_{\beta(z)}$ for all $z \geq l$. There exists a unique player $z^* = \beta^{l-1^{-1}}(l-1)$. For all $z \neq z^*$ we have $\beta^l(z) = \beta^{l-1}(z)$ and so $\overline{a}^{l-1}_{\beta^l(z)} = \overline{a}^{l-2}_{\beta^{l-1}(z)} = \overline{a}_{\beta(z)}$. For z^* we have $\beta^l(z^*) = \beta^{l-1}(l-1)$ implying

$$\overline{a}^{l-1}_{\beta^l(z^*)} = \overline{a}^{l-1}_{\beta^{l-1}(l-1)} = \overline{a}^{l-2}_{l-1} = \overline{a}^{l-2}_{\beta^{l-1}(z^*)} = \overline{a}_{\beta(z^*)}.$$

Without loss of generality let $j = 1$ [or 're-index players' so that $j = 1$]. Note that $i = \beta(j)$. Given that i and j are δ substitutes

$$\left|u_j(\overline{a}^1) - u_i(\overline{a})\right| \leq \delta.$$

Given that $\overline{a}^2_j = \overline{a}^1_j$ and players 2 and $\beta^1(2)$ are δ substitutes we have

$$\left|u_j(\overline{a}^2) - u_j(\overline{a}^1)\right| \leq \frac{\delta}{n}.$$

Iterating this argument and using $\overline{a}^n = \overline{a}'$ we obtain

$$|u_j(\overline{a}') - u_i(\overline{a})| \leq \delta + \frac{n-2}{n}\delta$$

completing the proof.■

References

[1] Aumann R.J. (1974): "Subjectivity and correlation in randomized strategies", *Journal of Mathematical Economics* 1, 67-96.

[2] Aumann, R.J. (1987) "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica* 55: 1-18.

[3] Brandenburger, A. and E. Dekel (1987) "Rationalizability and correlated equilibria," *Econometrica* 55:1391-1402.

[4] Cartwright, E.J. (2003) "Learning to play approximate nash equilibria in games with many players, " University of Warwick Department of Economics Discussion Paper 671.

[5] Friedman, D. (1996) "Equilibrium in evolutionary games: Some experimental results," *Economic Journal* 106: 1-25.

[6] Hargreaves-Heap, S. and Y. Varoufakis (2002) "Some experimental evidence on the evolution of discrimination, cooperation and perceptions of fairness," *The Economic Journal*, 112: 679-703.

[7] Harsanyi, J.C. (1973) "Games with randomly distributed payoffs: a new rationale for mixed strategy equilibrium points:" *International Journal of Game Theory* 2: 1-23.

[8] Hart, S., (2005) "Adaptive heuristics", *Econometrica* 73: 1401-1430.

[9] Kovalenkov, A. and M. Wooders (2003) "Approximate cores of games and economies with clubs," *Journal of Economic Theory* 110, 87-120.

[10] Myerson, R.B. (1997) *Game Theory: Analysis of Conflict*, Harvard University Press.

[11] Rapoport, A, D. Seale and E. Winter (2001) "Coordination and learning behavior in large groups with asymmetric players," *Games and Economic Behavior:* 1-26.

[12] Sugden, R., (1995) "A theory of focal points," *Economic Journal* 105: 533-550.

[13] Van Huyck, J., R.C. Battalio and F.W. Rankin (1997) "On the origin of convention: evidence from coordination games," *Economic Journal*, 107: 576-597.

[14] M. Wooders, E. Cartwright and R. Selten (2001) "Social conformity and equilibrium in pure strategies in games with many players," University of Warwick working paper no 589, revised. To appear in *Games and Economic Behavior.*