AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

SOCIAL CONFORMITY AND BOUNDED RATIONALITY IN ARBITRARY GAMES WITH INCOMPLETE INFORMATION: SOME FIRST RESULTS

Edward Cartwright

And

Myrna Wooders

No 672

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Conformity and bounded rationality in games with incomplete information*

Edward Cartwright
Department of Economics
University of Warwick
Coventry CV4 7AL, UK
E.J.Cartwright@warwick.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/

First version: January 2002
This version: August 2003

Abstract

It has been frequently observed, in both economics and psychology, that individuals tend to conform to the choices of other individuals with whom they identify. Can such conformity be consistent with self-interested behavior? To address this question we use the framework of games with incomplete information. For a given game we first put a lower bound on ε so that there exists a Nash ε-equilibrium in pure strategies consistent with conformity. We also introduce a new concept of conformity that allows players to conform and yet perform different actions. This is achieved by the endogenous assignment of roles to players and by allowing actions to be conditional on roles. We conclude by relating our research to some experimental literature.

*This paper is a revision of Cartwright and Wooders (2003a). The main results of this 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

A fundamental question of game theory is whether Nash equilibrium play can emerge in populations composed of boundedly rational individuals. The importance of this question follows from the limits to individual rationality and the importance of the Nash equilibrium concept in game theory and economics. We suggest that, at least in some contexts, two common elements of bounded rationality are (1) the use of pure strategies (and not mixed strategies)[1] and (2) conformity in the sense that similar individuals play similar strategies.[2] In this paper, we ask whether this behavior can be consistent with Nash equilibrium play. We do so by determining, for an arbitrary game, a lower bound on ε permitting existence of a Nash ε-equilibrium in pure strategies that is consistent with conformity. Our results are related to some experimental literature.

Informally, we say that an equilibrium is consistent with conformity if similar players play the same strategy. A principal motivation, however, of the current paper is to introduce a concept of conformity where individuals can be seen as conforming even if they are not performing the same action. Consider, for example, the division of labour within a male-female household and suppose social norms dictate that women do the housework and men go out to work. Even though we observe that men and women behave differently, we can think of all individuals as conforming to a common norm. For this reason we assume that players can endogenously create a set of roles. A player is randomly assigned a role and can make his action conditional on his role. Within a household the women takes up her role by doing the housework and the man takes up his role by going out to work. In our context, roles serve no purpose in the game other than as signalling devices and the probability distribution through which roles are assigned is determined as part of the equilibrium.

[1] It has long been recognised that motivating the use of mixed strategies is not always straightforward (Harsanyi 1973). Experimental evidence supports the view that individuals typically do not play mixed strategies (Friedman 1996) and if they do, there may be seriel correlation (Walker and Wooders 2001).

[2] Experimental evidence on conformity amongst similar individuals is long standing (e.g. Deutsch and Gerard 1955). In an economic setting there are many reasons we may expect conformity. For example, if a person is boundedly rational or has imperfect information then he may imitate a person he believes is better informed (Gale and Rosenthal 1999, Schleifer 2000). Alternatively, in a coordination game with multiple equilibria, a player may be able to make a more informed strategy choice by observing the actions of others (Scharfstein and Stein 1990, Ellison and Fudenberg 1995, Young 2001). Finally, due to normative influences, a person may be motivated by desires for prestige, popularity or acceptance or more generally, to 'fit in' with a social norm (Bernheim 1994).

When allowing endogenous roles, some additional conditions are required to retain a meaningful notion of conformity. To this end we make two assumptions on the probability distribution through which roles are assigned. We say that a class (or group) of players is conforming if: (a) any two players in the class have identical probabilities of being assigned each role; (b) the number of players in the class that will have each role is known ex-ante; and (c) any two players in the class play the same strategy. We propose that these conditions allow a class of players to be interpreted as conforming to some common norm. Condition (a) can be viewed as an anonymity requirement while (b) requires that the structure of a class – the number of players assigned each role – is determined. Together, these two conditions are quite strong.

To address whether there exists an approximate Nash equilibrium in pure strategies consistent with conformity we first introduce the notion of *approximate substitute players* of a non-cooperative game. We also define the concept of a (δ, Q)-class game. A (δ, Q)-class game has the property that the player set can be partitioned into Q *classes* of players. Players in the same class are approximate substitutes and the dissimilarity of players in a class is bounded by the parameter δ. A (δ, Q)-class game is also required to have the property that, when a 'small' number of players from each class change strategy, payoffs for the remaining players change by less than δ. We note that any finite game is a (δ, Q)-class game for any Q and some, perhaps large, δ. A feature, therefore, of the (δ, Q)-class game framework is that it allows us to draw conclusions on arbitrary games (and an arbitrary level of conformity as measured by Q). Moreover, our conclusions can be stated in terms of the parameters describing the games, δ and Q.

Given a (δ, Q)-class game we say that a Nash equilibrium is consistent with conformity if players within every class are conforming (in the sense of a-c above). Our main result can be summarized:

If Γ is a (δ, Q) class game then, for some probability distribution over roles, for any $\varepsilon \geq 10\delta$ there exists a Nash ε-equilibrium in pure strategies consistent with conformity.

The value 10δ could be interpreted as a bound on the distance from full rationality of players using pure strategies and conforming. Two preliminary results treat conformity and the use of pure strategies in isolation. Our first result can be summarized:

If Γ is a (δ, Q) class game then for any $\varepsilon \geq 4\delta$ there exists a Nash ε-equilibrium in pure strategies.

Our second result can be summarized:

If Γ is a (δ, Q) class game then for any $\varepsilon \geq 4\delta$ there exists a Nash ε-equilibrium consistent with conformity.

Our results allow us to draw conclusions for arbitrary games on the individual rationality of conformity and the use of pure strategies. It is crucial, however, to have some general examples of (δ, Q) class games for arbitrary values of δ. In a companion paper (Cartwright and Wooders 2003b) we connect the concept of games with approximate substitutes to games with many players derived from a noncooperative 'pregame'.[3] This allows us to apply the results of this paper to games with many players and demonstrate that, for a broad family of games, the bound on ε required for existence of a Nash ε-equilibrium with conformity can be arbitrarily small.

In the following subsection we provide some motivating examples. Subsequently we proceed as follows: Section 2 introduces notation and Section 3 defines a (δ, Q)-class game. Section 4 treats purification and Section 5 conformity. Section 6 looks at conformity in pure strategies. In Section 7 we discuss our findings in the light of some experimental literature before looking at other related literature in Section 8. Section 9 concludes and remaining proofs are contained in an Appendix.

1.1 Motivating Examples

We provide three simple examples to demonstrate that there need not exist an approximate Nash equilibrium in pure strategies consistent with conformity. We being with an example showing that there need not exist an approximate Nash equilibrium in pure strategies.

Example 1: Players are characterized as *rich* or *poor*. Each player has chooses one of two locations A or B. Poor players prefer living with rich players. Thus, the payoff of a poor player is equal to the proportion of rich players who choose the same location as himself. Rich players prefer to avoid

[3] A noncooperative pregame takes as given a metric space Ω of player attributes and a finite set S of pure strategies. An attribute $\omega \in \Omega$ is interpreted as specifying the characteristics of a player and the metric on attribute space Ω allows measurement of similarity of players. A universal payoff function h is also taken as given. These three elements, Ω, S, and h, constitute a pregame. Given a finite set of players and an attribute function, ascribing a point in attribute space to each player, a pregame induces a game, according to standard definitions, on the player set. A pregame thus allows us to model a family of games all induced from a common strategic structure.

poor players. Thus, the payoff of a rich person is equal to the proportion of poor players who choose a different location to himself.

It can easily be seen that games of the form outlined in Example 1 need not have a pure strategy Nash equilibrium. Suppose, for example, that there is only one rich player and many poor players; given any pure strategy vector some player can gain by at least one from changing his strategy. Note, however, that there does exist a Nash equilibrium consistent with conformity - 'everybody choose location A with probability one half and location B with probability one half' is a Nash equilibrium and has the property that every player plays the same strategy.

In the second example we demonstrate that there need not exist an approximate Nash equilibrium consistent with a meaningful notion of conformity, which at a minimum should require that at least some players choose the same or similar actions.

Example 2: There are n players and n locations. Each player chooses a location. Every player has a unique preferred location. If a player chooses his preferred location he gets a payoff of one; otherwise he gets a payoff of zero.

Games of the form outlined in Example 2 trivially have a Nash equilibrium in pure strategies. They need not, however, have an approximate equilibrium consistent with conformity since preferences may be too diverse. Suppose, for example, that the preferred location of each player i is location i. Thus, each player has a different preferred location.

Example 3 combines Examples 1 and 2 to demonstrate why there need not exist a Nash equilibrium in pure strategies consistent with conformity.

Example 3: Players are characterized as *rich* or *poor*. There are two *countries* A and B. In each country there are n *towns* (thus there are $2n$ towns in total). Each player chooses a town (and thus indirectly a country). Rich players prefer to avoid poor players. Thus, the payoff of a rich person is equal to the proportion of poor players whose choose a different *country* to himself. The payoff of a poor player is composed of two elements. A poor player prefers living in the same country as rich players. Thus, assume the payoff of a poor player is equal to the proportion of rich players who choose the same country as himself. In addition, in each country a poor player has a unique preferred town; a poor player receives an increment to his payoff of one if he lives in his preferred town (and no increment otherwise).

Suppose that there is one rich player and n poor players. Further, suppose that in each country no two poor players have the same preferred town. In this example there does not exist an approximate Nash equilibrium in pure strategies for the same reasons as in Example 1. There also does not exist a Nash equilibrium consistent with conformity for the same reasons as in Example 2.

2 A Bayesian Game - definitions and notation

A *Bayesian game* Γ is given by a tuple (N, A, T, p, u) where N is a finite *player set*, A is a set of *action profiles*, T is a set of *type profiles*, p is a *set of player beliefs* and u is a set of *utility functions*. Before formally defining each of these components a brief overview may be helpful. The game Γ randomly assigns (or 'Nature' assigns) each player a type. Informed of his own type but not the types of his opponents, each player chooses an action. His payoff depends on the actions and types of all players, including himself. Since a player does not know the types of other players he bases his action choice on beliefs about the types of others.

Let $N = \{1, ..., n\}$ be a finite player set and let $\mathcal{T}$ be a finite set of *types*. Let $T = \mathcal{T}^N$ be the set of *type profiles* where, given type profile $t = (t_1, ..., t_n)$, the type of player i is t_i. We say that a game is *a game of perfect information* if $|\mathcal{T}| = 1$. Types are randomly assigned to players by Nature. Each player $i \in N$ forms beliefs about the type profile as given by function p_i where $p_i(t_{-i}|t_i)$ denotes the probability that player i assigns to type profile (t_i, t_{-i}) given that he is of type t_i. We note that $\sum_{t_{-i} \in T_{-i}} p_i(t_{-i}|t_i) = 1$ for each $t_i \in \mathcal{T}$. Let $p = \{p_1, ..., p_n\}$ be the set of beliefs about the type profile. We say that there are *consistent beliefs* if there exists a probability distribution g over the set of type profiles such that

$$p_i(t_{-i}|t_i) = \frac{g(t_i, t_{-i})}{\sum_{t'_{-i} \in T_{-i}} g(t_i, t'_{-i})} \tag{1}$$

for all $i \in N$ and $t_i \in \mathcal{T}$.[4] Thus, informally stated, there is consistent beliefs if every player has the same beliefs over the type profile.

There exists a finite set $\mathcal{A}$ of *actions*. Let $A = \mathcal{A}^N$ be the set of *action profiles* where for action profile $a = (a_1, ..., a_n)$ the action of player is given by a_i. For each player $i \in N$ there exists a *utility function* $u_i : A \times T \to \mathbb{R}$.

[4] We assume that the denominator of (1) is always positive - i.e. there is positive probability that a player $i \in N$ will be of each type $t_i \in \mathcal{T}$.

The interpretation is that $u_i(a,t)$ denotes the payoff of player i if the action profile is a and the type profile t. Let $u = \{u_1, ..., u_n\}$ denote the set of player utility functions.

Knowing his own type, a player chooses an action. Thus, a *pure strategy* details the action a player will take for each type $t^z \in \mathcal{T}$ and is given by a function $s^k : \mathcal{T} \to \mathcal{A}$ where $s^k(t^z)$ is the action played by the player if he is of type t^z. Denote the set of pure strategies by $\mathcal{S}$ where we let $K = |\mathcal{A}|^{|\mathcal{T}|} = |\mathcal{S}|$ denote the number of pure strategies.

A (mixed) *strategy* is given by a probability distribution over the set of pure strategies. The set of strategies is thus $\Delta(\mathcal{S})$. Given a strategy x we denote by $x(s^k)$ the probability that a player plays pure strategy $s^k \in \mathcal{S}$. We say that pure strategy $s^k \in support(x)$ if and only if $x(s^k) > 0$. We denote by $x(a^l|t^z)$ the probability that player i plays action a^l given that he is of type t^z. We note that $\sum_{a^l \in \mathcal{A}} x(a^l|t^z) = 1$ for all $t^z \in \mathcal{T}$. Let $\Sigma = \Delta(\mathcal{S})^N$ denote the set of *strategy vectors*. We refer to a strategy vector m as a *degenerate* if for all $i \in N$ and $t^z \in \mathcal{T}$ there exists some a^l such that $m_i(a^l|t^z) = 1$.

We assume that players are motivated by expected payoffs. Recall that a player chooses an action knowing his own type but not those of the complementary player set. Consider a player $i \in N$. Given a strategy vector σ, a type $t^z \in \mathcal{T}$ and beliefs about the type profile p_i the probability that player i puts on the action profile-type profile pair $a = (a_1, ..., a_n)$ and $t = (t_1, ..., t_{i-1}, t^z, t_{i+i}, ..., t_n)$ is given by:

$$\Pr(a, t_{-i}|t^z) = p_i(t_{-i}|t^z)\sigma_1(a_1|t_1)...\sigma_i(a_i|t^z)...\sigma_n(a_n|t_n).$$

Thus, given any strategy vector σ, for any type $t^z \in \mathcal{T}$ and any player i of type t^z, the expected payoff of player i can be calculated. Let $U_i(\cdot|t^z) : \Sigma \to \mathbb{R}$ denote the expected utility function of player i conditional on the type of player i being t^z where:

$$U_i(\sigma|t^z) = \sum_{a \in A} \sum_{t_{-i} \in T_{-i}} \Pr(a, t_{-i}|t^z) u_i(a, t_z, t_{-i}).$$

The standard definition of a Bayesian Nash equilibrium applies. A strategy vector σ is a *Bayesian Nash ε-equilibrium* if and only if:

$$U_i(\sigma_i, \sigma_{-i}|t^z) \geq U_i(s^k, \sigma_{-i}|t^z) - \varepsilon$$

for all $s^k \in \Delta(\mathcal{S})$, all $t^z \in \mathcal{T}$ and for all $i \in N$. We say that a Bayesian Nash ε equilibrium m is a *Bayesian Nash ε-equilibrium in pure strategies* if m is degenerate.

3 Approximate substitutes

Given a Bayesian game $\Gamma = (N, T, A, p, u)$ 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. This requires us to formulate a metric by which to compare players.[5] 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. In contrast, we say that players i and j are individual substitutes if they have similar payoff functions.[6] Combining both measures together, we refer to players i and j as *approximate substitutes* if they are both interaction and individual substitutes.

Take as given a Bayesian game $\Gamma = (N, A, T, p, u)$. A partition $\{N_1, ..., N_Q\}$ of N is a *δ-interaction substitute partition* when, for any two strategy vectors $\sigma^1, \sigma^2 \in \Sigma$, if:

$$\sum_{i \in N_q} \sigma_i^1(s^k) = \sum_{i \in N_q} \sigma_i^2(s^k),$$

for all N_q and all $s^k \in \mathcal{S}$, then:

$$\left| U_i(x, \sigma_{-i}^1 | t^z) - U_i(x, \sigma_{-i}^2 | t^z) \right| \leq \delta \tag{2}$$

for any player $i \in N$, and $t^z \in \mathcal{T}$ and any strategy $x \in \Delta(\mathcal{S})$.

A partition $\{N_1, ..., N_Q\}$ is a *δ-individual substitute partition* if, for any N_q, for any two players $i, j \in N_q$ and for any strategy vector $\sigma \in \Sigma$ such that $\sigma_i = \sigma_j$:

$$|U_i(x, \sigma_{-i} | t^z) - U_j(x, \sigma_{-j} | t^z)| \leq \delta \tag{3}$$

for any strategy $x \in \Delta(\mathcal{S})$.

We say that a partition $\{N_1, ..., N_Q\}$ is a *δ-substitute partition* if $\mathcal{N}$ is both a δ-interaction substitute partition and a δ-individual substitute partition. In this instance we say that two players belonging to a subset N_q

[5] For comparison, we note that in cooperative game theory the distance between players is typically measured by the maximum difference in value that players can add to coalitions. Thus, in cooperative game theory, a δ-substitute partition has the property that, given any coalition structure, 'swapping' players who are δ-substitutes, between coalitions, has an effect of less than δ on the worth of the coalitions (for example, Kovalenkov and Wooders 2003).

[6] The distinction between interaction substitutes and payoff substitutes is suggestive of distinction between crowding types and taste types (cf., Conley and Wooders 2001 and references there). This distinction also appears in Wooders, Cartwright, and Selten (2001 revised) for noncooperative games.

are δ-substitutes for each other. We note that the partition into singletons $\{\{1\},...,\{n\}\}$ is a 0-substitute partition. That is, each player is a 0-substitute for themselves. We also note that for any Bayesian game Γ and any $Q \leq N$ there exists a δ-substitute partition for some finite $\delta \geq 0$.

3.1 A (δ, Q)-class Bayesian game

We define a third type of partition. Roughly, this type of partition puts a bound on the change in payoffs when QK players, K from each class, change strategy. Formally, a partition $\{N_1, ..., N_Q\}$ is a *δ-strategy switching partition* when, for any two strategy vectors $\sigma^1, \sigma^2 \in \Sigma$, if:

$$\sum_{i \in N_q} \left| \sigma_i^1(s^k) - \sigma_i^2(s^k) \right| \leq 1, \tag{4}$$

for each N_q and all $s^k \in \mathcal{S}$ then:

$$\left| U_i(x, \sigma_{-i}^1 | t^z) - U_i(x, \sigma_{-i}^2 | t^z) \right| \leq \delta$$

for any player $i \in N$, all $t^z \in \mathcal{T}$ and any strategy $x \in \Delta(\mathcal{S})$.

We say that a Bayesian game Γ is a *$(\delta_I, \delta_P, \delta_C, Q)$-class Bayesian game* if there exists a Q member partition $\{N_1, ..., N_Q\}$ that is a δ_I-interaction substitute partition, a δ_P-individual substitute partition and a δ_C-strategy switching partition. We call such a partition *congruent*. We say that a Bayesian game Γ is a (δ, Q)-class Bayesian game if Γ is a *$(\delta_I, \delta_P, \delta_C, Q)$-class Bayesian game* where $\delta_I, \delta_P, \delta_C \leq \delta$. We refer to each N_q as a class of player and say that two players $i, j \in N_q$ are the same class of player. It can easily be seen that for any game Γ and for any Q the minimum value of δ such that Γ is a (δ, Q)-class game can, in principal, be calculated.

3.2 Discussion

We have already noted that the partition of N into singletons $\{\{1\},...,\{n\}\}$ is a 0-substitute partition. Note, however, that finding a δ for which there exists a δ-strategy switching partition would require comparing strategy vectors σ^1 and σ^2 in which any player $i \in N$ can change her strategy any way she wishes (from 4). Thus, unless the game is trivial, it cannot be a (δ, N)-class Bayesian game for any meaningful value of δ.

In contrast, consider the opposite extreme of a partition of the player set N into one set $\{\{1,...,n\}\}$. It is unlikely that this is a δ-substitute partition for a small value of δ. For it to be so would require that all players

have similar payoff functions (from 2) and that payoffs depend only on the 'population average' or the number of players playing each strategy (from 3). The partition may, however, be a δ-strategy switching partition for small δ as (4) is more restrictive the smaller is Q.

Between these two extremes we clearly find a trade off in trying to find the minimum δ for which a game Γ is a (δ, Q)-class Bayesian game. The larger the number of classes Q then the larger is likely to be the minimum value of δ for which there exists a δ-strategy switching partition. Conversely the larger is Q the smaller is likely to be the δ for which there exists a δ-substitute partition.[7]

In defining a (δ, Q)-class Bayesian game the role of incomplete information is not explicit. We highlight the following issue: if two players i and j use the same or similar strategies then this does not necessarily imply that their expected actions and types are similar. This is because the distribution over types of players i and j may differ. Thus, even though players i and j play similar strategies, expectations about their types and their actions may be dissimilar.

For example, suppose that there are two players, 1 and 2, in a class N_q. Player 1 is always of type t^1 and player 2 is always of type t^2. There are two actions a^1 and a^2. Let σ^1 be such that player 1 plays the strategy 'if type t^1 play action a^1 and if type t^2 play action a^2' and player 2 plays the strategy 'if type t^1 play action a^2 and if type t^2 play action a^1'. Consider a strategy σ^2 in which both players exchange strategies. It is clear that even though the 'aggregate strategy of the class' is invariant between strategy vectors σ^1 and σ^2 the expectations of what will happen, in terms of actions, change completely.

As the above suggests, the definition of a δ-interaction substitute partition, and also a δ-strategy switching partition, implicitly measures the variability in beliefs over types and the importance of such variations on payoffs. In particular, if a game is to be a (δ, Q)-class Bayesian game for small δ, then we would expect that either players of the same class are believed to have similar distributions over types or payoffs are relatively invariant to the type profile. This, however, seems reasonable; an assumption of common priors, for example, makes such issues irrelevant.

[7] This is discussed in more detail by Cartwright and Wooders (2003).

4 Purification of mixed strategies

Our first result places a bound on the distance from rationality of using pure strategies. The result requires two lemmas. Since these are quite technical and require additional definitions, we place them and the proof of Theorem 1 in the Appendix.[8]

Theorem 1: Let $\Gamma = (N, A, T, p, u)$ be a $(\delta_I, \delta_P, \delta_C, Q)$-class Bayesian game. Let ε be a positive real number where $\varepsilon \geq 2(\delta_I + \delta_C)$. Then the game Γ has a Bayesian Nash ε-equilibrium in pure strategies.

We note that the value of δ_P has no effect on the bound for which there exists a Bayesian Nash ε-equilibrium. That is, the existence of an approximate Bayesian Nash equilibrium in pure strategies does not require that players in the same class should have similar payoff functions. This will not be the case when we consider conformity. Note also that Q does not appear in the lower bound on ε. However, as discussed above, the parameters δ_I and δ_C depend on Q so dependence of ε on Q is implicit.

5 Conformity

Our second result treats the distance from rationality of conformity. We highlight that in this section no assumption is made that players use pure strategies. We first define the notion of a strategy vector consistent with conformity. Take as given a partition $\{N_1, ..., N_Q\}$ of player set N into classes. We say that a Bayesian Nash ε-equilibrium m is a Bayesian Nash ε-equilibrium *consistent with conformity* if and only if $m_i = m_j$ for all $i, j \in N_q$ and for each N_q. Thus, a strategy vector consistent with conformity has the property that any two players in the same class play the same strategy.

Theorem 2: Let Γ be any $(\delta_I, \delta_P, \delta_C, Q)$-class Bayesian game. Let ε be a positive real number where $\varepsilon \geq 2(\delta_I + \delta_P)$. The game Γ has a Bayesian Nash ε equilibrium m consistent with conformity.

[8] There is a growing literature on purification of mixed strategy equilibria in large games; see, for example, Schmeidler (1973), Mas-Colell (1984), Pascoa (1993), Khan, Rath and Sun (1997), Kalai (2000) and Wooders, Cartwright and Selten (2001). All of these papers treat games with a continuum of players or provide asymptotic results. Our result differs in that we treat an arbitrary game and obtain a bound on ε for the existence of an ε-equilibrium in terms of the parameters describing the games.

Proof: By Nash's Theorem there exists a Bayesian Nash equilibrium σ^* of the game Γ. Given that Γ is a $(\delta_I, \delta_P, \delta_C, Q)$-class game there is a congruent partition of N, say $\{N_1, ..., N_Q\}$. For each N_q and for each $s^k \in \mathcal{S}$ let $\sigma^*(q,k)$ be defined as

$$\sigma^*(q,k) = \frac{1}{|N|} \sum_{i \in N_q} \sigma_i^*(s^k).$$

Consider a strategy vector m satisfying the property that, for all $i \in N$, if $i \in N_q$ then $m_i(s^k) = \sigma^*(q,k)$. Clearly m is consistent with conformity. Given that σ^* is a Bayesian Nash equilibrium and that $\mathcal{N}$ is a δ_I-interaction substitute partition:

$$U_i(y, m_{-i}|t^z) \geq U_i(x, m_{-i}|t^z) - 2\delta_I$$

for all $x, y \in \Delta(\mathcal{S})$ where $support(y) \subset support(\sigma_i^*)$ and for all $t^z \in \mathcal{T}$. Given that $\mathcal{N}$ is a δ_P-individual substitute partition:

$$|U_i(x, m_{-i}|t^z) - U_j(x, m_{-j}|t^z)| \leq \delta_P$$

for any players $i, j \in N_q$ (for some q), all $t^z \in \mathcal{T}$ and $x \in \Delta(\mathcal{S})$. We note, that for any player $i \in N_q$ and for pure strategy $s^k \in support(m_i)$ there exists some player $j \in N_q$ such that $s^k \in support(\sigma_j^*)$. Thus:

$$U_i(m_i, m_{-i}|t^z) \geq U_i(x, m_{-i}|t^z) - 2(\delta_I + \delta_P)$$

for all $i \in N$, $t^z \in \mathcal{T}$ and $x \in \Delta(\mathcal{S})$. This completes the proof.■

Theorem 2 shows that if Γ is a (δ, Q) class game then for any $\varepsilon \geq 4\delta$ there exists a Bayesian Nash ε-equilibrium with the property that any two players belonging to the same class play the same strategy. We note that the value of δ_C is irrelevant for the bound on which there exists a Bayesian Nash ε-equilibrium consistent with conformity. We also note that Theorem 2 encompasses the special case in which $Q = n$. In this case, there exists a 0-substitute partition and so there exists a Bayesian Nash equilibrium inducing n classes. This is, of course, just an immediate application of the Nash Existence Theorem. This observation makes clear that we need to have some notion of how large classes need to be. In the next section we address this issue by introducing ex-post stability. In Cartwright and Wooders (2003b) we take an alternative approach by considering a family of arbitrarily large games where the number of classes can be fixed independent of the population size.

6 Conformity in pure strategies

As motivated in the introduction we wish to consider the possibility that players conform in strategy yet chose different actions. The existence of imperfect information makes this possible since every player may play the same pure strategy and yet the actions of players (which are conditional on type) differ. This 'purification' of a mixed strategy is only possible, however, if there is 'enough' uncertainty over types; it is not, for example, possible in games of complete information. We thus consider a possible exogenous creation of incomplete information.

Take as given a game of perfect information Γ with player set N, set of action profiles A and a set of utility functions $\{u_i\}_{i \in N}$ where $u_i : A \to \mathbb{R}$. Assume that there exists a set of roles $\mathcal{R} = \{r^1, ..., r^K\}$.[9] Let $R = \mathcal{R}^N$ be the set of *role profiles*. Take as given a *probability distribution f over the set of role profiles R* where $f(r)$ denotes the probability of role profile r. We consider a *Bayesian game with endogenous roles* Γ^f. In game Γ^f roles are types. Thus, roles are randomly allocated to players, a player can make his action choice conditional on his role and makes his choice of action knowing his role but not those of players in the complementary player set. A player's payoff, however, does not depend directly on the role profile. We assume that players have consistent beliefs with respect to the distribution over roles f. Formally, we can define game $\Gamma^f = (N, A, T^f, p^f, u^f)$ is defined to satisfy:

1. $T^f \equiv R$,
2. for all $r \in R$,
$$p_i^f(r_{-i}|r_i) = \frac{f(r_i, r_{-i})}{\sum_{r'_{-i} \in R_{-i}} f(r'_{-i}, r_i)}$$
3. $u_i^f(a, r) = u_i(a)$ for all $a \in A$, $r \in R$ and all $i \in N$.

Condition 1 states that roles are equivalent to types. Condition 2 states that players have consistent beliefs with respect to the distribution of roles. Condition 3 states that payoffs are not directly effected by the role profile.

As before, given a partition into classes $\{N_1, ..., N_Q\}$, we say that an equilibrium m of game Γ^f is consistent with conformity if and only if $m_i = m_j$ for any $i, j \in N_q$. Given, however, that actions can be made conditional on roles, is the fact that players play the same strategy sufficient to suggest

[9] The number of roles is as large as the number of actions.

conformity? Generally, we would suggest not; the examples of Section 1.1 allow us to demonstrate why this is so and to motivate two criteria we impose on the probability distribution over roles.

Our first condition requires that any two players within a class, in addition to playing the same strategy, must have the same probability of being assigned each role. One might think of this as an 'equal opportunity' condition within classes.

Within class anonymity: A probability distribution over roles f satisfies *within class anonymity* if the probability that a player from a class N_q will have role r^k is (a priori) identical for all players belonging to that class. Formally, if $i, j \in N_q$ for some q then:

$$\sum_{r \in R: r_i = r^k} f(r) = \sum_{r \in R: r_j = r^k} f(r)$$

for all $r^k \in \mathcal{R}$.

To motivate this condition, we reconsider Example 2. With endogenous roles, the set of strategies can be as large as the set of possible actions so, even though all players may choose the same strategy, this does not suggest any meaningful conformity. Consider a pure strategy of the form 'if role i choose location i' and a probability distribution over roles that says player i is always role i. Clearly, there is really no conformity. This type of problem is overcome by the condition of within class anonymity since all the players in each class must be assigned the same probability distribution over roles.

To motivate the next requirement, consider again the example in the introduction of a male-female household following the roles of 'he goes out to work, she stays home'. For this norm to be successful, it is necessary that no player, knowing the structure of society – the number of players with each role in his or her class – after roles are assigned, wishes to change role assignment. Thus we impose the following condition:

Within class determination: Given a role profile r let $h(r, k, q)$ be the number of players in class N_q who have role r^k. A probability distribution over roles f is *within class determined* if for any class N_q and for any two role profiles r and $\overline{r}$, if $f(r), f(\overline{r}) > 0$ then $h(r, k, q) = h(\overline{r}, k, q)$ for all classes q and for all $r^k \in \mathcal{R}$.

Now reconsider Example 1 and suppose there is only one rich player. Let us assume that the rich player constitutes a class. Further, we assume that

there are two roles A and B, the rich player chooses location A if and only if allocated role A and the probability distribution over roles is such that the rich player is allocated role A with probability one half. This permits a Nash equilibrium in pure strategies. Clearly, however, it is difficult to justify this as an equilibrium in pure strategies. Within class determination overcomes this problem since it requires that ex-ante the number of players in each class who will have each role should be fixed.

In sum, within class anonymity requires that each player in a class has an equal probability of being allocated each role. Within class determination implies that the number of players who have each role can be known with certainty ex ante - the only uncertainty is who will have each role. These are strong requirements on f. We propose they capture the notion that players in the same class who play the same strategy are conforming to some norm of behavior.

Before stating our final result we introduce one further definition. To overcome the problem of class (or society) size we use the concept of ex-post information proofness as introduced by Kalai (2000). Information proofness implies that, knowing the action profile and the type profile, no player has a strong incentive to change her own action. Formally, an action profile, type profile pair a, t is said to be ε *information proof* if for all $i \in N$:

$$u_i(a, t) \geq u_i(a^k, a_{-i}, t) - \varepsilon$$

for all $a^k \in \mathcal{A}$. A strategy profile σ is said to be a *Bayesian Nash ε information proof equilibrium* if it yields an ε information proof action profile, type profile pair with probability one. Note that if a strategy profile σ is a Bayesian Nash ε information proof equilibrium then there exists a Bayesian Nash information proof ε-equilibrium that is degenerate. Moreover, there exists such an equilibrium with the property that each player i is choosing some pure strategy in the support of σ_i. If a strategy vector is a Bayesian Nash information proof equilibrium then, as discussed further by Kalai (2000), no player would wish to change his action after knowing the types (or roles) and the actions of the other players.

We now state our final result.

Theorem 3: Let Γ be any $(\delta_I, \delta_P, \delta_C, Q)$-class Bayesian game with perfect information. Let ε be a positive real number where $\varepsilon \geq 2(2\delta_I + \delta_P + 2\delta_C)$. Then there exists a partition of N into Q classes and a Bayesian game with endogenous roles $\Gamma(f)$ (where f is within class anonymous and within class determined) such that $\Gamma(f)$ has a Bayesian Nash ε information proof equilibrium m in pure strategies that is consistent with conformity.

Proof: From Theorem 1 the game Γ has a Bayesian Nash $2(\delta_I + \delta_C)$ equilibrium in pure strategies m^*. Given that game Γ is a $(\delta_I, \delta_P, \delta_C, Q)$-class game let $\mathcal{N} = \{N_1, ..., N_Q\}$ be a partition of N into Q classes. For any degenerate strategy profile $\sigma \in \Sigma$ define

$$h(\sigma, k, q) = \Big|\{i \in N_q : \sigma_i(s^k) = 1\}\Big|.$$

Given strategy vector m^* let f be such that for any role profile $r \in R$, if $f(r) > 0$ then $h(r, k, q) = h(m^*, k, q)$. Let every player $i \in N$ play the strategy (for Bayesian game with endogenous roles $\Gamma(f)$) 'play pure strategy s^k (from game Γ) if of role r^k'. We claim that this gives the desired equilibrium. This can be demonstrated by showing that any action profile, type profile pair that can arise with positive probability (in game $\Gamma(f)$) is ε-information proof.

Consider any degenerate strategy profile $\overline{m}$ such that $h(\overline{m}, k, q) = h(m^*, k, q)$. Pick an arbitrary player $i \in N_q$. Define $y = \overline{m}_i$. There must exist some player $j \in N_q$ where $m^*_j = y$ and where

$$U_j(m^*_j, m^*_{-j}|t^z) \geq U_j(x, m^*_{-j}|t^z) - 2(\delta_I + \delta_C)$$

for all $x \in \Delta(\mathcal{S})$. Consider the strategy vector $\overline{\overline{m}}$ where $\overline{\overline{m}}_i = y = m^*_j$ and $\overline{\overline{m}}_l = m^*_l$ for all other $l \in N$. Given that $\mathcal{N}$ is a δ_C-strategy switching partition

$$U_j(\overline{\overline{m}}_j, \overline{\overline{m}}_{-j}|t^z) \geq U_j(x, \overline{\overline{m}}_{-j}|t^z) - 2(\delta_I + 2\delta_C)$$

for all $x \in \Delta(\mathcal{S})$ and $t^z \in \mathcal{T}$. Given that $\mathcal{N}$ is a δ_P-individual substitute partition and that $\overline{\overline{m}}_i = \overline{\overline{m}}_j$,

$$\big|U_j(x, \overline{\overline{m}}_{-j})|t^z) - U_i(x, \overline{\overline{m}}_{-i}|t^z)\big| \leq \delta_P$$

for all $x \in \Delta(\mathcal{S})$ and $t^z \in \mathcal{T}$. Thus,

$$U_i(\overline{\overline{m}}_i, \overline{\overline{m}}_{-i}|t^z) \geq U_i(x, \overline{\overline{m}}_{-i}|t^z) - 2(\delta_I + \delta_P + 2\delta_C)$$

for all $x \in \Delta(\mathcal{S})$ and $t^z \in \mathcal{T}$. Note, however, that this implies

$$U_i(\overline{m}_i, m^*_{-i})|t^z \geq U_i(x, m^*_{-i}|t^z) - 2(\delta_I + \delta_P + 2\delta_C)$$

for all $x \in \Delta(\mathcal{S})$ and $t^z \in \mathcal{T}$. Given that $\mathcal{N}$ is a δ_I-interaction substitute partition,

$$\big|U_i(x, m^*_{-i}|t^z) - U_i(x, \overline{m}_{-i}|t^z)\big| \leq \delta_I$$

for all $x \in \Delta(\mathcal{S})$ and $t^z \in \mathcal{T}$. Thus,

$$U_i(\overline{m}_i, \overline{m}_{-i})|t^z) \geq U_i(x, \overline{m}_{-i}|t^z) - 2(2\delta_I + \delta_P + 2\delta_C)$$

for all $x \in \Delta(\mathcal{S})$ and $t^z \in \mathcal{T}$. This completes the proof.■

7 Some Experimental Evidence

The purpose of this section is to relate our analysis to some experimental evidence on learning in coordination games (with 'many' players). The relevant literature includes, amongst other works, Van Huyck, Battalio and Beil (1990, 1991), Van Huyck, Battalio, Mathur, Ortmann and Van Huyck (1995), Friedman (1996), Van Huyck, Battalio and Rankin (1997) and Hargreaves-Heap and Varoufakis (2002).

A typical experiment proceeds over a number of distinct periods and revolves around a stage game such as the following two player symmetric Hawk-Dove game:

	H	D
H	-2	8
D	0	4

At the beginning of a period subjects simultaneously choose either strategy H or D. Each player is then matched to play the Hawk-Dove game against each of his opponents and receives the mean average payoff.[10] In a 'one population' case each player is a possible opponent for any other player. Thus, in every period all combinations of players are matched to play the Hawk-Dove stage game. In a 'two population' case the player set is split into a population of 'row players' and a distinct population of 'column players'. The possible opponents for a 'row player' are the set of 'column players' and vice versa (Friedman 1996).[11]

It is simple to calculate that the one population case is a $(0, 0, \frac{4}{n}, 1)$-class Bayesian game while the two population case is a $(0, 0, \frac{8}{n}, 2)$-class Bayesian game, where n is the number of subjects. In applying Theorems 1, 2 and 3 we see that in large populations there exists an approximate Nash equilibrium in pure strategies consistent with conformity. The Nash equilibria can, of course, be explicitly calculated. A Nash equilibrium takes one of three basic forms that, assuming $n = 12$, can be summarized: (1) A *'mixed strategy'* equilibrium where every subject plays strategy - 'play H with probability $\frac{2}{3}$

[10] A variety of other treatments are possible: including random matching whereby a subject is matched to play just one randomly chosen opponent in each period (Friedman 1996).

[11] Again, there are a variety of alternative two population models we might consider. Hargreaves and Heap (2002), for example, endow each player with either the color blue or red. Each subject plays every other subject but knows the colour of his opponent and can make strategy conditional on that colour. Van Huyck, Battalio and Rankin (1997) consider other possibilities.

and play D with probability $\frac{1}{3}$'. (2) A *'pure strategy'* equilibrium of which one example is - '8 players play H and 4 players play D'. (3) In the two population case, a *'discriminatory equilibrium'* of which one example is - 'all row players choose H and all column players choose D'.

The analysis of this paper would suggest some basic hypothesis: (i) Approximate Nash equilibrium play is observed in the long run; we would expect this given that both the use of pure strategies and conformity are consistent with Nash equilibrium play. (ii) A pure strategy equilibrium is 'more likely' to emerge than a mixed strategy equilibrium.[12] Indeed, in a pure strategy equilibrium we would view players as taking a role - either Hawk or Dove. Note, however, that in the one population case there is no signal or mechanism to guide role distribution. In contrast, in the two population case, there is a clear distinction between two population groups that may guide the distribution of roles. Reflecting on this we would hypothesize that: (iii) The discriminatory equilibrium will emerge in the two population case as a player responds to his assignment to a population as an assignment of role. (iv) A player is more likely to be seen to use a pure strategy in the two population case than the one population case because he can more easily identify his role.

The existing experimental literature allows us to draw some conclusions on the above hypotheses. Basically, our hypothesis do not seem inconsistent with the data. Before discussing the literature in a little more detail it may prove useful and interesting to provide data on two representative experimental runs taken from Friedman (1996). Figure 1 (below) provides details on the evolution of play in a run of the one population case and Figure 2 provides details for a two population case.

We see in Figure 1 the number playing D is never more than 2 away from the Nash equilibrium level of 4 and is typically within 1. Thus, Nash-like play has emerged. It is clear that the mixed strategy equilibrium is not a good characterization of play; six players, for example, are observed to play the same strategy, either H or D, in ten or more consecutive periods of the experimental run. In comparison, the pure strategy equilibrium does not seem an unreasonable characterization of play.[13]

[12] In reality, we would expect at least some subjects to change strategy (and thus not 'strictly' play a pure strategy) due to motivations such as curiosity or boredom.

[13] We highlight that some change of stage game strategy is always to be expected due to experimentation and the like. Also some changes in strategy are to be expected, under our hypothesis, while the assignment of roles is emerging.

PLAYER \ PERIOD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
3	•	•	□	•	□	•	•	•	•	□	•	•	□	•	•	•
4	•	•	□	□	□	•	□	□	□	□	□	□	□	□	□	□
5	□	•	□	•	•	•	•	□	□	•	•	•	□	•	•	•
6	•	•	•	•	•	•	□	□	•	•	•	•	•	•	•	•
7	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
8	□	□	•	□	□	•	•	•	•	•	•	•	•	•	•	•
9	•	•	•	•	□	•	•	•	•	•	•	•	•	•	•	•
10	□	•	•	□	•	□	□	□	□	•	•	•	•	•	□	□
11	•	•	•	•	□	•	□	•	•	□	•	□	□	•	□	•
12	□	•	□	•	•	•	•	□	•	•	□	□	□	□	□	•

Figure 1a: The choices of players in each period. The symbol • indicates a player chose H and □ indicates that a player chose D.

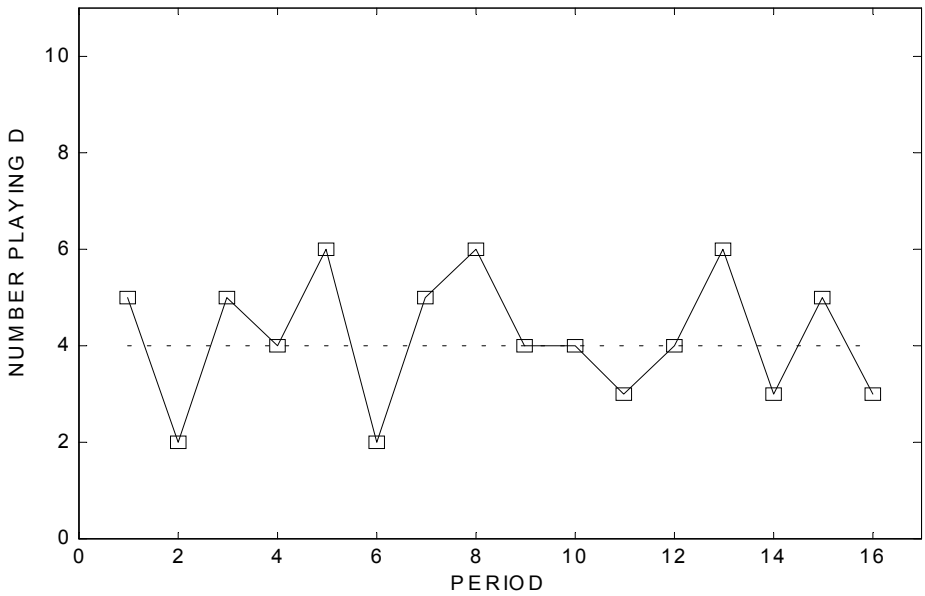

Figure 1b: The aggregate number of subjects playing D in each period.

Figure 1: An experimental run of the one population, no labels Hawk-Dove game. The data corresponds to run 9 in session 6 of the Friedman (1996) experiments. The experiment lasted 16 periods and involved 12 subjects.

PLAYER		PERIOD 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
GROUP 1	1	□	●	●	●	●	●	●	●	●	●	●	□	□	□	□	□
	3	●	●	□	□	□	□	□	●	●	●	□	□	□	□	□	□
	5	□	●	□	□	●	●	□	□	□	●	□	□	●	□	□	□
	7	□	□	●	●	●	□	□	□	□	□	●	●	●	●	□	□
	9	●	●	□	□	□	●	□	□	●	●	□	●	□	□	□	●
	11	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
GROUP 2	2	□	●	●	□	□	□	□	□	□	□	□	●	□	□	□	□
	4	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
	6	□	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
	8	●	□	●	●	□	●	●	●	●	●	●	□	●	●	●	●
	10	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
	12	□	□	●	□	□	□	□	●	□	●	●	□	□	●	□	●

Figure 2a: The choices of players in each period. The symbol • indicates a player chose H and □ indicates that a player chose D.

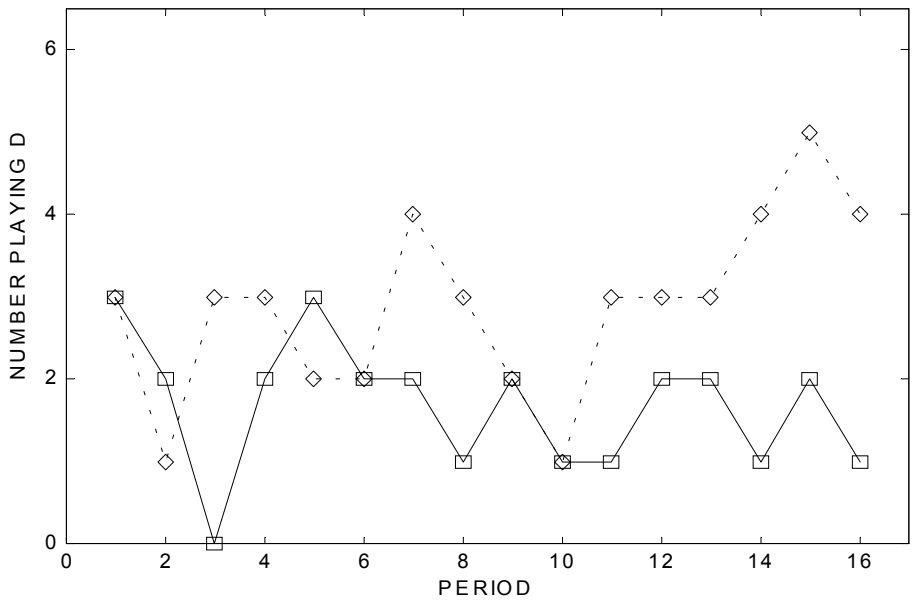

Figure 2b: The aggregate number of subjects in each group playing D in each period. The solid line indicates the number in Group 2.

Figure 2: An experimental run of the two population, no labels Hawk-Dove game. The data corresponds to run 5 in session 23 of the Friedman (1996) experiments. The experiment lasted 16 periods and involved 12 subjects.

We see in Figure 2 that there are consistently more players in Group 1 that are playing D than in Group 2. This provides some evidence that the discriminatory equilibrium characterizes actual play. It also appears that Group 2 quickly becomes a group that is predominantly Hawkish. Group 1, in contrast, appears to have become primarily Dovish but does so less quickly and with a greater number of 'deviations' to Hawkish behavior. This is, however, easily explained by the fact that the Doves 'lose out' relative to the Hawks.

The experimental runs discussed above are fairly representative of the general results obtained by Friedman (1996). Friedman (1996) draws the conclusion that aggregate play typically does appear to converge towards approximate Nash equilibrium play. This emergence of approximate Nash equilibrium play in experimental evolutionary games is confirmed by a number of authors. Friedman (1996) is also able to reject (with a high degree of confidence for most runs) the null hypothesis 'that all players are independently choosing H with probability $\frac{2}{3}$' against the alternative 'that players change their action less frequently'. In a two population setting Friedman (1996) concludes that play typically converges to the discriminatory equilibrium. The emergence of a discriminatory equilibrium is also observed by Hargreaves-Heap and Varoufakis (2002).[14] In experiments where players are labelled as blue or red, Hargreaves-Heap and Varoufakis (2002) strongly reject the null hypothesis that 'the frequency with which a red player plays H is equal to the frequency with which a blue player plays H'. The emergence of discriminatory equilibria is also observed by Van Huyck, Battalio and Mathur (1995) and Van Huyck, Battalio and Rankin (1997) in treating stage games other than Hawk-Dove.

8 Relationships to the literature

A large related literature addresses the possible motivations for players to use mixed strategies (a recent paper on this topic is Reny and Robson 2002). The central issue is whether a mixed strategy equilibrium can be seen as approximately equivalent to a pure strategy equilibrium. This is plausible because imperfect information and the resultant exogenous uncertainty make explicit randomization unnecessary. Aumann at el. (1983) provide sufficient conditions, on the exogenous uncertainty, such that any mixed strategy vec-

[14] Interestingly, it takes around 20 periods for discrimination to strongly emerge (Hargreaves–Heap and Varoufakis 2002) which is longer than the length of the experimental runs (of 16 periods) used by Friedman (1996).

tor can be approximately purified. Harsanyi (1973) argues that a game with perfect information should be considered as an idealization of nearby games in which there is a small amount of payoff uncertainty. Harsanyi (1973) shows that such uncertainty implies any mixed strategy equilibrium can be approximately purified.

This literature would suggest that for a wide class of games and for any mixed strategy equilibrium of such games there exists an approximately equivalent pure strategy equilibrium. This may appear to generalize the results in this paper as our results suggest there are games for which an approximate equilibrium in pure strategies does not exist. Further, this remains the case even when we allow players to introduce some exogenous uncertainty. The literature, however, primarily questions the motivations of rational players while the focus of this paper is one of questioning the bounded rationality of using pure strategies. The implications of this different approach are felt in both the assumptions of conformity and that the number of roles is no larger than the number of actions. For example, it is typical of the literature, as in Aumann et. al. (1983) to assume a continuum of types. It is difficult to envisage a boundedly rational player being able to condition actions on a continuum of types.

Another approach to the question of whether a mixed strategy equilibrium can be seen as approximately equivalent to a pure strategy equilibrium is taken in Kalai (2000) where he shows that in incomplete information games with many players, every Bayesian Nash equilibrium is approximately information proof. As a consequence, in complete information games with many players, every Nash equilibrium generates an approximate Nash equilibrium in pure strategies where the goodness of the approximation is given by the law of large numbers. Our Theorem 3 shows that for an arbitrary game with perfect information there exists a Bayesian Nash ε information proof equilibria, and provides a lower bound on ε, stated in terms of the parameters describing the game. Thus, Kalai's results and ours are quite distinct.

The model of this paper is also related to the literature on correlated equilibria (for an introduction see Myerson 1997). Correlated equilibrium is motivated as reflecting the incentives on players to coordinate their actions. For example, in a Battle of the sexes game the two players can coordinate their actions by utilizing exogenous signals. Such signals may reflect preplay communication or readily observable 'sunspots'. Theorem 3 could be interpreted as showing the existence of an approximate correlated equilibrium consistent with conformity. As above, however, we can point to a slight difference in emphasis between this paper and the literature. In particular,

the literature on correlated equilibria is motivated by considerations of how rational players can coordinate their actions. Roles, as introduced in this chapter, are motivated by considerations of how boundedly rational players may be able to approximate rational behavior through conformity.

A further related literature concerns the evolution of institutions (see, for example, Durlauf and Young 2001, Young 2001 and references therein). This literature addresses the question of how conventions or institutions can evolve, through individual interactions, to create coordination on a large scale. Such a literature helps in understanding how roles could become endogenised in the way we assume in this chapter.

9 Conclusion

This paper has introduced the concepts of approximate substitutes in non-cooperative games. Doing so allows us to put a bound on the distance from rationality of using pure strategies and of conformity. We use and motivate a model of conformity that permits players to conform and yet play different actions. This is possible through imperfect information and the existence of roles. In particular, players can make action choice conditional on their role and roles are assigned to players randomly. To retain a meaningful notion of conformity we impose two restrictions on how roles are allocated: that each player has the same probability of being assigned each role (so players cannot be differentiated on the basis of these probabilities) and that the number of players who will have each role is ex-ante known (so that after roles are assigned, no player would significantly benefit by changing his behavior). Our results are related to the experimental research, in particular to Friedman (1996) and Hargreaves-Heap and Varoufakis (2002). In a companion paper (Wooders and Cartwright 2003b) we apply the results of this paper to demonstrate that in large games the use of pure strategies and conformity can be consistent with individual rationality. In particular, we connect the concepts of approximate substitutes in non-cooperative games and non-cooperative pregame frameworks to provide a family of (δ, Q)-class games for arbitrarily small δ.

10 Appendix

Theorem 1 makes use of two lemmas. The first is due to Wooders, Cartwright and Selten (2001). First we introduce some notation. Given real number h let $\lfloor h \rfloor$ denote the nearest integer less than or equal to h and $\lceil h \rceil$ the nearest

integer greater than h (i.e. $\lfloor 9.5 \rfloor = 9$ and $\lceil 9.5 \rceil = 10$. Also note that $\lfloor 9 \rfloor = 9$ and $\lceil 9 \rceil = 10$).

Lemma 1: For any strategy vector $\sigma = (\sigma_1, ..., \sigma_n)$ there exists a degenerate strategy vector $m = (m_1, ..., m_n)$ such that $support(m_i) \subset support(\sigma_i)$ for all $i \in N$ and:

$$\left\lceil \sum_{i=1}^{n} \sigma_i(s^k) \right\rceil \geq \sum_{i=1}^{n} m_i(s^k) \geq \left\lfloor \sum_{i=1}^{n} \sigma_i(s^k) \right\rfloor$$

for all $s^k \in \mathcal{S}$.

Lemma 2: Given any two strategy vectors m and σ where m is degenerate and where

$$\left\lceil \sum_{i=1}^{n} \sigma_i(s^k) \right\rceil \geq \sum_{i=1}^{n} m_i(s^k) \geq \left\lfloor \sum_{i=1}^{n} \sigma_i(s^k) \right\rfloor$$

for all $s^k \in \mathcal{S}$ there exists a strategy vector $\overline{\sigma}$ such that,

$$\sum_{i=1}^{n} \overline{\sigma}_i(s^k) = \sum_{i=1}^{n} \sigma_i(s^k)$$

and

$$\sum_{i=1}^{n} \left| \overline{\sigma}_i(s^k) - m_i(s^k) \right| \leq 1$$

for all $s^k \in \mathcal{S}$.

Proof: Given such a σ and m we proceed by constructing an appropriate $\overline{\sigma}$. Let K^- denote the set of pure strategies for which

$$\sum_{i=1}^{n} m_i(s^k) = \left\lfloor \sum_{i=1}^{n} \sigma_i(s^k) \right\rfloor$$

and let K^+ denote the set of strategies for which

$$\sum_{i=1}^{n} m_i(s^k) = \left\lceil \sum_{i=1}^{n} \sigma_i(s^k) \right\rceil.$$

We note that for every $s^k \in \mathcal{S}$ either $s^k \in K^-$ or $s^k \in K^+$. For each s^k let,

$$A(k) = \sum_{i=1}^{n} \sigma_i(s^k) - \left\lfloor \sum_{i=1}^{n} \sigma_i(s^k) \right\rfloor$$

and let,

$$A^+ = \sum_{k:s^k \in K^+} A(k).$$

Provisionally, set $\overline{\sigma}_i = m_i$ for all i. Then, for each $s^{\overline{k}} \in K^+$ identify a player $i_{\overline{k}}$ such that $m_{i_{\overline{k}}} = 1$. For each $s^{\overline{k}} \in K^+$ re-allocate player $i_{\overline{k}}$ the strategy defined, for each k, as follows:[15]

$$\overline{\sigma}_{i_{\overline{k}}}(s^k) = \begin{cases} A(k) \text{ if } k = \overline{k} \\ 0 \text{ if } s^k \in K^+ \text{ and } k \neq \overline{k} \\ A(k)\frac{1-A(\overline{k})}{|K^+|-A^+} \text{ otherwise.} \end{cases}$$

We conjecture that this strategy vector $\overline{\sigma}$ satisfies the required conditions. First, we should check that $\overline{\sigma}$ as defined above is indeed a strategy vector. For those players i for whom $\overline{\sigma}_i = m_i$ there is nothing to check. Consider a player $i_{\overline{k}}$ for some $s^{\overline{k}} \in K^+$. We begin by noting that $1 \geq \overline{\sigma}_{i_{\overline{k}}}(s^k) \geq 0$ for all k which is clear after noting that $|K^+| \geq 1 + A^+ - A(\overline{k})$. We also note that

$$\begin{aligned} \sum_k A(k) &= \sum_k \left(\sum_{i=1}^n \sigma_i(s^k) - \left\lfloor \sum_{i=1}^n \sigma_i(s^k) \right\rfloor \right) \\ &= n - \sum_{s^k \in K^-} \sum_{i=1}^n m_{ik} - \sum_{s^k \in K^+} \left[\sum_{i=1}^n m_{ik} - 1 \right] \\ &= \left|K^+\right|. \end{aligned}$$

Thus,

$$\sum_{s^k \in K^-} A(k) = \left|K^+\right| - A^+.$$

Therefore,

$$\begin{aligned} \sum_k \overline{\sigma}_{i_{\overline{k}}}(s^k) &= A(\overline{k}) + \frac{1 - A(\overline{k})}{|K^+| - A^+} \sum_{k:s^k \in K^-} A(k) \\ &= 1. \end{aligned}$$

Thus, $\overline{\sigma}$ is a strategy vector.

[15] Note that if K^+ is non-empty then $\left|K^+\right| > A^+$.

We note that for all $s^k \in K^-$,

$$\begin{aligned}\sum_{i=1}^{n} \overline{\sigma}_i(s^k) &= \left\lfloor \sum_{i=1}^{n} \sigma_i(s^k) \right\rfloor + \frac{A(k)}{|K^+| - A^+} \sum_{s^{\overline{k}} \in K^+} \left(1 - A(\overline{k})\right) \\ &= \sum_{i=1}^{n} \sigma_i(s^k).\end{aligned}$$

Clearly, for all $s^k \in K^+$ we have that $\sum_{i=1}^{n} \overline{\sigma}_i(s^k) = \sum_{i=1}^{n} \sigma_i(s^k)$. Finally, for all $s^k \in K^-$

$$\sum_{i=1}^{n} \left|\overline{\sigma}_i(s^k) - m_i(s^k)\right| = \frac{A(k)}{|K^+| - A^+} \sum_{s_{\overline{k}} \in K^+} \left(1 - A(\overline{k})\right) = A(k) \leq 1,$$

and for all $s^k \in K^+$

$$\sum_{i=1}^{n} \left|\overline{\sigma}_i(s^k) - m_i(s^k)\right| = (1 - A(k)) \leq 1.$$

This completes the proof.■

Theorem 1: Let $\Gamma = (N, A, T, p, u)$ be a $(\delta_I, \delta_P, \delta_C, Q)$-class Bayesian game. Let ε be a positive real number where $\varepsilon \geq 2(\delta_I + \delta_C)$. Then the game Γ has a Bayesian Nash ε-equilibrium in pure strategies.

Proof: Using Nash's Theorem there must exist a Bayesian Nash Equilibrium strategy vector σ^*. This implies, for all $i \in N$, that:

$$U_i(y, \sigma^*_{-i}|t^z) \geq U_i(x, \sigma^*_{-i}|t^z) \tag{5}$$

for all $x, y \in \Delta(\mathcal{S})$ where $support(y) \subset support(\sigma^*_i)$ and for all $t^z \in \mathcal{T}$.

Given that Γ is a $(\delta_I, \delta_P, \delta_C, Q)$-class game there is a partition of N into Q classes. Let $\{N_1, ..., N_Q\}$ be such a partition. We apply Lemma 1 in turn to each N_q. Doing so implies that there exists a strategy vector $m \in \Sigma$ where $support(m_i) \subset support(\sigma^*_i)$ and m_i is degenerate for all $i \in N$, and where:

$$\left| \sum_{i \in N_q} \sigma^*_i(s^k) - \sum_{i \in N_q} m_i(s^k) \right| \leq 1$$

for all N_q and all $s^k \in \mathcal{S}$. By Lemma 2 this implies that there exists a strategy vector $\overline{\sigma} \in \Sigma$ such that

$$\sum_{i \in N_q} \overline{\sigma}_i(s^k) = \sum_{i \in N_q} \sigma_i^*(s^k)$$

and

$$\sum_{i \in N_q} \left| \overline{\sigma}_i(s^k) - m_i(s^k) \right| \leq 1$$

for all s^k and all N_q.

From the definition of a δ-interaction substitute partition:

$$\left| U_i(x, \overline{\sigma}_{-i} | t^z) - U_i(x, \sigma_{-i}^* | t^z) \right| \leq \delta_I$$

for any $x \in \Delta(\mathcal{S})$, $t^z \in \mathcal{T}$ and $i \in N$. Using the definition of a δ-strategy switching partition:

$$|U_i(x, \overline{\sigma}_{-i} | t^z) - U_i(x, m_{-i} | t^z)| \leq \delta_C$$

for any $x \in \Delta(\mathcal{S})$, $t^z \in \mathcal{T}$ and $i \in N$. Thus:

$$\left| U_i(x, \sigma_{-i}^* | t^z) - U_i(x, m_{-i} | t^z) \right| \leq \delta_I + \delta_C$$

for any $x \in \Delta(\mathcal{S})$, $t^z \in \mathcal{T}$ and $i \in N$. Therefore, given (5),

$$U_i(m_i, m_{-i} | t^z) - U_i(x, m_{-i} | t^z) \geq -2(\delta_I + \delta_C) \geq -\varepsilon$$

for any $x \in \Delta(\mathcal{S})$, $t^z \in \mathcal{T}$ and $i \in N$.■

References

[1] Aumann, R.J., Y. Katznelson, R. Radner, R.W. Rosenthal,and B. Weiss (1983) "Approximate purification of mixed strategies," *Mathematics of Operations Research* 8: 327-341.

[2] Blume, L.E. (1993) "The statistical mechanics of strategic interaction," *Games and Economic Behavior* 5: 387-424.

[3] Cartwright, E.J. (2003) "Learning to play approximate nash equilibria in games with many players, " University of Warwick Department of Economics Discussion Paper 671. On-line at http://www2.warwick.ac.uk/fac/soc/economics/research/papers/twerp671.pdf.

[4] Cartwright. E.J. and M.H. Wooders (2003a) "Social conformity and bounded rationality in arbitary games with many players," University of Warwick Working Paper no 672. On-line at http://www2.warwick.ac.uk/fac/soc/economics/research/papers/twerp672.pdf.

[5] Cartwright. E.J. and M.H. Wooders (2003b) "Conformity and bounded rationality in games with many players," University of Warwick Working Paper (to appear).

[6] Cheung, Y.W. and D. Friedman (1998) "A comparison of learning and replicator dynamics using experimental data," *Journal of Economic Behavior and Organization:* 35: 263-280.

[7] Conley, J. and M.H. Wooders (2001) "Tiebout economics with differential genetic types and endogenously chosen crowding characteristics," *Journal of Economic Theory* 98: 261-294.

[8] Durlauf, S. and H.P. Young (2001) *Social Dynamics*, MIT Press.

[9] Ellison, G. and D. Fudenberg (1995) "Word-of-mouth communication and social learning," *Quarterly Journal of Economics* 110: 93-125.

[10] Follmer, H. (1974) "Random economies with many interacting agents," *Journal of Mathematical Economics* 1: 51-62.

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

[12] Gale, D. and R. Rosenthal (1999) "Experimentation, imitation, and strategic stability," *Journal of Economic Theory* 84: 1-40.

[13] 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.

[14] 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.

[15] Hildenbrand, W. (1971) "Random preferences and equilibriumanalysis," *Journal of Economic Theory* 3: 414-429.

[16] Kalai, E. (2000) "Private information in large games," Northwestern University Discussion Paper 1398.

[17] Kagel, J.H. and A.E. Roth, 1995, *The Handbook of Experimental Economics*, Princeton University Press.

[18] Khan, A., K.P. Rath and Y.N. Sun (1997) "On the existence of pure strategy equilibria with a continuum of players," *Journal of Economic Theory* 76:13-46.

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

[20] Mas-Colell, A. (1984) "On a theorem of Schmeidler," *Journal of Mathematical Economics* 13: 206-210.

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

[22] Pascoa, M. (1993) "Approximate equilibrium in pure strategies for nonatomic games," *Journal of Mathematical Economics* 22: 223-241.

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

[24] Reny, P.J. (1999) "On the existence of pure and mixed strategy Nash equilibria in discontinuous games," *Econometrica* 67: 1029-1056.

[25] Reny, P.J. and A. Robson (2002) "Reinterpreting mixed strategy equilibria: a unification of the classical and Bayesian views," on-line at http://publish.uwo.ca/~arobson/homepage.html.

[26] Scharfstein, D.S. and J.C. Stein (1990) "Herd behaviour and investment," *The American Economic Review*, vol 80, no. 3: 465-479.

[27] Schleifer, A., (2000) *Inefficient Markets, an introduction to behavioral finance*, Oxford University Press.

[28] Schmeidler, D. (1973) "Equilibrium points of nonatomic games," *Journal of Statistical Physics* 7: 295-300.

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

[30] Van Huyck, J., R.C. Battalio and R.O. Beil, 1990, "Tacit coordination games, strategic uncertainty, and coordination failure," *American Economic Review*, 80: 234-248.

[31] Van Huyck, J., R.C. Battalio, S. Marthur, A. Ortmann and P.P. Van Huyck (1995) "On the origin of convention: Evidence from symmetric bargaining games," *International Journal of Game Theory*, 24: 187-212.

[32] 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.

[33] Walker, M. and J. Wooders (2001) "Minimax play at Wimbledon," *American Economic Review* 91: 1521-1538.

[34] 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.

[35] Young, H.P. (2001) *Individual Strategy and Social Structure*, Princeton University Press.