CONFORMITY, EQUITY AND CORRELATED EQUILIBRIUM

by

Edward Cartwright and Myrna Wooders

Working Paper No. 08-W06

February 2008

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Conformity, equity and correlated equilibrium*

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, Vanderbilt University
Nashville, TN 37235 USA
and
Department of Economics, University of Warwick
Coventry CV4 7AL, UK
M.Wooders@warwick.ac.uk

July 2007

Keywords: conformity, equity, stereotyping, correlated equilibrium, stereotyped beliefs, subjective correlated equilibrium

Abstract

We explore the potential for correlated equilibrium to capture conformity to norms and the coordination of behavior within social groups. Given a partition of players into social groups we propose three properties one may expect of a correlated equilibrium: within-group anonymity, group independence and stereotyped beliefs. Within-group anonymity requires that players within the same social group have equal opportunities and equal payoffs. Group independence requires that there be no correlation of behavior between groups. If beliefs are stereotyped then any two members of a social group are expected to behave identically.

*This paper is a 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, at Northwestern University in August 2002, at the Purdue Economic Theory conference in honor of C.D. Aliprantis and at Hebrew University in 2006. We thank the participants for their interest and comments.

We demonstrate that there are subjective correlated equilibrium satisfying within-group anonymity, group independence and stereotyping. These results apply when players within social groups are 'similar', and not necessarily identical. A number of related issues, such as fairness, are also discussed.

1 Introduction

It is well documented that individuals can coordinate their behavior to mutual advantage even if doing so requires different individuals to perform different actions and receive inequitable rewards (e.g. Schelling 1960, Hayek 1982, Sugden 1989, Friedman 1996, Van Huyck et al. 1997, Rapoport et. al. 2001, Hargreaves-Heap and Varoufakis 2002). Interaction in social groups (whether groups be teenager, heavy metal fan, football supporter, economist, Harvard professor, Christian, etc.) occupies much of our everyday life and facilitates such coordination (Johnson and Johnson 1987). In particular, individuals, often subconsciously, conform to established rules and norms of behavior of the groups to which they belong, resulting in coordination of behavior (Hayek 1960, 1982, Akerlof and Kranton 2000).[1] Memberships in social groups also determine a person's social identity, framing how he is expected to behave by others (Sherif 1966, Tajfel 1978, Brown 2000). Indeed, different norms may or may not emerge in different social contexts (Roth et. al.1991, Brandts, Saijo and Schram 2004).

In situations that require individuals to perform different but coordinated actions, coordination can be achieved through the conditioning of actions on signals or roles (Johnson and Johnson 1977, Selten 1980, Sugden 1989, Hogg et al. 1995). For example, as cars meet on a narrow road, roles may be "first to arrive," "second to arrive," and so forth, and actions can be conditioned on these roles, with the first to arrive taking the action "do not give way" and the second to arrive taking the action "give way" etc. A more complex example is the organization of an academic department. There are a variety of roles to be filled within the department – some explicitly recognized, such as chair, director of graduate studies, and chair of the recruitment committee, while others are implicitly desired, such as people willing to attend seminars and contribute to the discussion. A successful department is one in which these roles are advantageously allocated amongst

[1] The economic literature on conformity includes Akerlof (1980), Elster (1989) and Bernheim (1994).

faculty members.[2]

The starting point for the present paper is that the coordination of individual actions can be nicely captured by the concept of correlated equilibrium (Aumann 1974, 1987). In contrast to Nash equilibrium, correlated equilibrium allows player actions to be statistically dependent on some random event external to the model. This could be, for example, whether it rains or whether the player is male or female. More generally, imagine a mediator (or some device) that instructs players to take actions according to some commonly known probability distribution. In this paper we think of the mediator as distributing roles. If it is in the interests of each player to assume the role assigned to him by the mediator's instructions, then the probability distribution over roles is a correlated equilibrium (cf., Aumann 1987, Forges 1986, Dhillon and Mertens 1996). Note that a correlated equilibrium will induce a probability distribution over actions. If roles are distributed independently across players, then a correlated equilibrium is equivalent to a Nash equilibrium of the original game. Given that roles may be correlated across individuals, the set of correlated equilibria is generally larger than the set of Nash equilibria.[3]

Correlated equilibrium has the potential to capture the coordination of behavior through the observance of rules and conformity to norms. In particular, each player can be thought of as having a pure strategy that maps instructions into actions of the form "if told to play action x then play action x"; in a correlated equilibrium, all players use the same pure strategy of conforming to the instructions they are given.

A motivation for the current paper is that additional insights are possible when one considers how the concept of correlated equilibrium fits into a system of social identity. In particular, social groups may serve as a medium through which roles are distributed. This suggests that certain properties are to be expected of an equilibrium. We will consider three properties: within-group anonymity and group independence, both properties of the correlated equilibrium, and stereotyping, a property of beliefs. These prop-

[2] Hayek (1960, 1982) distinguishes between organization that results from "purposeful construction" and organization that result from "self-generating" spontaneous order. We shall treat these as equivalent. Importantly, however, we do model order as arising through free choice: i.e., whether an individual is told to do something (become chair of graduate studies) or feels he should do something (attend seminars), he is free to make his own decisions.

[3] The set of correlated equilibria has many appealing properties, for example, it is nonempty, compact, convex and easy to describe (Aumann 1974, 1987). See also Hart (2005) for a discussion of recent work (in collaboration with Mas-Colell) on how adaptive learning leads to correlated equilibrium play.

erties imply that the resulting probability distribution over roles appears 'simple' (and therefore one that could potentially emerge) in the sense that roles are distributed 'anonymously' within social groups. Also, the probability distribution is one that most players would not find objectionable.[4] Note that the three properties act as a form of equilibrium selection device on the set of correlated equilibria.

The property of within-group anonymity concerns fairness or equity within groups and is motivated by evidence that 'fairness' is important within social groups. Any individual who perceives group opportunities and outcomes as biased against himself is likely to become disillusioned and leave the group. Similarly anyone perceived as doing better than everyone else may be forced out of the group or aspire to a group that is a better fit (Brown 2000). Either way, the success of social groups depends, in part, on perceptions of within-group equity (Johnson and Johnson 1987, Tajfel 1978, Rabin 1993 and Fehr and Schmidt 1999). For example, a second driver to arrive at an intersection may be happy to conform to the "second driver gives way norm" because he expects that over time he will be first to arrive as often as he is second to arrive. Similarly, faculty members may be willing to do more than their fair share in the expectation that others will take over in due course or that eventually their efforts will be rewarded. The within-group anonymity property requires that any two individuals within the same group have the same probability of being allocated roles. This results in *equity of opportunity*, whereby the distribution of roles within a group can be seen as fair, and equity of *expected payoff*, whereby outcomes can be seen as fair. Within-group anonymity also eliminates the necessity for the correlating device to be able to identify players within groups and thus greatly simplifies the allocation of roles.

The property of group independence concerns correlation between groups. That social group membership binds players together, creating a "them-and-us" attitude, is well known (Tajfel 1978, Hogg and Vaughan 2005). In this context the correlation of actions between players in different groups is unlikely or at least questionable.[5] The group independence property requires that the distribution of roles is statistically independent between different groups. This rules out the correlation of roles across groups. Group independence also, like within-group anonymity, simplifies the allocation of

[4] It could be argued that, because a correlated equilibrium is individually rational, players would behave according to a correlated equilibrium whether they find it objectionable or not. But, as we discuss in the conclusion, the device that decides how roles are allocated is ultimately endogenous and so should be one to which most players 'have no objections'.

[5] Ruling out correlation between groups does not rule out coordination between groups.

roles.

An important point to bear in mind, particularly with regard to within-group equity, is that *perceptions* and not necessarily reality matter (Hogg and Vaughan 2005). If an outcome is perceived as equitable, it may not matter whether it is in fact equitable. For example, one faculty member may take on more responsibility than any other but all may perceive this as fair. The distinction between perception and reality led us to look at *subjective* correlated equilibrium. A subjective correlated equilibrium differs from a correlated equilibrium in that in a subjective equilibrium players may have differing beliefs about the probability with which roles are distributed. We say beliefs are stereotyped if a player believes that players in the same social group have the same probability of being allocated each role. Thus, players expect within-group anonymity, whether or not it actually happens. That stereotyping exists is widely acknowledged (Hogg and Vaughan 2005).[6]

Underlying our approach is the assumption that groups are composed of similar individuals. An expectation of homogeneity within groups follows from the social identity and social categorization derived from group membership. If an individual orients himself in society through group membership, judging himself and being judged by others according to the group(s) he belongs to, then an individual aims to belong to the group that he best fits. This leads to like-minded people grouping together (Akerlof and Kranton 2000, Brown 2000, Hogg and Vaughan 2005). In practice players in the same group can be expected to be similar but not necessarily identical. To capture this phenomenon we introduce a concept of approximate player substitutability. This concept provides a metric that can be used to measure differences between players both in how they are affected by the actions of others and in how the actions of others affect them.

Given that this paper is motivated by a belief that correlated equilibrium can be used to capture conformity to norms and the coordination of behavior with groups we have two objectives. First, to show that correlated equilibrium can capture conformity and norms within groups, and second, to show that doing so provides useful insights. With the first objective in mind we shall demonstrate the existence of a correlated equilibrium satisfying within group anonymity and group independence. Specifically, if players are approximate substitutes, then we obtain an approximate correlated equilibrium satisfying within group anonymity and group independence. Given that any two players are approximate substitutes for some level of approxi-

[6] We do not regard either individual rationality or stereotyping as necessarily desirable characterisics of members of a society.

mation our results apply to any (finite) game and any set of social groups. With the second objective in mind we shall look at four issues, stereotyping, fairness, optimal group size and semi-anonymity. We demonstrate, for example, that stereotyping can be consistent with individual rationality.

We proceed as follows: Section 2 introduces the model, Section 3 discusses properties of social groups and Section 4 provides the main results. In Section 5 we address the further issues of stereotyping, fairness, optimal group size and semi-anonymity. Section 6 concludes the paper and indicates directions for further research. An appendix contains additional 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 $\mathbb{R}$.

Given game $\Gamma = (N, A, \{u_i\}_{i\in N})$ we think of a *correlating device* p that sends a private signal to each player. The signal suggests to each player an action. Once a player receives his signal, but does not observe those of others, he chooses an action.[7] To distinguish suggested action from actual behavior we equate the signal with an *assignment to a role*. Thus, the correlating device p assigns each player a role from set A and the player then chooses an action that may or may not equate with his assigned role.

Given that the correlating device assigns a role to each player we can formally think of the correlating device as being given by a *probability distribution p over action profiles* where $p(\overline{a})$ denotes the probability that players will be assigned roles consistent with action profile $\overline{a}$. We shall denote by p_i the marginal distribution of p, where $p_i(k)$ denotes the probability that player i is assigned role k.[8] Let $P := \Delta(R^N)$ denote the set of possible correlating devices.

Given game Γ and correlating device $p \in P$ we refer to the *game with roles* Γ^p. Given that action choice can be made conditional on assigned role a *strategy* for a player i, in game Γ^p, is a function s_i mapping the set of signals A to the set of actions A. In interpretation, $s_i(k)$ is the action performed by player i if he is assigned role k. Of primary interest is the *strategy profile* $\overline{s}^* \in S^N$ where $s_i^*(k) = k$ for all k and i. That is, the strategy profile where

[7] In this sense signal can be equated to Harsanyi type.

[8] Formally, $p_i(k) = \sum_{\overline{a}:\overline{a}_i=k} p(\overline{a})$

each player plays the action consistent with their assigned role. Let S denote the set of strategies.

Note that a player's payoff does not depend directly on his role or the roles of other players although it may depend on roles indirectly through the choice of action that a distribution of roles induces. We shall assume for the present that correlating device p is common knowledge and players have consistent beliefs with respect to p. We relax this assumption in Section 4. Given this we can define a 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 assigned according to distribution p and players behave according to $\overline{s}^*$.

For any $\varepsilon \geq 0$ we say that p is a *correlated ε-equilibrium of game* Γ if and only if

$$U_i(p) \geq \sum_{\overline{a} \in A^N} p(\overline{a}) u_i(s_i(\overline{a}_i), \overline{a}_{-i}) - \varepsilon \tag{1}$$

for all $s_i \in S$.[9] Thus, no player would wish to deviate from playing the action corresponding to his allocated role. We refer to a correlated 0-equilibrium as a correlated equilibrium. 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 Nash ε-equilibrium is defined in the standard way).

2.1 Two simple examples

To illustrate the ideas as we proceed through the remainder of the paper consider the following simple example.

Example 1: There are three players $N = \{1, 2, 3\}$ and two possible actions $A = \{E, Z\}$. We shall think of members of an economics department having to teach intermediate microeconomics with actions "put in effort" (E) or "be lazy" (Z). Players 1 and 2 constitute one social group and player 3 constitutes a second social group. For example, players 1 and 2 might be teaching assistants (TAs) and player 3 might be a professor.

A player's payoff is equal to a bonus minus an effort cost. The bonus, the same for all players, depends on effort and is non-zero if and only if player 3

[9] This definition equates to a natural approximation to the standard definition of correlated equilibrium, although 'role' is often termed 'signal' (Fudenberg and Tirole 1998). Note, however, that the use of the term ε-correlated equilibrium by Myerson (1986) has a different meaning to the one here.

and at least one of players 1 and 2 choose E. If players 1 and 3 choose E the bonus is 10; if instead players 2 and 3 choose E the bonus is $b \leq 10$ for some real number b; if players 1, 2 and 3 all choose E the bonus is 13. Basically, the success of intermediate micro depends on the effort put in and requires a professor plus at least one TA to put in effort. If the course is a success then all receive a bonus. Effort does, however, come at a cost. If players 1 or 2 choose action E then they pay effort cost e_1 and e_2 respectively (and if they choose Z they pay effort cost 0). We shall fix values for b, e_1 and e_2 as we proceed. Payoffs can be summarized:

		Payoff of player		
Set of players who play E	Bonus	1	2	3
$\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}$	0	0 or $-e_1$	0 or $-e_2$	0
$\{1,3\}$,	10	$10 - e_1$	10	10
$\{2,3\}$	b	b	$b - e_2$	b
$\{1,2,3\}$	13	$13 - e_1$	$13 - e_2$	13

A correlating device suggests to each player what action to take, either E or Z, and is a correlated equilibrium if each player does best to play the suggested action. That is, each player performs his assigned role. To illustrate, the concept of correlated equilibrium suppose that it is efficient for one, and only one, of the TAs to put in effort. Further, suppose that the correlating device consists of a signal sent to each of the TAs indicating whether he is the more intelligent of the two. The TA who thinks that he is the most intelligent puts in effort. We can then contrast three different types of correlating device all of which may be correlated equilibria. (1) With some positive probability, each TA independently thinks that he is the most intelligent. This means that the two TAs will independently decide whether to put in effort and results in no correlation of actions. (2) The TA with the best exam performance is considered by both as the most intelligent and therefore puts in effort while the other does not. Note that which TA will have the best exam performance may be a random event but both TAs know for certain, ex-post, who has the best exam performance. Thus, one and only one randomly selected TA will put in effort. (3) The TAs receive a correlated signal of relative intelligence. For example, the TAs may believe the outcome of an exam is in part random, in a way that means that at least one of the TAs and potentially both of them believe that they are the most intelligent. This guarantees that one TA will put in effort and means that with positive probability both will put in effort. We shall see this type of correlating device in Sections 5.1 and 5.3.

Example 1 will prove sufficient for most that we wish to illustrate but a slight extension will prove useful on occasions.

Example 2: The action E is decomposed into E_1 and E_2 where we could interpret E_1 as "to put in effort using textbook 1" and E_2 as "to put in effort using textbook 2". We assume that the payoffs of players 1 and 2 are not effected by the choice of E_1 or E_2 and so still think of the two TAs as choosing E or Z. Player 3 may, however, get a higher payoff if he plays E_1 when player 1 plays E and plays E_2 when player 2 plays E. That is, the professor may do better to match his choice of textbook with the TA who puts in effort. To return to the earlier interpretation of a correlating device, the professor can also receive a signal indicating which of the two TAs is most intelligent. This allows him to form expectations over who will put in effort and may allow him to coordinate his choice of textbook with the TA or TAs likely to put in effort.

3 Social groups

As discussed in the introduction, we wish to study correlated equilibrium relative to a set of social groups. Formally, a partition $\Pi = \{N_1, ..., N_G\}$ of the player set into G subsets is taken as given. We refer to Π as a *partition into social groups* and refer to each N_g as a *social group.* In the examples it is natural to think of the two TAs as constituting one social group $N_1 = \{1, 2\}$ and the professor as a second social group $N_2 = \{3\}$.

An important part of our analysis is the effects of interchanging actions of members of the same social group. This leads us to the formalism of a permutation of an action profile. One action profile $\overline{a}'$ is a *permutation* of another action profile $\overline{a}$ if the number of players in each social group playing each action (or assigned each role) is the same.[10] Thus, in Example 1 the action profile (E, Z, E) is a permutation of action profile (Z, E, E) because, in both cases, one TA puts in effort while the other is lazy. Given action profile $\overline{a}$ let $\mathcal{P}^{\Pi}(\overline{a})$ denote the set of action profiles that are permutations of $\overline{a}$.

Taking a game Γ and a partition into social groups Π as given, two properties of a correlating device, motivated in the introduction, are within-group anonymity and group independence.

[10] More precisely, given a game Γ, a partition into groups $\Pi = \{N_1, ..., N_G\}$ and an action profile $\overline{a}$, let $h^{\Pi}(\overline{a}, k, g) = |\{i \in N_g : \overline{a}_i = k\}|$ denote the number of players in group N_g who play action k. We say that action profile $\overline{a}'$ is a *permutation of* $\overline{a}$ if $h^{\Pi}(\overline{a}, k, g) = h^{\Pi}(\overline{a}', k, g)$ for all k and N_g.

Within-group anonymity: Correlating device p satisfies within-group anonymity *(WGA)* if the distribution p treats players from the same social group identically. Formally, given any two action profiles $\overline{a}$ and $\overline{a}'$, if $\overline{a}' \in \mathcal{P}^{\mathrm{II}}(\overline{a})$ then:

$$p(\overline{a}) = p(\overline{a}').$$

WGA captures two important aspects of group behavior that we wish to model: equity and conformity. A probability distribution satisfying WGA provides equality of opportunity within groups because any two players belonging to the same social group have the same probability of being allocated each role within the group.[11] In our example, if exam scores of the TAs are the correlating device, and the TA with the highest exam score puts in effort while the other TA is lazy, then there is within group anonymity if each has an equal chance of the highest exam score. As we shall see in Section 5.2, this equality of opportunity results in an equality of expected payoff. WGA also implies conformity within social groups is observed in a correlated equilibrium satisfying WGA because any two players belonging to the same social group are, ex-ante, expected to behave identically: they have the same probability of being allocated each role and behave in identical ways once allocated a role. Further, it is individually rational for players to conform to the behaviors expected of their assigned roles.

We now turn to the group independence property.

Group independence: Let i and j be any two players belonging to different social groups. If player j is assigned role a_j, let $p_i(k|a_j)$ denote the probability that player i has role k. Correlating device p satisfies *group independence (GI)* if there is no correlation of actions between groups. Formally, it requires that $p_i(k) = p_i(k|a_j)$ for all $k, a_j \in R$.

If social groups are distinct then correlation of actions between groups may be unlikely. This reflects how correlating actions between groups may be difficult because players in different social groups do not as easily identify or communicate with each other. The professor, for instance, may not be 'familiar enough' with the two TAs to correlate his actions with their actions. Specifically, in Example 2, GI would not permit the professor to correlate his actions with the TAs by for example using textbook 1 if player 1 puts in effort and textbook 2 if player 2 puts in effort. Note, however, that GI does not stop coordination between groups. For instance, the professor

[11] For instance, if $i, j \in N_g$, then $p_i(k) = p_j(k)$ for all $k \in A$.

could coordinate his actions with the TAs by choosing textbook 2 in the knowledge that player 2 will put in effort.

4 Equilibrium

In this section we demonstrate the existence of a correlated equilibrium satisfying WGA and GI. We begin by describing a special case in which players in social groups are exact substitutes. We then generalize to the case of approximate substitutes before generalizing further to permit subjective beliefs. Our main result, Theorem 1, is obtained for the most general case (and stated in Section 4.5).

4.1 Exact substitute players

Take as given a game $\Gamma = (N, A, \{u_i\}_{i \in N})$. We say that two players i and j are *exact substitutes* if they have the same payoff function and influence other players identically. Thus, whenever i and j exchange strategies they also exchange payoffs while any third player l is indifferent. Formally, if players i and j are *exact substitutes* then $u_i(\overline{a}) = u_j(\overline{a}')$ and $u_l(\overline{a}) = u_l(\overline{a}')$ for all $l \neq i, j$ and any two action profiles $\overline{a}, \overline{a}' \in A^N$, where (1) $\overline{a}_l = \overline{a}'_l$ for all $l \neq i, j$, (2) $\overline{a}_i = \overline{a}'_j$ and (3) $\overline{a}_j = \overline{a}'_i$. A partition of the player set into social groups $\Pi = \{N_1, ..., N_G\}$ is an *exact substitute partition* if, for any social group N_g, any two players $i, j \in N_g$ are exact substitutes.

Returning to the example, it may be that players 1 and 2, the TAs, are equally skillful and thus player 3, the professor's, payoff is unaffected when the two TAs exchange actions. This would require that $b = 10$. For players 1 and 2 to be exact substitutes, we also require that $e_1 = e_2$, implying that both TAs have the same effort cost. In this case when 1 and 2 exchange strategies they also exchange payoffs. In particular, if 1 and 3 teach then the payoffs for players 1 and 2 are $10 - e_1$ and 10 respectively while if, instead, 2 and 3 teach the payoffs are 10 and $10 - e_2$. Note that the professor is not an exact substitute for a TA because if the professor and a TA exchange the strategies of "put in effort" for "be lazy" then all payoffs will change.

The following two results are derived from our main result, Theorem 1, and proved in the Appendix.[12] The first result states that if the player set is partitioned into social groups, each consisting of exact substitutes, then for any correlated equilibrium p^* there is a correlated equilibrium p' that

[12] These corollaries could also be obtained as a consequence of the well-known fact that the set of correlated equilibria is convex.

satisfies WGA and where all players in the same social group receive the same payoff, equal to the average within-group payoff from equilibrium p^*.

Corollary 1: Given a game Γ, an exact substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$ and a correlated equilibrium p^*, there exists a correlated equilibrium p' satisfying WGA, where

$$\frac{1}{|N_g|} \sum_{j \in N_g} U_j(p^*) = U_i(p') \tag{2}$$

for all $i \in N_g$ and all $N_g \in \Pi$.[13]

In terms of the example, (if $e_1, e_2 \leq 10$) the p^* could be $p^*(E, Z, E) = 1$. Thus, the first TA, player 1, and the professor, player 3, will have assigned role "put in effort". This does not satisfy WGA because the two TAs, players 1 and 2, have an unequal chance of being assigned role E and therefore of putting in effort. [Note that this results in player 2 getting a higher payoff than player 1.] There does exist a correlated equilibrium p' where $p'(E, Z, E) = p'(Z, E, E) = 0.5$. That is, one of the TAs is randomly selected to put in effort. This does satisfy WGA and both players 1 and 2 have the same expected payoff.

Our second result states that if the player set is partitioned into social groups consisting of exact substitutes then there exists a correlated equilibrium satisfying WGA and GI.

Corollary 2: Given a game Γ and an exact substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$, there exists a correlated equilibrium p^* of Γ satisfying WGA and GI, where $U_j(p^*) = U_i(p^*)$ for all $i, j \in N_g$ and all N_g.

The equilibrium p' described above, for Example 1, satisfies both WGA and GI.

4.2 Approximate substitute players

While players within social groups may be similar, in general one would not expect them to be identical. For example, two TAs are likely to be different, even if only slightly. This leads us to introduce a notion of approximate substitutes that will allow us to apply our results to any game and to any

[13] The set of correlated equilibria satisfying WGA is convex. Also note that p' preserves the properties of p^*, so, for example, if p^* is an acceptable correlated equilibrium (as defined by Myerson 1986), then p' is an acceptable correlated equilibrium.

partition of the player set into social groups, irrespective of within-group player heterogeneity.

Players i and j are δ-*substitutes* (or, informally, approximate substitutes) if they are both δ-interaction substitutes and δ-individual substitutes. 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. If the actions of i and j are interchanged, then the payoffs to other players are only slightly affected. In our example, if b is near to 10, then, since the TAs have similar abilities, the professor is relatively indifferent to which TA puts in effort. Players i and j are individual substitutes if they have similar payoff functions. In our example, this would require that e_1 is close to e_2 and so the two TAs have similar effort costs. Combining both measures, we refer to players i and j as substitutes if they are both interaction substitutes and individual substitutes.

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

$$\left|u_l(\overline{a}) - u_l(\overline{a}')\right| \leq \frac{\delta}{n} \tag{3}$$

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

$$\left|u_i(\overline{a}) - u_j(\overline{a}')\right| \leq \delta. \tag{4}$$

Players i and j are δ-*substitutes* (or informally approximate substitutes) if they are both δ-interaction substitutes and δ-individual substitutes. Note that, for some δ, any two players are δ substitutes; thus, δ substitutability provides a metric on the player set N.

Returning to the example the two TAs are $3(10 - b)$-interaction substitutes. The closer therefore is b to 10 the closer the approximation. The two TAs are also $10 - b + e_2 - e_1$ individual substitutes. So, the closer is b to 10 and e_1 to e_2 the closer the approximation.

A δ-*substitute partition into social groups* $\Pi = \{N_1, ..., N_G\}$ is a partition of player set N with the property that any two players belonging to the same social group are δ-substitutes; i.e., if $i, j \in N_g$ for some $N_g \in \Pi$, then i and j are δ-substitutes. Note that the partition into singletons $\{\{1\}, ..., \{n\}\}$ is a 0-substitute partition. Also, for any game Γ and any partition Π of N, for some finite $\delta \geq 0$ the partition Π is a δ-substitute partition into social groups. See Section 5.3 for more discussion.

Corollaries 1 and 2 can be extended to treat the more general case of approximate substitutes.[14]

Corollary 3: Consider game Γ, a δ-substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$ and a correlated equilibrium p^*. For some $\varepsilon \leq 4\delta$ there exists a correlated ε-equilibrium p' satisfying WGA where

$$\left| \frac{1}{|N_g|} \sum_{j \in N_g} U_j(p^*) - U_i(p') \right| \leq 2\delta \tag{5}$$

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

Thus, given any correlated equilibrium p^* there exists an approximate correlated equilibrium p' satisfying WGA and where the average payoff in each social group remains approximately the same. Recall that, for some δ, any partition into groups is a δ-substitute partition; thus Corollary 3 provides a lower bound on the size of ε required to ensure that there exists a correlated ε-equilibrium satisfying WGA.

Corollary 3 only shows the existence of an *approximate* correlated equilibrium. There need not exist a correlated equilibrium satisfying WGA. For example, suppose that $b = 10$, $e_1 = 9.5$ and $e_2 = 10.5$. In this case there exists a unique correlated equilibrium $p^* = (E, Z, E)$ in which players 1 and 3 are assigned role E. As already discussed this does not satisfy WGA. The distribution over roles p' where $p'(E, Z, E) = p'(Z, E, E) = 0.5$ does satisfy WGA but is not a correlated equilibrium because when allocated role E player 2 would do better to play Z (and get 0 rather than $-0.5 = 10 - 10.5$). Distribution p' is a correlated 0.5-equilibrium. The equilibrium p' thus requires player 2 to 'sacrifice 0.5' in order that the outcome satisfies WGA.

Corollary 4: Consider game Γ and a δ-substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$. For some $\varepsilon \leq 4\delta$, there exists a correlated ε-equilibrium p' of Γ satisfying WGA and GI.

Note that Corollary 4 only provides for the *existence* of an *approximate* correlated equilibrium. There may be a correlated equilibrium for which there is no nearby approximate correlated equilibrium satisfying WGA and GI. In Example 2 when there are actions E_1 and E_2 there may be a correlated equilibrium requiring the professor to use action E_1 if player 1 plays E and

[14] Note that Corollaries 3 and 4, unlike Corollaries 1 and 2, cannot be obtained from convexity of the set of correlated equilibria.

action E_2 if player 2 plays E. That is, it requires the textbook used to match the TA who will teach on the course. It may be that such an equilibrium could not be equated with a 'nearby equilibrium' that satisfies GI because it requires the professor to correlate his action with the TAs.

4.3 Subjective Beliefs

The assumption that players know the probability distribution p shall now be relaxed. Instead, players are modelled as having subjective beliefs about the correlating device. Specifically, there exists a *set of beliefs* $\{\beta^i\}_{i\in N}$, where $\beta^i \in P$ denotes the beliefs of player i. Thus, $\beta^i(\overline{a})$ denotes the probability that player i places on roles being assigned according to action profile $\overline{a}$. We say that the set of beliefs $\{\beta^i\}_{i\in N}$ constitutes 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 (as given by (1)) in the natural way by requiring no individual i to expect a payoff gain from changing strategy *given his beliefs, β^i*.

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). A framework of social identity, however, suggests certain properties, including stereotyping, that one might expect beliefs to satisfy. We propose a definition of stereotyped beliefs. Given game Γ, a δ-substitute partition into social groups Π and set of correlating devices $\mathcal{B} = \{p^i\}_{i\in N}$, we say that beliefs are *stereotyped and determined by $\mathcal{B}$* if

$$\beta^i(\overline{a}) = \frac{1}{|\mathcal{P}^{\Pi}(\overline{a})|} \sum_{\overline{a}'\in\mathcal{P}^{\Pi}(\overline{a})} p^i(\overline{a}') \qquad (6)$$

for all i and $\overline{a}$. Thus, each player expects the correlating device to satisfy WGA; this suggests stereotyping because players in the same social group are expected to behave identically (whether they actually do so or not). So, the professor may expect the two TAs to behave in the same way.

Different players may have different beliefs over the correlating device. We can imagine an objective device actually distributing roles according to a probability distribution $p \in P$. If $\beta^i = p$, then player i has beliefs consistent with the objective device. If $p^i = p$, then i's beliefs will not be consistent with distribution p if $p_j(k) \neq p_l(k)$ for some $j, l \in N_g$ and some social group

N_g. The beliefs of i are, however, related to the distribution p but subject to stereotyping. For instance, in Example 1 it may be that player 1 will put in effort and player 2 will be lazy so that $p(E, Z, E) = 1$. If player 3's beliefs are stereotyped then he may set $p(E, Z, E) = p(Z, E, E) = \frac{1}{2}$ and so expect one of the TAs to put in effort but does not know which one. If $p^i \neq p$, then i's beliefs are potentially far from consistent with the given allocation of roles. We discuss stereotyping in more detail in Section 5.1.

Given game Γ, let $\mathcal{CE}^\Gamma$ denote the set of correlated equilibria of game Γ. The following result demonstrates that if $\mathcal{B} \subset \mathcal{CE}^\Gamma$ and beliefs are determined by $\mathcal{B}$, then the result is an approximate correlated equilibrium of game Γ. Thus, it is an equilibrium for players to expect players in the same social group to behave in an identical fashion whether or not they actually do so in reality.

Theorem 1: Consider game Γ, a δ-substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$ and set of correlating devices $\mathcal{B} = \{p^i\}_{i \in N}$. If $\mathcal{B}$ is a subset of $\mathcal{CE}^\Gamma$ and beliefs are stereotyped and determined by $\mathcal{B}$, then $\{\beta^i\}_{i \in N}$ is a subjective correlated ε-equilibrium where $\varepsilon \leq 4\delta$.

Informally, this result implies that for every correlated equilibrium there exists a corresponding approximate subjective correlated equilibrium with stereotyped beliefs. From this result it is simple to derive Corollaries 1 to 4 (see the Appendix). Note that p^i and β^i can differ across agents. If $p^i = p$ for all i and some p then $\beta^i = \beta$ for all i and some β. This may seem like a natural case and merely generalizes the previous analysis (Corollaries 1 to 4) to one with stereotyping. By contrast, if $p^i \neq p^j$ for some i and j then $\beta^i \neq \beta^j$ and so i and j not only have stereotyped beliefs but also have different beliefs. This may be appropriate if i and j belong to different social groups and have, say, different beliefs about the distribution over types in their own social group. So, for instance the two TAs and the professor may have different beliefs.

5 Further Discussion

In this section we discuss four issues, using the framework introduced in the previous two sections, to demonstrate the insights that are possible from the framework. These issues are stereotyping, fairness, optimal group size and semi-anonymity.

5.1 Stereotyping

An interesting issue is whether stereotyping is costly to players. It can be costly for players to stereotype within their own social group as we can illustrate by setting $b = 10$ and $e_1 = e_2 = 6$ in Example 1. If we set player 3 as choosing E players 1 and 2 (the two TAs) play the following matrix game

	E	Z
E	$7, 7, 13$	$4, 10, 10$
Z	$10, 4, 10$	$0, 0, 0.$

Consider correlating device p where $p(E, E, E) = \frac{1}{3}$ and $p(E, Z, E) = \frac{2}{3}$. The device p is not a correlated equilibrium. To see why, note that player 2 knows that player 1 has probability one of being assigned role E. If player 1 behaves according to his assigned role then it is therefore better for player 2 to choose Z than E. So, if player 2 is assigned role E he should choose Z. Basically, if player 2 knows that player 1 will consider himself the most intelligent TA then player 2 does better to be lazy.

Suppose now that $\mathcal{B} = \{p, p, p\}$ and beliefs are stereotyped and determined by $\mathcal{B}$. This implies that, $\beta^2(E, E, E) = \frac{1}{3}, \beta^2(E, Z, E) = \frac{1}{3}$ and $\beta^2(Z, E, E) = \frac{1}{3}$ (with the same for players 1 and 3). If player 2 is assigned role E then he expects that player 1 is assigned role E with probability 0.5 and role Z with probability 0.5. This implies that player 2's expected payoffs from E and Z are 5.5 and 5 respectively and so it is individually rational for player 2 to choose E if assigned role E. Basically, if player 2 believes that he is the most intelligent TA then he expects that player 1 will not consider himself the most intelligent TA with probability one half. Given device p this is actually an erroneous belief. The set $\{\beta^1, \beta^2, \beta^3\}$ is, however, a subjective correlated equilibrium. The 'cost' to player 2 of stereotyping in this example is 1 because with probability $\frac{1}{3}$ he will be assigned role E and would earn 3 more by choosing Z.

Stereotyping proves costly or distortionary in this example because a player stereotypes behavior within his own social group. Player 2, because of stereotyping, has incorrect beliefs about the probability that player 1 is allocated to each role. This has the consequence that, depending on whether or not he stereotypes, a player can have different beliefs about his own probability of being allocated each role. In the example, when beliefs are stereotyped, player 2 expects to have role E with probability $\frac{2}{3}$. When beliefs are not stereotyped, he expects to have role E with probability $\frac{1}{3}$.

It is natural to think of stereotyping in terms of a player forming beliefs on how players outside of his social group will behave. It is not so clear whether stereotyping should influence a player's belief about his own (marginal) probability distribution of being assigned each role. If stereotyping does not influence a player's belief about his own probability of being assigned each role then we find that stereotyping is less distortionary or costly. To formalize this consider permutations of an action profile $\overline{a}$ for which player i's action does not change. More precisely, given action profile $\overline{a}$ and player i (and the set $\mathcal{P}^{\Pi}(\overline{a})$ of action profiles that are permutations of $\overline{a}$) let $\mathcal{P}_i^{\Pi}(\overline{a})$ denote the subset of $\mathcal{P}^{\Pi}(\overline{a})$ where $\overline{a}'_i = \overline{a}_i$. Given game Γ, a δ-substitute partition into social groups Π and set of correlating devices $\mathcal{B} = \{p^i\}_{i \in N}$, we say that beliefs are *other-stereotyped* and determined by $\mathcal{B}$ if

$$\beta^i(\overline{a}) = \frac{1}{\left|\mathcal{P}_i^{\Pi}(\overline{a})\right|} \sum_{\overline{a}' \in \mathcal{P}_i^{\Pi}(\overline{a})} p^i(\overline{a}') \tag{7}$$

for all i and $\overline{a}$. This revises the definition of stereotyped and determined by $\mathcal{B}$ in the sense that player i only stereotypes the behavior of players other than himself. Beliefs about his own probability of being assigned each role, given by β^i, are the same as those given by p^i. In Example 1, if $\mathcal{B} = \{p, p, p\}$ and beliefs are other-stereotyped and determined by $\mathcal{B}$ then β^i is the same as p for all i and so beliefs are the same as the objective correlating device. Note that this implies WGA does not hold. Players essentially expect WGA to hold for everyone but themselves.

If beliefs are stereotyped then we have seen above that the expected payoff can be significantly different to the payoff that would be expected if beliefs were consistent with the correlating device. The following result shows that this does not hold if beliefs are other stereotyped.

Proposition 1: Consider game Γ, a δ-substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$ and a set of correlating devices $\mathcal{B} = \{p^i\}_{i \in N}$. If beliefs are other-stereotyped and determined by $\mathcal{B}$, then,

$$\left|U_i(p^i) - U_i(\beta^i)\right| \leq \delta$$

for all i and all $s_i \in S$.[15]

[15] One could imagine a device signalling a set of roles for the entire player set, or alternatively, devices signalling roles for all the players in each social group. Players in one group may not be able (or willing) to differentiate between people in some other group (possibly with some distinguishing characteristic such as gender) and thus assign the average role distribution to members of the other group.

If Π is an exact substitute partition and so $\delta = 0$ then Proposition 1 suggests that there are no costs or distortions from a player stereotyping the behavior of others. In particular, his expected payoff is the same whether or not his beliefs are stereotyped. If players within social groups are only approximate substitutes, then there may be potential costs to stereotyping, depending on the amount of heterogeneity.

To illustrate, consider again Example 2 with actions E_1 and E_2. If player 3 plays action E_1 suppose that he gets bonus 10 if player 1 or player 2 choose E. If he chooses E_2 then he gets bonus 9.9 if player 1 chooses E and bonus 10.3 if player 2 chooses E. So, if the professor knows that player 1 will put in effort he is better using textbook 1 instead of 2. Correlating device $p(E, Z, E_2) = 1$ is not therefore a correlated equilibrium because player 3 would do better to play E_1 if assigned role E_2. Suppose, however, that player 3 has stereotyped beliefs β^3, where $\beta^3(Z, E, E_2) = \beta^3(E, Z, E_2) = 0.5$. It is now individually rational for player 3 to play E_2 because his expected payoff from doing so is 10.1. Basically, if player 3 does not know which TA will put in effort then he does better by choosing textbook 2. We see, therefore, that β^3 is a correlated equilibrium. If player 3 did not have stereotyped beliefs but actually knew the correlating device then he would do better by 0.1.

More generally, it follows from Proposition 1, that for any player i whose beliefs β^i are other stereotyped and determined by p^i, that if it is individually rational for player i to play k when assigned role k then player i could gain by at most 2δ by deviating from k when assigned role k if he has non-stereotyped beliefs p^i. The definition of other stereotyped could be relaxed further so that a player only stereotypes players in certain groups. For instance, he may not stereotype those in his own group. This would lead to the same conclusions.

A further thing to note from the previous example is how the two TAs may receive a lower or higher payoff if player 3 stereotypes. In particular, that player 3 stereotypes results in him using E_2 rather than E_1. The two TAs may or may not prefer this and indeed the amount of payoff difference could be arbitrarily large or small. Thus, while stereotyping may have little impact on the player with stereotyped beliefs (player 3 in this example) it can have a large impact on those who are being stereotyped (players 1 and 2 in this example).

5.2 Fairness and social contract

One motivation for WGA is to have equality of opportunity within groups. This also implies approximate equality of expected payoff. In particular, Corollary 3 implies that players in the same social group should get approximately the same payoff given a correlated equilibrium p' satisfying WGA. Specifically, it implies that $|U_i(p') - U_j(p')| \leq 4\delta$ for all i, j belonging to the same social group. We can, however, do better than this.

Proposition 2: Consider game Γ, a δ-substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$ and a correlating device p satisfying WGA. Then

$$|U_i(p) - U_j(p)| \leq \delta$$

for all $i, j \in N_g$, and all N_g.

Proposition 2 suggests that if a correlating device satisfies WGA, then expected payoffs are fair. Ex-post payoffs may, however, neither be fair or individually rational. We can see this from the example of Section 5.1 where $b = 10$ and $e_1 = e_2 = 6$. There exists correlated equilibrium p' where $p'(E, E, E) = p'(E, Z, E) = p'(Z, E, E) = \frac{1}{3}$. With probability two thirds the outcome is unfair with one of the TAs having an effort cost and the other not. With probability one third the outcome is not individually rational with both TAs having an incentive to choose Z because the other is choosing E.

To use a Rawlsian thought experiment one could suppose that all players (in the same social group) are identical or, in other words, exact substitutes. A correlated equilibrium p^* that satisfies WGA would have the property that all players expect to get the same payoff. This would be an acceptable social contract under Rawls's reasoning (Rawls 1972). The criticism often made, however, of the Rawls notion of social contract (e.g., Binmore 1989) is that ex-post outcomes need be neither fair nor individually rational, leading to questions of whether such a notion represents an appropriate form of social contract. In particular, given the lack of individual rationality, it is questionable whether commitments could be enforced. The following result is, however, easily obtained.

Corollary 5: Consider game Γ and a δ-substitute partition into social groups $\Pi = \{N_1, ..., N_G\}$. If Γ has a pure-strategy Nash equilibrium, then there exists a correlated 4δ-equilibrium p' satisfying WGA and GI and where, if $p'(\overline{a}) > 0$, then $\overline{a}$ is a Nash 4δ-equilibrium of game Γ.

This result does not solve the problem of ex-post fairness but does at least solve the problem of ex-post individual rationality. For instance, in Example 1 there exists a correlated equilibrium p' where $p'(E, Z, E) = p'(Z, E, E) =$

$\frac{1}{2}$. Ex-post the outcome is not fair as payoffs differ by 6 but the outcome is self-enforcing given that both potential outcomes are individually rational.

5.3 Group size

If players within social groups are not exact substitutes, within-group anonymity can lead to costs, as illustrated by the fact that only approximate correlated equilibria satisfying WGA may exist. Thus, if correlation between social groups is possible, the optimal group size would be one. The larger and more heterogeneous is a group, then the larger are the potential costs of within-group anonymity. If, however, group independence is assumed, then there are countervailing gains to larger groups because the larger is the social group, then the greater is the scope for mutually beneficial correlation of actions.[16] These trade-offs suggest an optimal group size determined by the heterogeneity of players and the potential gains from correlating actions. [17]

To illustrate we can again return to the example of Section 5.1 where $b = 10$ and $e_1 = e_2 = 6$. If we think of all three players as constituting a social group (i.e. we do not group the two TAs together) then the best we can do if we impose GI is correlated equilibrium p^1 where $p^1(E, Z, E) = 1$. This gives payoffs, $4, 10$ and 7 and an average payoff of 8. Grouping players 1 and 2 into a social group means that correlated equilibrium p' where $p'(E, E, E) = p'(E, Z, E) = p'(Z, E, E) = \frac{1}{3}$ now satisfies GI and WGA. This gives expected payoffs of $7, 7$ and 11 with an average payoff of $8\frac{1}{3}$. Thus, grouping players 1 and 2 together potentially increases the average payoff. Finally we can consider grouping all three players together. This results in there being no correlated equilibrium that satisfies WGA [except (Z, Z, Z). The best we can do is correlated 3-equilibrium p^2 where $p^2(E, E, E) = 1$ giving payoffs of 7. The optimal partition into groups in this example clearly groups players 1 and 2 together.

More formally, consider a game Γ, let $\mathcal{CE} \subset P$ denote the set of correlated equilibrium of Γ and let $\mathcal{U} \subset \mathbb{R}^n$ denote the set of expected payoff vectors attainable through a correlated equilibrium. That is, if $\{\overline{U}_1, ..., \overline{U}_n\} \in \mathcal{U}$, then there exists some $p \in \mathcal{CE}$ such that $\overline{U}_i = U_i(p)$ for all i. Now, consider

[16] If WGA is not required, the optimal group size is n.

[17] Related is the issue of the number of nations as modeled, for example, by Alesina and Spolaore (1997); larger countries imply benefits from greater internal efficiencies, security and ability to cope with external shocks but also imply greater heterogeneity and thus a problem in 'keeping everyone happy'. Similar conditions arise in economies with clubs and/or local public goods; see, for example, the survey articles Conley and Smith (2005), Demange (2005) and Le Breton and Weber (2005).

a partition of N into social groups $\Pi = \{N_1, ..., N_G\}$. Denote by $\mathcal{CE}^{\Pi}_{WGA}$ the set of correlated equilibria that satisfy WGA and let $\mathcal{U}^{\Pi}_{WGA}$ denote the corresponding set of payoffs. Clearly, if $G = n$ (implying that groups are singletons), then $\mathcal{U}^{\Pi}_{WGA} = \mathcal{U}$. If, however, at least some social groups are non-singleton, then it is to be expected that $\mathcal{U}^{\Pi}_{WGA}$ is a strict subset of $\mathcal{U}$. More precisely, if a partition of N into social groups $\Pi' = \{N_{1'}, ..., N_{G'}\}$ has the property that $N_g \subset N_{g'}$ for every $N_g \in \Pi$ and some $N_{g'} \in \Pi'$, then $\mathcal{U}^{\Pi'}_{WGA} \subset \mathcal{U}^{\Pi}_{WGA}$. Informally, the set of payoff vectors realizable through a correlated equilibrium satisfying WGA is decreasing in the size of groups. Denote by $\mathcal{CE}^{\Pi}_{GI}$ the set of correlated equilibria that satisfy GI (and not necessarily WGA) and denote by $\mathcal{U}^{\Pi}_{GI}$ the corresponding set of payoffs. Now, if $G = 1$, then $\mathcal{U}^{\Pi}_{GI} = \mathcal{U}$. If, however, the player set is partitioned into smaller groups, it is to be expected that $\mathcal{U}^{\Pi}_{GI}$ is a strict subset of $\mathcal{U}$. To re-use the notation above it is the case that $\mathcal{U}^{\Pi'}_{GI} \supset \mathcal{U}^{\Pi}_{GI}$. Informally, the set of payoff vectors realizable through a correlated equilibrium satisfying GI is increasing in the size of groups.

The issues of optimal group size appears to be interesting but complicated. For example, the issue of what constitutes an optimal group must first be addressed. One way to proceed (which could be done in the example above) is to use a Pareto criterion to rank outcomes. Another way would be to look at coalition formation with, for instance, players 1 and 2 in the above example having incentives to group together. A second issue is to what extent individual rationality can be traded for 'better outcomes' in the sense that players would be willing to sacrifice individual rationality for an equilibrium that satisfies both WGA and GI. Finally, one could question whether partitioning players into social groups always leads to social stratification where similar people group together (Gravel and Thoron 2007, for example). A full investigation of these issues is beyond the scope of this paper.

5.4 Semi-anonymous games

In this section we connect the concepts of semi-anonymous games (Kalai 2004, Wooders, Cartwright and Selten 2006) with δ-substitutability. This will provide a class of games where the player set can be partitioned into social groups of approximate substitutes. Let Ω denote a metric space of *player attributes*, let C denote a space of *crowding attributes* and let T denote a space of *taste attributes*, where $\Omega \equiv C \times T$. In interpretation, a player's attributes detail how he influences others (his crowding attribute) and his payoff function (his taste attribute). In the example, attributes detail the

b, e_1 and e_2. Two players can be seen as exact substitutes if they have the same attributes. A game is semi-anonymous if a player's payoff depends on the number of players of each *crowding* attribute playing each action.[18] For example, the professor's payoff depends on the number of TAs who choose high effort (and not their identities). If a game is semi-anonymous, then two players of the same crowding attribute are 0-interaction substitutes. More formally:

Given a finite player set N, an *attribute function* α maps N into Ω, where $\alpha(i) \equiv (\alpha_C(i), \alpha_T(i))$ details the attributes of player i. A *weight function* w maps $C \times A$ (where A, as previously, denotes a finite set of actions) into the set of non-negative integers $\mathbb{Z}_+$ and is said to be relative to action profile $\overline{a}$ if

$$w_{\overline{a}}(c,a) \equiv |\{i \in N : \alpha_C(i) = c \text{ and } \overline{a}_i = a\}|$$

for all $c \in C$ and $a \in A$. The integer $w_{\overline{a}}(c,a)$ is the number of players with crowding type c who are assigned the action a by the action profile $\overline{a}$. To an attribute function α we associate a game $\Gamma(N,\alpha) \equiv (N, A, \{u_i^\alpha\}_{i \in N})$. If $u_i^\alpha(\overline{a}) = u_i^\alpha(\overline{a}')$ for any i and any $\overline{a}, \overline{a}'$, where $w_{\overline{a}} = w_{\overline{a}'}$ and $\overline{a}_i = \overline{a}'_i$, then game $\Gamma(N,\alpha)$ is *semi-anonymous*. Semi-anonymity has bite only if multiple players have the same or similar crowding attributes. If all players have the same crowding attribute, then the payoff to a player depends only on his action and the proportion of players playing each action.

If two players have approximately the same attribute, then we can think of them as approximate substitutes. Let *dist* be the metric on the space of player attributes. We say that a partition Π of the player set N into social groups is a δ-attribute partition if $dist(\alpha(i), \alpha(j)) \leq \delta$ for any $i, j \in N_g$ and any $N_g \in \Pi$. Suppose a *continuity in attributes* assumption holds whereby

$$\left| u_i^\alpha(\overline{a}) - u_i^{\alpha'}(\overline{a}) \right| \leq 2\delta$$

for all $i \in N$, any $\delta \geq 0$, any action profile $\overline{a}$ and any game $\Gamma(N,\alpha')$, where $dist(\alpha(i), \alpha'(i)) \leq \delta$. Thus, a slight perturbation of attributes only slightly changes payoffs. In the example, a slight perturbation in b, e_1 or e_2 would only change payoffs slightly. We obtain:

Proposition 3: Let $\Gamma(N,\alpha)$ be any game, let Π be a δ-attribute partition and let $\overline{a}$ be an action profile. If $\overline{a}' \in \mathcal{P}^\Pi(\overline{a})$ and $\overline{a}'_j = \overline{a}_i$, where $i, j \in N_g$

[18] The term 'semi-anonymous games' is from Kalai (2004) while the term 'crowding attribute,' motivated by earlier research of Conley and Wooders (2001), is from Wooders, Catwright and Selten (2006).

for some $N_g \in \Pi$, then

$$\left| u_i^\alpha(\overline{a}) - u_j^\alpha(\overline{a}') \right| \leq 2\delta.$$

A look at Lemma 1 and the proof of Theorem 1 in the Appendix shows that all, using Proposition 3, our results could be stated for a δ-attribute partition as opposed to a δ-substitute partition. Indeed, this could be seen as a generalization because two players belonging to the same social group in a δ-attribute partition need not be δ-substitutes.

6 Concluding remarks

The concept of correlated equilibrium is a very appealing one. In interpretation, however, one needs to question where the role-allocation device comes from and how players learn to interpret signals. In the current paper we argue that social identity acts as a form of equilibrium-selection device whereby only correlated equilibria satisfying certain properties are likely to emerge – namely, equilibria satisfying within-group anonymity and group-independence properties. We demonstrate the existence of an approximate correlated equilibrium satisfying these properties. Also, we show that stereotyping of individuals within social groups is not costly.

One may want the role-allocation device to be endogenous; that is, it would be desirable to be able to model how players can endogenously develop a coordinated way of recognizing and interpreting random signals from nature or pre-play communication. An endogenous role-allocation device would enable one to determine from the model whether or not WGA and GI can be expected to emerge as properties of the device. The research of Hart and Mas-Colell (summarized in Hart 2005) demonstrates how naive learning heuristics such as regret matching can lead to aggregate play corresponding to a correlated equilibrium. The approach of Hart and Mas-Colell is, however, framed in a myopic setting in which correlation arises without any social context or social influence. What would happen if an element of social influence, such as desires for within-group fairness, are added to learning? More generally, do we observe play converging on correlated equilibria that exhibit the properties considered in this paper?

One would also like social groups to be endogenous. It may be possible to address this as a coalition-formation problem either in a noncooperative/cooperative framework such as in Perry and Reny (1996), or more recent work on economies with local public goods or many-to-any matching problems, such as Konishi and Unver (2006), or through a network approach similar to those described in Jackson (2005). The issue of optimal group size and composition is also related to the issue of endogeneity of group sizes.

A related issue is to consider communication as opposed to correlated equilibrium (Forges 1987). Communication equilibrium is the extension of correlated equilibrium to games in extensive form where communication and signals are possible during the game and not just prior to play. In endogenizing the allocating device and social group membership, it would be natural to model more explicitly the process of communication between players, not only before the game but during the play of the game (or, if thinking of repeated plays of a stage game, between plays of the stage game).

7 Appendix

Lemma 1: Given game Γ, a δ-substitute partition into social groups Π and any action profile $\overline{a}$, if $\overline{a}' \in \mathcal{P}^{\Pi}(\overline{a})$ and $\overline{a}'_j = \overline{a}_i$, where i and j are δ-substitutes then

$$\left|u_i(\overline{a}) - u_j(\overline{a}')\right| \leq \delta + \frac{n-1}{n}\delta.$$

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

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

- $\overline{a}^1_1 = \overline{a}_{\gamma(1)}$ and $\overline{a}^1_{\gamma(1)} = \overline{a}_1$ while $\overline{a}^1_z = \overline{a}_z$ for all other $z \in N$
- $\gamma^2(\gamma^{-1}(1)) = \gamma(1)$ and $\gamma^2(z) = \gamma(z)$ for all other z.

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

We shall demonstrate that $\overline{a}^n_l = \overline{a}'_l$ for all l. To do so we proceed in two stages, showing that (i) if $\overline{a}^l_l = \overline{a}'_l$, then $\overline{a}^n_l = \overline{a}'_l$ and (ii) $\overline{a}^l_l = \overline{a}'_l$.

Stage (i). The value of $\gamma^{-1}(l) = z^*$ is unique for all l. So, if $\gamma(l) = l$ then by construction $\gamma^h(l) = l$ for all h and $\gamma^h(z) \neq l$ for all $z \neq l$. Thus, if $\overline{a}_l^l = \overline{a}_l'$, then $\overline{a}_l^n = \overline{a}_l'$. Suppose that $\gamma^{-1}(l) = z^* \neq l$. We need to show that $\gamma^h(h) \neq l$ for all $h > l$. If $z^* > l$ then, by construction, $\gamma^{l^{-1}}(l) = z^*$ and $\gamma^{l+1}(z^*) = \gamma^l(l) \neq l$. Thus, there can be no agent $h > l$ such that $\gamma^h(h) = l$. If $z^* < l$ then, by construction, $\gamma^{z^*+1}(\gamma^{z^{*-1}}(z)) = \gamma^{z^*}(z^*) = l$ and $\gamma^{z^*+1}(z) \neq l$ for all $z \neq \gamma^{z^{*-1}}(z)$. If $\gamma^{z^{*-1}}(z) > l$ then we are done. Ultimately, $\gamma^{z^{*-1}}(z) \geq l$. Thus, if $\overline{a}_l^l = \overline{a}_l'$, then $\overline{a}_l^n = \overline{a}_l'$.

Stage (ii). We shall show that $\overline{a}_{\gamma^l(z)}^{l-1} = \overline{a}_{\gamma(z)}$ for all $2 \leq l \leq n$ and all $z \geq l$. If true this would imply that $\overline{a}_l^l = \overline{a}_{\gamma^l(l)}^{l-1} = \overline{a}_{\gamma(l)} = \overline{a}_l'$ and so $\overline{a}_l^n = \overline{a}_l'$. We use proof by induction. Set $l = 2$. There exists a unique player $z^* = \gamma^{-1}(1)$. For any $z \neq \gamma^{-1}(1)$ we have $\gamma^2(z) = \gamma(z)$, and so clearly $\overline{a}_{\gamma^2(z)}^1 = \overline{a}_{\gamma(z)}$. For $z^* = \gamma^{-1}(1)$ we have $\gamma^2(z^*) = \gamma(1)$. Thus, $\overline{a}_{\gamma^2(z^*)}^1 = \overline{a}_{\gamma(1)}^1 = \overline{a}_1 = \overline{a}_{\gamma(z^*)}$. Now suppose the conjecture holds for $l \geq 2$. Thus, $\overline{a}_{\gamma^l(z)}^{l-1} = \overline{a}_{\gamma(z)}$ for all $z \geq l$. There exists a unique player $z^* = \gamma^{l^{-1}}(l)$. For all $z \neq z^*$ we have $\gamma^{l+1}(z) = \gamma^l(z)$, and so $\overline{a}_{\gamma^{l+1}(z)}^l = \overline{a}_{\gamma^l(z)}^{l-1} = \overline{a}_{\gamma(z)}$. For z^* we have $\gamma^{l+1}(z^*) = \gamma^l(l)$, implying $\overline{a}_{\gamma^{l+1}(z^*)}^l = \overline{a}_{\gamma^l(l)}^l = \overline{a}_l^{l-1}$ where the last equality follows by construction. Finally, $\overline{a}_l^{l-1} = \overline{a}_{\gamma^l(z^*)}^{l-1} = \overline{a}_{\gamma(z^*)}$.

Now observe that $\overline{a}^1$ is a permutation of $\overline{a}$, and $\overline{a}^l$ is a permutation of $\overline{a}^{l-1}$. Without loss of generality, let $j = 1$ or re-index players so that $j = 1$. Note that $i = \gamma(j)$. Given that i and j are δ substitutes, we have

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

Given that $\overline{a}_j^2 = \overline{a}_j^1$ and players 2 and $\gamma^1(2)$ 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

$$\left|u_j(\overline{a}') - u_i(\overline{a})\right| \leq \delta + \frac{n-1}{n}\delta$$

completing the proof.■

Proof of Proposition 1: Given the proof of Lemma 1 and that $\overline{a}_i' = \overline{a}_i$ for all $\overline{a}' \in \mathcal{P}_i^{\Pi}(\overline{a})$, we have that,

$$\left|u_i(\overline{a}) - u_i\left(\overline{a}'\right)\right| \leq \delta$$

for any $\overline{a}' \in \mathcal{P}_i^{\Pi}(\overline{a})$ because players within the same social group are δ-interaction substitutes. The statement of the Proposition is now clear.■

Proof of Proposition 2: First note that, given $\overline{a} \in \mathcal{P}^{\Pi}(\overline{a})$ and $\overline{a}' \in \mathcal{P}^{\Pi}(\overline{a})$, then $\mathcal{P}^{\Pi}(\overline{a}') = \mathcal{P}^{\Pi}(\overline{a})$ for all $\overline{a}', \overline{a} \in A^N$. Thus, the set of action profiles A^N can be partitioned into a finite set of sets of actions profiles $\Psi^1, \Psi^2, ..., \Psi^L$, where $\Psi^l = \mathcal{P}^{\Pi}(\overline{a})$ for some $\overline{a} \in A^N$. We refer to a set Ψ^l as a *permutation class of action profiles*. Pick a permutation class Ψ^l and two players $i, j \in N_g$ for some N_g. For every $\overline{a} \in \Psi^l$ there exists $\overline{a}' \in \Psi^l$, where $\overline{a}_i = \overline{a}'_j, \overline{a}_j = \overline{a}'_i$ and $\overline{a}_l = \overline{a}'_l$ for all other $l \in N$. Given that i and j are δ-individual substitutes,

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

Finally, given that p satisfies WGA, we know $p(\overline{a}) = p(\overline{a}')$. So, letting $\overline{a}^l \in \Psi^l$

$$\begin{aligned} |U_i(p) - U_j(p)| &= \left|\sum_{\Psi^l} \sum_{\overline{a} \in \Psi^l} p(\overline{a}) u_i(\overline{a}) - \sum_{\Psi^l} \sum_{\overline{a} \in \Psi^l} p(\overline{a}) u_j(\overline{a})\right| \\ &\leq \sum_{\Psi^l} p(\overline{a}^l) \left|\Psi^l\right| \delta = \delta. \end{aligned}$$

This completes the proof.■

Proof of Proposition 3: Fix two players $i, j \in N_g$ for some N_g and fix two action profiles $\overline{a}$ and $\overline{a}' \in \mathcal{P}^{\Pi}(\overline{a})$ where $\overline{a}'_j = \overline{a}_i$. There exists a one-to-one mapping γ from N to N such that $\overline{a}'_{\gamma(l)} = \overline{a}_l$ and $dist(\alpha(\gamma(l)), \alpha(l)) \leq \delta$ for all $l \in N$ and $\gamma(i) = j$. Define attribute function α' where $\alpha'(l) := \alpha(\gamma(l))$ for all $l \in N$. Clearly, $dist(\alpha(l), \alpha'(l)) \leq \delta$ for all l and so $\left|u_i^{\alpha}(\overline{a}) - u_i^{\alpha'}(\overline{a})\right| \leq 2\delta$ for all i by continuity in attributes. Next note that pair $(\alpha'(l), \overline{a}_l) = (\alpha(\gamma(l)), \overline{a}'_{\gamma(l)})$ for all $l \in N$. Thus, $u_i^{\alpha'}(\overline{a}) = u_{\gamma(i)}^{\alpha}(\overline{a}')$ implying that $\left|u_i^{\alpha}(\overline{a}) - u_j^{\alpha}(\overline{a}')\right| \leq 2\delta$.■

Proof of Theorem 1: We need to show that

$$\sum_{\overline{a} \in A^N} \beta^i(\overline{a}) u_i(\overline{a}) \geq \sum_{\overline{a} \in A^N} \beta^i(\overline{a}) u_i(s_i(\overline{a}_i), \overline{a}_{-i}) - 4\delta \tag{8}$$

for all $i \in N$ and $s_i \in S$. We conjecture (*) that

$$\left|\sum_{\overline{a} \in A^N} \beta^i(\overline{a}) u_i(s(\overline{a}_i), \overline{a}_{-i}) - \frac{1}{|N_g|} \sum_{j \in N_g} \sum_{\overline{a} \in A^N} p^i(\overline{a}) u_j(s(\overline{a}_j), \overline{a}_{-j})\right| \leq 2\delta \tag{9}$$

for all $s \in S$, all $i \in N_g$ and all N_g. Given that p^i is a correlated equilibrium,

$$\frac{1}{|N_g|}\sum_{j\in N_g}\sum_{\overline{a}\in A^N} p^i(\overline{a})u_j(\overline{a}) \geq \frac{1}{|N_g|}\sum_{j\in N_g}\sum_{\overline{a}\in A^N} p^i(\overline{a})u_j(s(\overline{a}_j),\overline{a}_{-j}) \tag{10}$$

for all $s \in S$ and all N_g. Using (9) on both sides of the inequality in (10) gives (8) and the desired result.

It remains to prove the conjecture (*). Let $s \in S$ be any strategy, and let $i \in N_g$ be any player. By Lemma 1, for any permutation class Ψ^l [see the proof of Proposition 2 for the definition of a permutation class] and any $j \in N_g$,

$$\left|\sum_{\overline{a}\in\Psi^l} u_i(s(\overline{a}_i),\overline{a}_{-i}) - \sum_{\overline{a}\in\Psi^l} u_j(s(\overline{a}_j),\overline{a}_{-j})\right| \leq 2\delta$$

implying,

$$\left|\sum_{\overline{a}\in\Psi^l} u_i(s(\overline{a}_i),\overline{a}_{-i}) - \frac{1}{|N_g|}\sum_{j\in N_g}\sum_{\overline{a}\in\Psi^l} u_j(s(\overline{a}_j),\overline{a}_{-j})\right| \leq 2\delta.$$

Letting $\overline{a}^l \in \Psi^l$ and ϱ be some real number where $|\varrho| \leq 2\delta$ and using the definition of beliefs that are stereotyped and determined by $\mathcal{B}$ given by (6) we can obtain

$$\begin{aligned}
\sum_{\overline{a}\in\Psi^l} \beta^i(\overline{a})u_i(s(\overline{a}_i),\overline{a}_{-i}) &= \beta^i(\overline{a}^l)\left|\Psi^l\right|\sum_{\overline{a}\in\Psi^l} u_i(s(\overline{a}_i),\overline{a}_{-i}) \\
&= \beta^i(\overline{a}^l)\left|\Psi^l\right|\left[\frac{1}{|N_g|}\sum_{j\in N_g}\sum_{\overline{a}\in\Psi^l} u_j(s(\overline{a}_j),\overline{a}_{-j}) + \varrho\right] \\
&= \sum_{\overline{a}\in\Psi^l} p^i(\overline{a})\left[\frac{1}{|N_g|}\sum_{j\in N_g}\sum_{\overline{a}\in\Psi^l} u_j(s(\overline{a}_j),\overline{a}_{-j}) + \varrho\right] \\
&= \frac{1}{|N_g|}\sum_{j\in N_g}\sum_{\overline{a}\in\Psi^l} p^i(\overline{a})u_j(s(\overline{a}_j),\overline{a}_{-j}) + \varrho\sum_{\overline{a}\in\Psi^l} p^i(\overline{a}).
\end{aligned}$$

Summing over all Ψ^l, we get (9).■

Proof of Corollaries 1 to 5: To prove Corollary 3 (and therefore Corollary 1) set

$$p'(\overline{a}) = \frac{1}{|\mathcal{P}^{\Pi}(\overline{a})|}\sum_{\overline{a}\in\mathcal{P}^{\Pi}(\overline{a})} p^*(\overline{a}) \tag{11}$$

for all $\overline{a}$. By construction p' satisfies WGA and by Theorem 1 p' is a correlated 4δ-equilibrium. The equation (5) can then be derived from (9).

To prove Corollary 4 (and therefore Corollary 2) note that by standard Nash existence theorems game Γ has a Nash equilibrium (in mixed strategies). Thus, there exists a correlated equilibrium p^*, where roles are stochastically independent across players. Clearly p^* satisfies GI. Now, set p' as in (11). By construction p' satisfies WGA and GI. By Theorem 1 p' is a correlated 4δ-equilibrium.[19] To prove Corollary 5 we just note that by picking the pure strategy Nash equilibrium and applying the above reasoning we get the desired answer. ■

References

[1] Akerlof, G. A., (1980) "A theory of social custom of which unemployment may be one consequence," *Quarterly Journal of Economics* 94: 749-75.

[2] Akerlof, G. and R. Kranton, (2000) "Economics and identity," *The Quarterly Journal of Economics* 115: 715-753.

[3] Alesina, A. and E. Spolaore, (1997) "On the number and size of nations," *The Quarterly Journal of Economics* 112: 1027-1056.

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

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

[6] Bernheim, B. D. (1994) "A theory of conformity," *The Journal of Political Economy*, Vol. 102, No. 5. (Oct., 1994), pp. 841-877.

[7] Binmore, K., (1989) "Social contract I: Harsanyi and Rawls," *The Economic Journal* C: 84-102.

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

[19] Note that p' need not be 'fully' independent across players. For example (with three actions and two players) it may be that $p^*(1,3) = p^*(2,3) = \frac{1}{2}$ implying that p' will be such that $p'_1(3) = 0.5$ but $p'_1(3|r_2 = 1) = 1$.

[9] Brandts, J., T. Saijo and A. Schram (2004) "How universal is behavior? A four country comparison of spite and cooperation in voluntary contribution mechanisms", *Public Choice* 119: 381-424.

[10] Brown, R. (2000) "Social identity theory: past achievements, current problems and future challenges," *European Journal of Social Psychology* 30: 745-778.

[11] Conley, J. and S. Smith (2005) "Coalitions and clubs; Tiebout equilibrium in large economics," in *Group Formation in Economics; Networks, Clubs and Coalitions*, G. Demange and M. Wooders (eds.), Cambridge University Press, Cambridge UK.

[12] 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.

[13] Demange, G. (2005) "Group formation; The interaction of increasing returns and preference formation," in *Group Formation in Economics; Networks, Clubs and Coalitions*, G. Demange and M. Wooders (eds.), Cambridge University Press, Cambridge UK.

[14] Dhillon A., and J. F. Mertens, (1996) "Perfect correlated equilibria", *Journal of Economic Theory* 68: 279-302.

[15] Elster, J., (1989) "Social norms and economic theory," *Journal of Economic Perspectives* 3: 99-117.

[16] Fehr, E. and K. Schmidt (1999) "A Theory of Fairness, Competition and Cooperation" Quarterly Journal of Economics 114: 817-868.

[17] Forges, F., (1986) "An approach to communication equilibria," *Econometrica* 54: 1375-1385.

[18] Fudenberg, D. and J. Tirole (1998) *Game Theory,* MIT Press.

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

[20] Gravel, N. and S. Thoron (2007) "Does endogenous formation of jurisdictions lead to wealth-stratification?" *Journal of Economic Theory* 132: 569-583

[21] Haimanko, O., M. Le Breton and S. Weber (2004) “Voluntary formation of communities for the provision of public projects,” *Journal of Economic Theory* 115: 1-34.

[22] Hayek, F.A., (1960) *The Constitution of Liberty* Routledge and Keegan Paul Ltd. London.

[23] Hayek, F.A., (1982) *Law, Legislation and Liberty*, T.J. Press, Padstow.

[24] 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.

[25] 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.

[26] Hart, S., (2005) “Adaptive heuristics,” *Econometrica* 73: 1401-1430.

[27] Hogg, M., D. Terry and K. White, (1995) “A tale of two theories: a critical comparison of Identity Theory with Social Identity Theory,”, *Social Psychology Quarterly* 58: 255-269.

[28] Hogg, M.D., and G.M. Vaughan, (2005) *Social Psychology*, Pearson/Prentice Hall, Marlow.

[29] Jackson, M. (2005) “A survey of models of network formation: Stability and efficiency,” Chapter 1 in *Group Formation in Economics; Networks, Clubs and Coalitions* , edited by G. Demange and M. Wooders, Cambridge University Press, Cambridge U.K.

[30] Johnson, D.W., and F. Johnson, (1987) *Joining Together: Group Theory and Group Skills*, Prentice Hall, London.

[31] Konishi, H. and M. Utku Unver (2006), “Credible group-stability in many-to-may matching problems,” *Journal of Economic Theory* 129: 57-80.

[32] Kalai, E. (2004) “Large robust games,” *Econometrica* 72: 1631-1666.

[33] Myerson, R.B. (1986) “Acceptable and predominant correlated equilibria,” *International Journal of Game Theory* 15: 133-154.

[34] Rabin, M. (1993) “Incorporating fairness into game theory and economics,” *American Economic Review* 83: 1281-1302.

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

[36] Rawls, J., (1972) *A Theory of Justice* Oxford University Press.

[37] Roth, A., V. Prasnikar, M. Okuno-Fujiwara and S. Zamir (1991) "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: an Experimental Study", *American Economic Review* 81: 1068-1095.

[38] Sherif, M. (1966) *In Common Predicament: Social Psychology of Intergroup Conflict and Cooperation* Houghton Mifflin: Boston.

[39] Schelling, T. C. (1960) *The Strategy of Conflict*. Cambridge, Harvard University Press, 1960.

[40] Selten, R. (1980) "A note on Evolutionary Stable Strategies in Asymmetric Animal Conflicts", *Journal of Theoretical Biology* 84: 93-101.

[41] Tajfel, H., (1978) *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations*, Academic Press, London.

[42] 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.

[43] Wooders, M., E. Cartwright and R. Selten (2006) "Behavioral conformity in games with many players," *Games and Economic Behavior*, 57: 347-360 .