On Equilibrium in Pure Stategies in Games with Many Players

Edward Cartwright

And

Myrna Wooders

No 686

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

On Equilibrium in Pure Strategies in Games with Many Players*

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

October 26, 2003

*This paper continues research initiated in Wooders, Cartwright and Selten (2001). We are indebted to Bhaskar Dutta and Ehud Kalai for stimulating and helpful comments.

Abstract: Motivated by issues of imitation, learning and evolution, we introduce a framework of noncooperative games, allowing both countable sets of pure actions and player types and demonstrate that for all games with sufficiently many players, every mixed strategy Nash equilibrium can be used to construct a Nash ε-equilibrium in pure strategies that is 'ε-equivalent'. Our framework introduces and exploits a distinction between crowding attributes of players (their external effects on others) and their taste attributes (their payoff functions and any other attributes that are not directly relevant to other players). The set of crowding attributes is assumed to be compact; this is not required, however, for taste attributes. We stress that for studying issues such as conformity, the case of a finite set of types and actions, while illuminating, cannot yield completely satisfactory results. Our main theorems are based on a new mathematical result, in the spirit of the Shapley-Folkman Theorem but applicable to a countable (not necessarily finite dimensional) strategy space.

1 Motivation for the study of approximate Nash equilibrium in pure strategies

The concept of a Nash equilibrium is at the heart of much of economics and game theory. It is thus fundamental to question when Nash equilibrium provides a good description of human behaviour. A number of challenges are posed by the evidence. Experimental evidence, for example, supports the view that individuals typically do not play mixed strategies (cf., Friedman

1996) and if they do, there may be serial correlation.[1] Challenges are also posed by the observed imitative nature of human behaviour (cf., Offerman, Potters and Sonnemans 2002). Moreover, an individual may typically only imitate others in certain groups of individuals with which he identifies (Gross 1996). The importance of equilibrium in pure strategies is evidenced by numerous papers in the literatures of game theory and economics (from, for example, Rosenthal 1973 to Cripps, Keller and Rady 2002). An important question for game theory is whether Nash equilibrium can be consistent with conformity in choice of strategy by 'similar' individuals and with the use of pure strategies.

Our prior research investigating these issues addressed the question of whether social conformity – that is, roughly, situations where most individuals imitate similar individuals – can be consistent with approximate Nash equilibrium.[2] It was assumed, throughout this research, that social conformity requires the use of pure strategies. In this paper, we treat in isolation the most basic question – the existence of an approximate Nash equilibrium in pure strategies. Our main result is that in games with many players, all induced from a common underlying structure, any Nash equilibrium in mixed strategies is approximately equivalent to an equilibrium in pure strategies, a 'purification' result.

Within our framework a player is characterized by his attribute, a point in a given set of attributes. An important feature incorporated into our model is a distinction between the crowding attribute of a player and his taste attribute.[3] A player's crowding attribute reflects those characteristics

[1] This has been demonstrated in a number of papers; see Walker and Wooders (2001) for a recent contribution and references therein.

[2] Wooders, Cartwright and Selten (2001).

[3] This terminology is taken from Conley and Wooders (1996,1997) who use the term

of the player that directly affect other players – for example, whether one chooses to go to a particular club may depend on the gender and composition of the membership and how attractive one finds a particular economics department may depend on the numbers of faculty engaged in various areas of research. In an evolutionary context, crowding attributes may be endogenously selected. In this paper we assume that the space of crowding attributes is a compact metric space but no such assumptions are made on the space of taste attributes.

We treat games of imperfect information. Thus, as well as having a certain attribute, a player is randomly assigned, by nature, a type (as in a standard game of incomplete information). In interpretation we can think of a player's crowding attribute as publicly observable while his taste attribute may or may not be observable. We allow a countable set of pure actions and a countable number of (Harsanyi) types. A new mathematical result, allowing us to approximate a mixed strategy vector by a pure strategy vector in which each player plays a pure strategy in the support of his mixed strategy, underlies our purification results, and allows the non-finiteness of strategy and type sets.

The framework of the current paper is, in important respects, more general than that treated in our prior research. In particular, our earlier work treated finite action and finite type sets and a compact set of attributes. In our research on conformity and research in progress studying evolutionary selection and imitation, the case of a finite set of types and actions and compact space of attributes, while illuminating, cannot yield completely satisfactory results. In particular, with the number of different actions and

'crowding types'.

different types of players uniformly bounded, then in large populations, some conformity must arise by necessity. To generalize the framework as we do in this paper is thus crucial.

We discuss related literature in more detail in Section 5 but we note that this paper contributes to a large literature on the 'purification' of a non-cooperative equilibrium as a consequence of a large numbers of players (e.g. Schmeidler 1973, Mas-Colell 1984, Khan 1989, 1998, Khan et al. 1997, Pascoa 1998, Khan and Sun 1999, Kalai 2002). Unlike the approach in this paper much of the literature treats games with a continuum of players. As is natural when modelling a game with many players we assume that a player's payoff depends on the actions of others through the induced joint distribution of strategies over crowding attributes, types and actions. In this respect our approach resembles that of Green (1984) and Pascoa (1993a, b,1998)[4]. One approach in the literature (e.g. Schmeidler 1973, Mas-Colell 1984, Kalai 2002[5]) is to assume that a player's payoff depends on the actions of others through an indiscriminating distribution over actions (or types and actions); this corresponds to a special case of our model in which there is at most a finite number of crowding attributes and types.[6] When payoffs depend on

[4]Note that these authors consider games of complete information with a continuum player set.

[5]We note that the research in Kalai (2002) is an outgrowth of an earlier 2000 working paper.

[6]Mas-Colell (1984) remarks that strategy sets can encode for a player's attribute. For example, the payoff function may be set up in such a way that a male would never rationally choose from a particular subset of strategies while a female may only rationally choose from that subset. Similarly, in games of incomplete information (as in Kalai 2002) a player's type may encode for his attribute. If, however, the set of strategies and the set of types are finite, as in Mas-Colell and in Kalai, then at most a finite number of crowding attributes can be encoded. We remark that, in contrast to our research in this paper

the distribution of actions over crowding attributes there are two alternative ways of interpreting a large player set. First, as a large number of players (of any attribute) or second as a large number of players with each attribute. Green (1984) assumes an uncountable number of players of each attribute while Pascoa (1993a,b,1998) considers either that there be an uncountable number of players with each attribute or a certain continuity property. In a finite setting we provide results for both possible interpretations of a large player set.

Besides a continuity condition, our main result requires an assumption on the universal payoff function - 'the large game property' - dictating that the actions of any 'small group' of players should have little influence on the payoffs of others. The large game property is sufficient to demonstrate that:

Purification: Given any $\varepsilon > 0$ there is an integer $\eta(\varepsilon)$ with the property that for every game Γ with at least $\eta(\varepsilon)$ players and for any Bayesian Nash equilibrium σ of Γ there exists a Bayesian Nash ε-equilibrium in pure strategies m that is ε equivalent (in payoffs) to σ.

Our second main result treat games in which for each player in a game there are many players who have similar crowding attributes - a 'thickness in the distribution of players over the set of crowding attributes'. In treating such games we are able to significantly weaken the assumption on payoff functions to a very mild continuity property.

Related literature is discussed in Section 5. We comment here, however, on a related literature concerning purification of Nash equilibria in finite

and also in Wooders, Cartwright and Selten (2001), these authors make no further use of dependence of payoffs on crowding attributes.

games with imperfect information. This literature demonstrates that if there sufficient uncertainty over the signals (or types) that players receive then any mixed strategy can be purified (e.g. Radner and Rosenthal 1982, Aumann et. al. 1983). Given that we model games of imperfect information it is important to emphasize that we do not treat this form of purification - our results also hold for games of perfect information.

We proceed as follows: Section 2 introduces definitions and notation. In Section 3 we treat purification, providing a simple example before defining the large game property and providing our two main results. In Section 4 we provide a brief discussion of the literature and Section 5 concludes the paper. Additional proofs are provided in an Appendix.

2 Bayesian games and noncooperative pregames

We begin this section by defining a Bayesian game and its components. The pregame framework is then introduced and we demonstrate how Bayesian games can be induced from a pregame. Next, we consider the strategies available to players in a Bayesian game and discuss expected payoffs. We finish by defining a Nash equilibrium.

2.1 A Bayesian game

A *Bayesian game* Γ is given by a tuple (N, A, T, g, u) where N is a finite *player set*, A is a set of *action profiles*, T is a set of *type profiles*, g is a *probability distribution over type profiles* and u is a set of *utility functions*. We define these components in turn.

Let $N = \{1, ..., n\}$ be a finite player set, let $\mathcal{A}$ denote a countable set of *actions* and let $\mathcal{T}$ denote a countable set of *types*. 'Nature' assigns each player a type. Informed of his own type but not the types of his opponents, each player chooses an action. Let $A \equiv \mathcal{A}^N$ be the set of *action profiles* and let $T \equiv \mathcal{T}^N$ be the set of *type profiles*. Given action profile a and type profile t we interpret a_i and t_i as respectively the action and type of player $i \in N$.

A player's payoff depends on the actions and types of players. Formally, in game Γ, for each player $i \in N$ there exists a *utility function* $u_i : A \times T \to \mathbb{R}$. In interpretation $u_i(a, t)$ denotes the payoff of player i if the action profile is a and the type profile t. Let u denote the set of utility functions.

A player, once informed of his own type, selects an action without knowing the types of the complementary player set. A player thus forms beliefs over the types he expects others to be. These beliefs are represented by a 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. Throughout we will assume *consistent beliefs*. Formally, for some probability distribution over type profiles g, we assume:

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

for all $i \in N$ and $t_i \in \mathcal{T}$.[7] We denote by $\mathcal{T}_i$ the set of types $t_i \in \mathcal{T}$ such that $\sum_{t'_{-i} \in T_{-i}} g(t_i, t'_{-i})$. Thus, player i will be a type $t^z \in \mathcal{T}_i$.

2.2 Noncooperative pregames

To treat a family of games all induced from a common strategic situation we first introduce a space of *player attributes*, denoted by Ω. An attribute $\omega \in \Omega$ is composed of two elements - a taste attribute and a crowding

[7] We do not require (1) to hold if $\sum_{t'_{-i} \in T_{-i}} g(t_i, t'_{-i}) = 0$; i.e. if there is no probability that player i is type t_i.

attribute. In interpretation, the crowding attribute of a player describes those characteristics that might affect other players, for example, gender, ability to do the salsa, educational level, and so on. Let $\mathcal{P}$ denote a set of *taste attributes* and let $\mathcal{C}$ denote a set of *crowding attributes*. We assume that $\mathcal{P} \times \mathcal{C} = \Omega$. If a player i has attribute $\omega = (\pi, c)$ then π is interpreted as giving her payoff function and c is interpreted as determining how her strategy choice influences the payoffs of others. We will assume that $\mathcal{C}$ is a compact metric space (while no assumptions are made on $\mathcal{P}$).

Let N be a finite *player set*. A function α mapping from N to Ω is called an *attribute function*. The pair (N, α) is a *population*. While an attribute consists of a taste attribute/crowding attribute pair, crowding attributes play a special role. Thus, given an attribute function α we denote by κ the projection of α onto $\mathcal{C}$. Given population (N, α) the attribute of player i is therefore $\alpha(i)$ and the crowding attribute of player i is $\kappa(i)$ where $\alpha(i) = (\pi, \kappa(i))$ for some $\pi \in \mathcal{P}$. Taking as given a countable set of actions $\mathcal{A}$ and types $\mathcal{T}$ a population (N, α) induces a Bayesian game $\Gamma(N, \alpha) \equiv (N, A, T, g^{\alpha}, u^{\alpha})$ as we now formalize.

Denote by W the set of all mappings from $\mathcal{C} \times \mathcal{A} \times \mathcal{T}$ into $\mathbb{Z}_{+}$, the non-negative integers. A member of W is called a *weight function.* Given population (N, α) we say that weight function $w_{\alpha,a,t}$ is *relative to action profile a and type profile t* if:

$$w_{\alpha,a,t}(c, a^l, t^z) = \left| \left\{ i \in N : \kappa(i) = c, a_i = a^l \text{ and } t_i = t^z \right\} \right| .$$

Thus, $w(c, a^l, t^z)$ denotes the number of players with crowding attribute c and type t^z who play action a^l. A *universal payoff function* h maps $\Omega \times \mathcal{A} \times \mathcal{T} \times W$ into $\mathbb{R}_{+}$. Given a population (N, α) the function h will determine the payoff function u_i^{α} of any player $i \in N$. The payoff of player i will depend on his attribute, his action, his type and the weight function induced by the attributes, actions and types of the complementary player set. Formally, given an action profile a and a type profile t:

$$u_i^\alpha(a,t) = h(\alpha(i), a_i, t_i, w_{\alpha,a,t}).$$

Denote by D the set of all mappings from $\Omega \times \mathcal{T}$ into $\mathbb{Z}_+$. A member of D is called a *type function*. Given population (N, α) we say that type function $d_{\alpha,t}$ is *relative to type profile* t if:

$$d_{\alpha,t}(\omega, t^z) = |\{i \in N : \alpha(i) = \omega \text{ and } t_i = t^z\}|\,.$$

Thus, $d_{\alpha,t}(\omega, t)$ denotes the number of players with attribute ω and type t^z.[8] A *universal beliefs function* b maps D into $[0,1]$. The value $b(d_{\alpha,t})$ is interpreted as the probability of type profile t. Formally:

$$g^\alpha(t) = b(d_{\alpha,t})$$

where g^α is a probability distribution over type profiles for the population (N, α) induced from the universal beliefs function b. Players are assumed to have consistent beliefs with respect to g^α. It is important to realize the differences between functions g^α and b. Function g^α is defined relative to a population (N, α) and its domain is $\mathcal{T}^N$. Function b, however, is defined independently of any specific game and has domain D.[9]

A *pregame* is given by a tuple $\mathcal{G} = (\Omega, \mathcal{A}, \mathcal{T}, b, h)$, consisting of a compact metric space Ω, countable sets $\mathcal{A}$ and $\mathcal{T}$, functions $b : D \longrightarrow [0,1]$ and

[8] Note that $d_{\alpha,t}$ is a projection of $w_{\alpha,a,t}$ onto $\Omega \times \mathcal{T}$.

[9] Also, summing g^α over its domain gives a value of one - because it describes a unique population - while the sum of b over its domain is non-finite - because it describes beliefs for any population.

$h : \Omega \times \mathcal{A} \times \mathcal{T} \times W \longrightarrow \mathbb{R}_{+}$. As discussed above we refer to a population (N, α) as inducing, through the pregame, a Bayesian game $\Gamma(N, \alpha) \equiv (N, A, T, g^{\alpha}, u^{\alpha})$.

2.3 Strategies and expected payoffs

Take as given a population (N, α) and induced Bayesian game $(N, A, T, g^{\alpha}, u^{\alpha})$. Knowing his own type, but not those of his opponents a player chooses an action. 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. Let $\mathcal{S}$ denote the set of 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 $x(k)$ the probability that a player plays pure strategy $k \in \mathcal{S}$. We denote by $x(a^l|t^z)$ the probability that a player 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.[10] Given a strategy vector σ, a type $t^z \in \mathcal{T}_i$ and beliefs about the type profile p_i^{α} 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:

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

[10] We use the vNM assumption for convenience but our results do not depend on it: To derive our main results we impose either a large game property or a continuity property and in doing so impose all the assumptions needed on the U_i^{α} functions. Neither the large game property or continuity property require the vNM assumption to hold.

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 $U_i^\alpha(\cdot|t^z) : \Sigma \to \mathbb{R}$ denote the expected utility function of player i conditional on the type of player i being t^z where:

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

2.4 Nash equilibrium and purification

The standard definition of a Bayesian Nash equilibrium applies. A strategy vector σ is a *Bayesian Nash ε-equilibrium* (or informally an approximate Bayesian Nash equilibrium) if and only if:

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

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

Given a game $\Gamma(N, \alpha)$ we say that two strategy profiles σ and m are *ε-equivalent* if, for all $i \in N$ and $t^z \in \mathcal{T}_i$:

$$|U_i^\alpha(m|t^z) - U_i^\alpha(\sigma|t^z)| \leq \varepsilon.$$

We say that a strategy profile σ can be *ε-purified* if there exists a strategy profile m that is degenerate and ε-equivalent to σ.[11]

[11] A related notion of ε-purification was introduced by Aumann et. al. (1983). There, the notion of ε-purification is relative to strategies and not strategy vectors. Thus, two strategies p and t are *ε-equivalent* for player i if $|U_i^\alpha(p, \sigma_{-i}) - U_i^\alpha(t, \sigma_{-i})| < \varepsilon$ for any $\sigma_{-i} \in \Sigma^{N \setminus \{i\}}$. This definition proves useful in considering games of incomplete information but is too restrictive to be of use in considering games of complete information.

3 Purification

Before providing our main results it may be useful to provide a simple example:

Example 1: There are two crowding attributes - *rich* and *poor*. Players must choose one of two pure strategies or locations A and B. A poor player prefers living with rich players and thus his payoff is equal to the proportion of rich players whose choice of location he matches. A rich player prefers to not live with poor players and thus his payoff is equal to the proportion of poor players whose choice of location he does not match.

Any game induced from this pregame has a Nash equilibrium. It is simple to see, however, that if there exists an odd number of either rich or poor players then there does not exist a Nash equilibrium in pure strategies. Also, if either the number of rich players or the number of poor players is small then there need not exist an approximate Nash equilibrium in pure strategies, no matter how large the total population.

Our first main result (Theorem 2) demonstrates that if a pregame satisfies a large game property then, in any induced game with sufficiently many players, any Nash equilibrium can be approximately purified. The pregame of Example 1 does not satisfy the large game property; the large game property requires that any small group of players have diminishing influence in populations with a larger player set.

Our second main result (Theorem 3) demonstrates that if a pregame satisfies a mild continuity property then, in any game 'with a thick distribution of attributes', there exists an approximate Nash equilibrium in pure strategies. The pregame of Example 1 satisfies the continuity property; applying Theorem 3 demonstrates that if there are sufficiently many players who are rich and also sufficiently many who are poor then there exists an approximate Nash equilibrium in pure strategies.

3.1 Approximating mixed strategy profiles by pure strategy profiles

This section states a preliminary result. Theorem 1 shows that given any strategy profile σ, there exists a degenerate strategy profile m such that (i) each player i is assigned a pure strategy k in the support of σ_i, and (ii) the number of players who play each pure strategy k is 'close' to the expected number who would have played k given strategy profile σ. With this result in hand our main results can be easily proved. We note now that, when applying Theorem 1 in the proofs of Theorems 2 and 3, the strategy profile σ is not (necessarily) to be thought of as 'the strategy profile of the population' but more as the strategy profile restricted to those players who have the same crowding attribute.

Theorem 1: For any strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ there exists a degenerate strategy profile $m = (m_1, ..., m_n)$ such that:

$$\mathit{support}(m_i) \subset \mathit{support}(\sigma_i) \tag{2}$$

for all i and:

$$\left| \sum_{i=1}^{n} m_i(k) - \sum_{i=1}^{n} \sigma_i(k) \right| \leq 1 \tag{3}$$

for all $k \in S$.

Observe that if σ, in Theorem 1, were a Nash equilibrium, then Theorem 1 states that there is an approximating pure strategy profile m where *every* player plays a pure strategy in his best response set for σ. This is crucial in proving our two main theorems in that allows us to 'aggregate' the strategies of players who have the same crowding attribute yet potentially different taste attributes.

We highlight the relationship between Theorem 1 and the related but distinct Shapley-Folkman Theorem and note that the Shapley-Folkman Theorem will not suffice for our purposes.[12] For the reader's convenience we state the Shapley-Folkman Theorem:

Shapley-Folkman Theorem :[13] If $A_1, ..., A_J$ is a collection of sets in $\mathbb{R}^m$, $J > m$, then for any $x \in con(\sum_j A_j)$ there exists a representation of x of the form: $x = \sum_{J_1} y_j + \sum_{J_2} z_j$, where for each $j \in J_1$, $y_j \in A_j$ and for each $j \in J_2$, $z_j \in con(A_j)$, $|J_1| + |J_2| = J$ and $|J_2| \leq m$.

If we let K denote the number of strategies then from the Shapley-Folkman Theorem we obtain: for any strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ there exists a degenerate strategy profile $m = (m_1, ..., m_n)$ such that: $support(m_i) \subset support(\sigma_i)$ for all i and:

$$\max_{k \in S} \left| \sum_{i=1}^{n} m_i(k) - \sum_{i=1}^{n} \sigma_i(k) \right| < K \tag{4}$$

In Theorem 1 we obtain a bound that is independent of the number of strategies K. This is clearly crucial in treating a non-finite set of strategies, which is permitted by our model. We leave the full relationship between the Shapley-Folkman Theorem and Theorem 1 as an open question.

[12] As discussed in Section 4 Rashid (1983) does make use of the Shapley-Folkman Theorem in proving a special case of our Theorem 2.

[13] See, for example, Green and Heller (1991) for a proof of the Shapley-Folkman Theorem.

3.2 Continuity in crowding attributes

To derive our purification results we make use of a natural and mild continuity assumption on crowding attributes, introduced in Wooders, Cartwright and Selten (2001), that will be assumed throughout. Given the strategy choices of other players, it is assumed that each player is nearly indifferent to a minor perturbation of the crowding attributes of other players (provided his own crowding attribute is unchanged). Formally:

Continuity in crowding attributes: We say that a pregame $\mathcal{G}$ satisfies continuity in crowding attributes if: for any $\varepsilon > 0$, any two populations (N, α) and $(N, \overline{\alpha})$ and any strategy profile $\sigma \in \Sigma^N$ if:

$$\max_{j \in N} dist(\kappa(j), \overline{\kappa}(j)) < \varepsilon$$

then for any $i \in N$ where $\alpha(i) = \overline{\alpha}(i)$:

$$\left| U_i^{\alpha}(\sigma_i, \sigma_{-i} | t^z) - U_i^{\overline{\alpha}}(\sigma_i, \sigma_{-i} | t^z) \right| < \varepsilon$$

all $t^z \in \mathcal{T}_i$. Where '*dist*' is the metric on the space of crowding attributes $\mathcal{C}$.

Note that in the definition of continuity in crowding attributes the strategy profile is held constant. Thus, the attributes of players may change but their strategies do not.

3.3 Large game property

To define the large game property, some additional notation and definitions are required. Denote by EW the set of functions mapping $\mathcal{C} \times \mathcal{A} \times \mathcal{T}$ into $\mathbb{R}_+$, the set of non-negative reals. We refer to $ew \in EW$ as an *expected* *weight function.* Given a population (N, α) we say that an expected weight function $ew_{\alpha,\sigma}$ is relative to strategy profile σ if and only if:

$$ew_{\alpha,\sigma}(c, a^l, t^z) = \sum_{a \in A} \sum_{t \in T} w_{\alpha,a,t}(c, a^l, t^z) \Pr(a, t)$$

for all ω, a^l and t^z. Thus, $ew_{\alpha,\sigma}(\omega, a^l, t^z)$ denotes the *expected* number of players of crowding-attribute c who will have type t^z and play action a^l. Note that this expectation is taken before any player is aware of his type.

Fix a population (N, α). Let EW_α denote the set of expected weight functions that may be realized given population (N, α). We define a metric on the space EW_α:

$$dist1(ew, eg) = \frac{1}{|N|} \sum_{a^l \in \mathcal{A}} \sum_{t^z \in \mathcal{T}} \sum_{c \in \mathcal{C}} \left| ew(c, a^l, t^z) - eg(c, a^l, t^z) \right|$$

for any $ew, eg \in EW_\alpha$. Thus, two expected weight functions are 'close' if the expected proportion of players with each crowding attribute and each type playing each action are close. We can now state our main assumption:

Large game property: We say that a pregame $\mathcal{G}$ satisfies the large game property if: for any $\varepsilon > 0$, any population (N, α) and any two strategy profiles $\sigma, \overline{\sigma} \in \Sigma^N$ with expected weight functions $ew_{\alpha,\sigma}, ew_{\alpha,\overline{\sigma}}$ satisfying:

$$dist1(ew_{\alpha,\sigma}, ew_{\alpha,\overline{\sigma}}) < \varepsilon$$

if $\sigma_i = \overline{\sigma}_i$ then:

$$|U_i^\alpha(\sigma_i, \sigma_{-i}|t^z) - U_i^\alpha(\overline{\sigma}_i, \overline{\sigma}_{-i}|t^z)| < \varepsilon$$

for all $t^z \in \mathcal{T}_i$.

If a pregame satisfies the large game property then we can think of games induced from the pregame as satisfying two conditions on payoff functions:

1. A player is nearly indifferent to a change in the proportion of players of each attribute playing each pure strategy (provided his own strategy is unchanged); thus, any one individual has near-negligible influence over the payoffs of other players.

2. A player is 'risk neutral' in the sense that the expected weight function largely determines his payoff; thus two strategy profiles that induce the same expected weight function give a similar payoff.

The first condition is reflective of the attribute of game under consideration and is crucial to obtaining our main result; Example 1, for instance, does not satisfy the large game property in this respect. The second condition is relatively mild given that we consider games with many players; it follows, for example, from the law of large numbers that in the case of a finite strategy set, with high probability, in a game with many players the realized weight function will be close to the expected weight function (Kalai 2002).[14]

Note that the large game property relates to changes in the strategies of players while their attributes do not change; this contrasts with the assumption of continuity in crowding attributes that relates to changes in attributes while strategies do not change. As a consequence a pregame may satisfy the large game property and yet there not be continuity in attributes and vice-versa.

3.4 Main Result; Approximate purification

Our main result demonstrates that in sufficiently large games with many players any Nash equilibrium can be approximately purified.

[14] Thus, it is not so much that players are risk neutral but rather that there is little risk.

Theorem 2: Consider a pregame $\mathcal{G} = (\Omega, \mathcal{A}, \mathcal{T}, b, h)$ satisfying continuity in crowding attributes and the large game property. Given any real number $\varepsilon > 0$ there is an integer $\eta(\varepsilon)$ with the property that, for any induced game $\Gamma(N, \alpha)$ where $|N| > \eta(\varepsilon)$ and for any Nash equilibrium σ of game $\Gamma(N, \alpha)$, there exists a Bayesian Nash ε-equilibrium in pure strategies m that is an ε-purification of σ.

Proof: Suppose not. Then there is some $\varepsilon > 0$ such that for each integer ν there is an induced game $\Gamma(N^\nu, \alpha^\nu)$ with $|N^\nu| > \nu$ and a Nash equilibrium σ^v with the property that there exists no Nash ε-equilibrium in pure strategies providing an ε-purification of σ. Given that σ^ν is a Nash equilibrium, for any $i \in N^\nu$ and for any strategy m_i^ν where $support(m_i^\nu)$ $\subset support(\sigma_i^\nu)$ we have:

$$U_i^{\alpha^\nu}(m_i^\nu, \sigma_{-i}^\nu | t^z) \geq U_i^{\alpha^\nu}(s, \sigma_{-i}^\nu | t^z) \tag{5}$$

for all $t^z \in \mathcal{T}_i$ and $s \in \Sigma$.

Use compactness of $\mathcal{C}$ to write $\mathcal{C}$ as the disjoint union of a finite number of non-empty subsets $\mathcal{C}_1, ..., \mathcal{C}_A$, each of diameter less than $\frac{1}{6}\varepsilon$. For each $a = 1, ..., A$, choose and fix a point $c_a \in \mathcal{C}_a$. For each ν, without changing taste attributes of players, we define the crowding attribute function $\overline{\kappa}^\nu$ by its coordinates $\overline{\kappa}^\nu(\cdot)$ as follows:

$$\text{for each } j \in N,\ \overline{\kappa}^\nu(j) = c_a \text{ if and only if } \kappa(j) \in \mathcal{C}_a.$$

Define new attribute functions $\overline{\alpha}^\nu$ by $\overline{\alpha}^\nu(j) = (\pi(j), \overline{\kappa}^\nu(j))$ when $\alpha^\nu(j) = (\pi(j), \kappa^\nu(j))$ for each $j \in N^\nu$. By applying Theorem 1 to each $c \in \overline{\kappa}^\nu(N)$, i.e. $c_1, ..., c_A$ it follows that there exists a sequence $\{m^\nu\}$ of degenerate strategy profiles such that:

1. for all $c \in \mathcal{C}$, $a^l \in \mathcal{A}$ and $t^z \in \mathcal{T}$

$$\lim_{\nu\to\infty} \frac{eg^{\nu}_{\overline{\alpha}^{\nu},m^{\nu}}(c,a^l,t^z)}{|N^{\nu}|} = \lim_{\nu\to\infty} \frac{ew^{\nu}_{\overline{\alpha}^{\nu},\sigma^{\nu}}(c,a^l,t^z)}{|N^{\nu}|}, \text{ and} \tag{6}$$

2. for all ν and $i \in N^{\nu}$,

$$support(m^{\nu}_i) \subset support(\sigma^{\nu}_i). \tag{7}$$

Pick an arbitrary ν and player $i \in N^{\nu}$. Consider the attribute function $\overline{\overline{\alpha}}^{\nu}$ where $\overline{\overline{\alpha}}^{\nu}(i) = \alpha^{\nu}(i)$ and $\overline{\overline{\alpha}}^{\nu}(j) = \overline{\alpha}^{\nu}(j)$ for all $j \notin i$. By continuity in crowding attributes:

$$\left| U_i^{\alpha^{\nu}}(s,\sigma^{\nu}_{-i}|t^z) - U_i^{\overline{\overline{\alpha}}^{\nu}}(s,\sigma^{\nu}_{-i}|t^z) \right| < \frac{\varepsilon}{6}$$

for all $t^z \in \mathcal{T}_i$ and $s \in \Sigma$, and:

$$\left| U_i^{\alpha^{\nu}}(s,m^{\nu}_{-i}|t^z) - U_i^{\overline{\overline{\alpha}}^{\nu}}(s,m^{\nu}_{-i}|t^z) \right| < \frac{\varepsilon}{6}$$

for any $t^z \in \mathcal{T}_i$ and $s \in \Sigma$. In view of (6) and the large game property it is clear if ν was sufficiently large:

$$\left| U_i^{\overline{\overline{\alpha}}^{\nu}}(s,\sigma^{\nu}_{-i}|t^z) - U_i^{\overline{\overline{\alpha}}^{\nu}}(s,m^{\nu}_{-i}|t^z) \right| < \frac{\varepsilon}{6}$$

for any $t^z \in \mathcal{T}_i$ and $s \in \Sigma$. Thus, for ν sufficiently large and for any $i \in N^{\nu}$:

$$\left| U_i^{\alpha^{\nu}}(s,m^{\nu}_{-i}|t^z) - U_i^{\alpha^{\nu}}(s,\sigma^{\nu}_{-i}|t^z) \right| < \frac{\varepsilon}{2}$$

for any $t^z \in \mathcal{T}_i$ and $s \in \Sigma$. Finally, given (7) and (5) for ν sufficiently large:

$$U_i^{\alpha^{\nu}}(m^{\nu}_i, m^{\nu}_{-i}|t^z) - U_i^{\alpha^{\nu}}(s,m^{\nu}_{-i}|t^z) > -\varepsilon$$

for all $t^z \in \mathcal{T}_i$ and $s \in \Sigma$. This gives the desired contradiction.■

3.5 Many players of each crowding attribute

Our third result demonstrates that, when for each player in an induced game there are sufficiently many players with similar attributes, weaker conditions are sufficient to approximately purify Nash equilibrium.

Fix a population (N, α). As before let EW_α denote the set of expected weight functions that may be realized given population (N, α). For each $c \in \mathcal{C}$ with $c \in \kappa(N)$, let $\rho(c) \equiv \left|\kappa^{-1}(c)\right|$ be the number of players with crowding attribute c. We define a second metric on the space EW_α:

$$dist2(ew, eg) = \sum_{c \in \kappa(N)} \left(\frac{1}{\rho(c)} \sum_{a^l \in \mathcal{A}} \sum_{t^z \in \mathcal{T}} \left| ew(c, a^l, t^z) - eg(c, a^l, t^z) \right| \right)$$

for any $ew, eg \in EW_\alpha$. Thus, two expected weight functions are 'close' if the expected proportions of players with each crowding attribute playing each pure strategy are close. This differs significantly from the earlier $dist1$ where closeness is judged on the proportions *relative to the total population* playing each pure strategy. It is immediate that $dist2(ew, eg) \geq dist1(ew, eg)$. We state a second assumption that weakens the large game property:

Continuity property: We say that a pregame $\mathcal{G}$ satisfies the continuity property if: for any $\varepsilon > 0$, any population (N, α) and any two strategy profiles $\sigma, \overline{\sigma} \in \Sigma^N$ where:

$$dist2(ew_{\alpha,\sigma}, ew_{\alpha,\overline{\sigma}}) < \varepsilon$$

if $\sigma_i = \overline{\sigma}_i$ then:

$$|U_i^\alpha(\sigma_i, \sigma_{-i}|t^z) - U_i^\alpha(\overline{\sigma}_i, \overline{\sigma}_{-i}|t^z)| < \varepsilon$$

for all $t^z \in \mathcal{T}_i$.

The continuity property appears mild. In particular, one player can have a large influence even in large populations if he is the only player with his crowding attribute. Thus, for example, the pregame of Example 1 satisfies the continuity property but not the large game property. This illustrates that the continuity property is not sufficient to obtain a result such as Theorem 2.

Let $B_\tau(c)$ denote a ball in crowding attribute space $\mathcal{C}$ centered on c of radius τ. We denote by $F(\eta,\tau)$ the set of populations where $(N,\alpha) \in F(\eta,\tau)$ if and only if

$$\sum_{c' \in B_\tau(c)} \rho(c') > \eta$$

for all $c \in \alpha(N)$. Thus, population $(N,\alpha) \in F(\eta,\tau)$ only if there is a certain 'thickness' to the distribution of players over crowding attributes. Note, however, that a population $(N,\alpha) \in F(\eta,\tau)$ may have the property that there is a 'large' subset Ω' of attribute space and no player $i \in N$ with attributes in Ω'.

We obtain the following result:

Theorem 3: Consider a pregame $\mathcal{G} = (\Omega, \mathcal{A}, \mathcal{T}, b, h)$ that satisfies continuity in crowding attributes and the continuity property. Given any real number $\varepsilon > 0$ there is an integer $\eta(\varepsilon)$ and a real number $\tau(\varepsilon) > 0$ such that for any population $(N,\alpha) \in F(\eta(\varepsilon),\tau(\varepsilon))$ and any Nash equilibrium σ of the induced game $\Gamma(N,\alpha)$ there exists a Bayesian Nash ε-equilibrium in pure strategies m that is an ε-purification of σ.

Proof: Suppose not. Then there is some $\varepsilon > 0$ such that for each integer ν there is a population $(N^\nu,\alpha^\nu) \in F\left(\nu, \frac{1}{36}\varepsilon\right)$ and a Nash equilibrium σ^v of the induced game $\Gamma(N^\nu,\alpha^\nu)$ with the property that there exists no Nash

ε-equilibrium in pure strategies that is an ε-purification of σ.

To simplify notation for any set $A \subset \mathcal{C}$ let $\rho^\nu(A) \equiv \left|\kappa^{\nu^{-1}}(A)\right|$ be the number of players in population (N^ν, α^ν) with crowding attribute $c \in A$.

We conjecture (*) that, by passing to a subsequence if necessary, there exists a partition of $\mathcal{C}$ into a finite number of non-empty subsets $\mathcal{C}_1, ..., \mathcal{C}_Q$, each of diameter less than $\frac{\varepsilon}{6}$ where the sequence $\{\rho^\nu(\mathcal{C}_q)\}_{\nu>0}$ either tends to infinity or converges to zero. Assume that conjecture (*) is correct. Given the continuity property it is simple to see that a contradiction can be obtained in a similar manner to the contradiction in the proof of Theorem 2. It thus remains to prove conjecture (*).

Define $\tau \equiv \frac{1}{36}\varepsilon$. Use compactness of $\mathcal{C}$ to write $\mathcal{C}$ as the disjoint union of a finite number of non-empty subsets $\mathcal{C}_1, ..., \mathcal{C}_R$, each of diameter less than τ. This initial partition is unlikely to satisfy the desired properties; the desired partition will be 'formed' by merging subsets together. By passing to a sub-sequence if necessary, we can assume, for each $\mathcal{C}_r$, that the sequence $Y_r \equiv \{\rho^\nu(\mathcal{C}_r)\}_{\nu=1}^{\infty}$ either tends to infinity or converges to a finite limit. Define subsets A^∞ and A^+ of $\{\mathcal{C}_1, ..., \mathcal{C}_R\}$ by $\mathcal{C}_r \in A^\infty$ if and only if Y_r tends to infinity and $\mathcal{C}_r \in A^+$ if and only if Y_r tends to a positive real number. (Note that A^∞ and A^+ do not necessarily comprise a partition of $\{\mathcal{C}_1, ..., \mathcal{C}_R\}$ since Y_r may be zero for some $\mathcal{C}_r$.)

Consider any $\mathcal{C}_r \in A^+$. There must exist a real number ν_r such that for any population (N^ν, α^ν) where $\nu > \nu_r$ there is at least one player $i^\nu \in N^\nu$ with $\kappa^\nu(i^\nu) \in \mathcal{C}_r$. Let $c^\nu \equiv \alpha^\nu(i^\nu)$ for all ν. By assumption:

$$\sum_{c' \in B_\tau(c^\nu)} \rho^\nu(c') > \nu$$

for all ν. Fix an arbitrary point $c_r \in \mathcal{C}_r$. Given that the diameter of $\mathcal{C}_r$ is τ

it holds that:

$$\sum_{c' \in B_{2\tau}(c_r)} \rho^{\nu}(c') > \nu$$

for all ν. Partitioning $\mathcal{C}$ into sets $\mathcal{C}_1, ..., \mathcal{C}_R$ also partitions the ball $B_{2\tau}(c_r)$ into a finite number of sets $B^1_{2\tau}(c_r), .., B^R_{2\tau}(c_r)$ where:

$$c \in B^a_{2\tau}(c_r) \text{ if and only if } c \in B^a_{2\tau}(c_r) \text{ and } c \in \mathcal{C}_a.$$

This implies that there must exist some $B^a_{2\tau}(c_r) \in A^{\infty}$. Furthermore, $dist(c_r, c_a) < 3\tau$ for all $c_a \in \mathcal{C}_a$ and all $c_r \in \mathcal{C}_r$.

From the above, it follows, that by an appropriate merging of the subsets $\mathcal{C}_1, ..., \mathcal{C}_R$ (and, in particular, merging a set $\mathcal{C}_r \in A^+$ with a set $\mathcal{C}_a \in A^{\infty}$) there must exist a partition of $\mathcal{C}$ into a finite number of non-empty subsets $\mathcal{C}_1, ..., \mathcal{C}_Q$, each of diameter less than $6\tau = \frac{\varepsilon}{6}$ where the sequence $\{\rho^{\nu}(\mathcal{C}_r)\}_{\nu>0}$ either tends to infinity or converges to zero. This proves conjecture (*) and thus Theorem 3.■

3.6 A remark on existence of equilibrium

With a countable set of strategies, a Nash equilibrium, even one in mixed strategies, may not exist. This is easy to see. Suppose, for example, the game is one where the prize goes to the player who announces the highest integer. If we add the requirement of compactness of the sets of actions and of types, however, then existence of a Bayesian-Nash equilibrium in mixed strategies can be obtained using, for example, the fixed point theorem of Glicksberg (1952).

4 Some further relationships to the literature

Two authors that provide results on purification with large but finite player sets are Rashid (1983) and Kalai (2002). Kalai (2002) provides sufficient conditions for the existence of an approximate Bayesian ex-post Nash equilibrium. One implication of Kalai's results is that every Nash equilibrium can be approximately purified.[15] In contrast to this paper and Wooders, Cartwright and Selten (2001), Kalai requires both a finite number of actions and a finite number of crowding types. We conjecture, but have not demonstrated, that Kalai's sort of purification result will hold in the context of our paper.

With a finite set of strategies and finite types of players, Rashid (1983) makes use of the Shapley-Folkman Theorem to prove his result on existence of approximate equilibrium in pure strategies. By assuming a linearity of payoff functions Rashid demonstrates that 'near' to any Nash equilibrium there is an approximate Nash equilibrium in which $|N| - K$ players use pure strategies (where K is the number of strategies) and K players may play mixed strategies. (See also Carmona 2003 who argues that an additional condition, equicontinuity of payoff functions for example, is required).

Many authors have contributed to the literature on the existence of a pure strategy non-cooperative equilibria in games with a continuum of players (including Schmeidler 1973, Mas-Colell 1984, Khan 1989, 1998, Khan et al. 1997, Pascoa 1993a, 1998 and Khan and Sun 1999). This literature, given various assumptions on the strategy space, has demonstrated the exis-

[15] Indeed, Kalai demonstrates that not only can a Nash equilibrium be purified but when a Nash equilibrium is played almost *any* realized set of strategy profiles must be an approximate Nash equilibrium – that is, with probability arbitrarily close to one, every Nash equilibrium self purifies. See also footnote 4.

tence of a non-cooperative equilibrium when payoffs depend on opponent's strategies through the distribution over pure strategies. Our Theorem 2 can be seen as providing a finite analogue to some of these continuum results.[16]

Within the literature on non-atomic games, the approach of Pascoa (1993a) appears most similar to our own. Pascoa (1993a) deals with non-anonymous games as introduced by Green (1984). A player in a non-anonymous game has a type (which could be thought as an attribute in our framework) and a player's payoff depends on his opponent's strategies through the distribution over types and pure strategies. More formally, let T denote a set of types and D the set of Borel probability measures over $T \times S$.[17] The payoff to a player of type t from playing strategy s when the strategies of opponents is $\mu \in D$ is given by $v(t, s, \mu)$. To obtain his results Pascoa assumes that $v(t, \cdot, \cdot)$ is jointly continuous, with respect to the weak* topology on D.[18] This corresponds to our assumption of a pregame that satisfies the large game property and continuity in crowding attributes. Pascoa (1993a,1998) also obtains existence results using conditions similar to those of our Theorem 3.

5 Conclusions

This paper introduces a framework for studying asymptotic properties of strategic games with growing numbers of players. Our framework extends those already in the literature. The major innovations are (a) our math-

[16] Note that this literature is typically concerned with the existence of a non-cooperative equilibrium and not (as in this paper) the purification of a non-cooperative equilibrium that is assumed to exist (exceptions include Pascoa 1998).

[17] Where S denotes as previously the set of strategies.

[18] Pascoa (1993a) assumes a compact metric space of strategies.

ematical result (Theorem 1), (b) allowing countable sets of actions and types, and (c) the formalization of the separation of crowding and taste attributes of players. This separation plays a role in other research on noncooperative games, particularly on games with many players where similar players conform (see Wooders, Cartwright and Selten 2001 and Cartwright and Wooders 2003). To relate this separation to other lines of research, in models of private goods economies where the tastes of an individual affect other individuals only through his demand for private goods, a separation of tastes from other attributes of a player, in particular, endowment, is implicit. In the literature of local public goods economies and economies with clubs, where the utility of an individual depends on the attributes of other individuals in the same clubs, a distinction similar to that of this paper is made.[19] While such a distinction may be implicit in numerous examples and could also have been built into some of the prior literature, except for our research, we are unaware of any formalization and use of this distinction in the prior literature of noncooperative game theory. In research in progress on noncooperative games, but following Conley and Wooders (1996, 2001) research on cooperative and price taking equilibrium, we endogenise choice of crowding attributes and consider evolution of observed patterns of crowding attributes.

6 Appendix

We introduce some additional notation. Let $a = (a_1, ..., a_n),\ b = (b_1, ..., b_n) \in \mathbb{R}^n$. We write $a \geq b$ if and only if $a_i \geq b_i$ for all $i = 1, ..., n$. Given any strat-

[19] We refer the reader to Wooders, Cartwright and Selten (2003) and Conley and Wooders (2001) and references there for further motivation and discussion of crowding types.

egy profile σ let $M(\sigma)$ denote the set of strategy profiles such that $m \in M(\sigma)$ if and only if (1) m is degenerate and (2) $support(m_i) \subseteq support(\sigma_i)$ for all $i \in N$. It is immediate that $M(\sigma)$ is non-empty for any σ.

Lemma 1: Let $N = \{1, ..., n\}$ be a finite set. For any strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ and for any function $\overline{g} : S \to \mathbb{Z}_+$ such that $\sum_i \sigma_i \geq \overline{g}$, there exists $m \in M(\sigma)$ such that

$$\sum_i m_i \geq \overline{g}.$$

Proof: Suppose the statement of the lemma is false. Then there exists a strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ and a function $\overline{g}$ where $\sum_{i \in N} \sigma_i \geq \overline{g}$, such that, for any vector $m = (m_1, ..., m_n) \in M(\sigma)$ there must exist at least one $\widehat{k}$ where $\widehat{k} \in S$ and $\sum_i m_i(\widehat{k}) < \overline{g}(\widehat{k})$. For each vector $m \in M(\sigma)$ let L be defined as follows:

$$L(m) = \sum_{k \in S : \sum_i m_i(k) < \overline{g}_k} \left(\overline{g}(k) - \sum_i m_i(k) \right)$$

We note that $L(m)$ must be finite and positive for all m.[20] Select $m^0 \in M(\sigma)$ for which $L(m)$ attains its minimum value over all $m \in M(\sigma)$. Intuitively the vector m^0 is 'as close' as we can get to satisfying the lemma. We remark that the method of proof will be one of 'shuffling' the pure strategies that players use so as to demonstrate the existence of a strategy profile m^* where $L(m^*) = L(m^0) - 1$. Providing the desired contradiction.

Pick a strategy $\widehat{k}$ such that $\overline{g}(\widehat{k}) - \sum_i m_i^0(\widehat{k}) > 0$. For any subset I of N let the set $S(I) \subset S$ be such that:

$$S(I) = \left\{ \widehat{k} \right\} \cup \left\{ k \in S : m_i^0(k) = 1 \text{ for some } i \in I \right\}$$

[20] Note that the set of k such that $\sum_i m_i(k) < \overline{g}_k$ need not be finite. Given, however, that $\sum_k \sum_i \sigma_i(k) = |N|$ it must be that $\sum_k \overline{g}(k) \leq |N|$ and thus $L(m)$ is finite.

We can now define sets N^t for $t = 0, 1, ...$ as follows:

$$
\begin{aligned}
N^0 &= \left\{ i \in N : m_i^0(\widehat{k}) = 1 \right\} \text{ and for all } t > 0 \\
N^t &= N^{t-1} \cup \left\{ \begin{array}{c} j \in N : \sigma_j(k) > 0 \text{ and } m_j^0(k) = 0 \\ \text{for some } k \in S\left(N^{t-1}\right) \end{array} \right\}
\end{aligned}
$$

Ultimately, for some $t^* \geq 1$ we must have that $N^{t^*+1} = N^{t^*} \equiv \overline{N}$. This is an immediate consequence of the finiteness of the player set. Let $S(\overline{N}) \equiv \overline{S}$.

Consider any pure strategy $k^* \in \overline{S}$. The construction of $\overline{N}$ and $\overline{S}$ imply that there must exist a chain of players $\{i_1, ..., i_{\overline{t}}\} \subset \overline{N}$ where (1) $m_{i_t}^0(k_t) = 1$ for $t = 1, ..., \overline{t} - 1$, (2) $m_{i_{\overline{t}}}(k^*) = 1$, (3) $\sigma_{i_t}(k_{t-1}) > 0$ for $t = 2, ..., \overline{t}$ and (4) $\sigma_{i_1}(\widehat{k}) > 0$. Thus, there exists a vector $m^* \in M(\sigma)$ such that:

$$
\begin{aligned}
& m_{i_1}^*(k_1) = 0 \text{ and } m_{i_1}^*(\widehat{k}) = 1, \\
& m_{i_{\overline{t}}}^*(k^*) = 0 \text{ and } m_{i_{\overline{t}}}^*(k_{\overline{t}-1}) = 1 \\
& m_{i_t}^*(k_t) = 0 \text{ and } m_{i_t}^*(k_{t-1}) = 1, \text{ for all } t = 2, ..., \overline{t} - 1, \text{and} \\
& m_i^*(k) = m_i^0(k) \text{ for all other } i \text{ and } k.
\end{aligned}
$$

Suppose that:

$$\sum_{i \in N} m_i^0(k^*) > \overline{g}(k^*).$$

This implies that:

$$\sum_{i \in N} m_i^0(k^*) \geq \overline{g}(k^*) + 1$$

and thus $L(m^*) = L(m^0) - 1$.

To avoid a contradiction we need:

$$\sum_{i \in N} m_i^0(k) \leq \overline{g}(k). \tag{8}$$

for all $k \in \overline{S}$. Using the definition of $\overline{S}$ there can exist no player $j \in N \backslash \overline{N}$ such that $\sigma_j(k) > 0$ for some $k \in \overline{S}$ unless $m_j^0(k) = 1$. This implies that:

$$\sum_{i \in N \backslash \overline{N}} m_i^0(k) \geq \sum_{i \in N \backslash \overline{N}} \sigma_i(k) \tag{9}$$

for all $k \in \overline{S}$. Using the definition of $\overline{S}$ we have that:

$$\sum_{k\in\overline{S}}\sum_{i\in\overline{N}} m_i^0(k) \geq \sum_{k\in\overline{S}}\sum_{i\in\overline{N}} \sigma_i(k). \tag{10}$$

Combining (9) and (10) and using the statement of the lemma, we see that:

$$\sum_{k\in\overline{S}}\sum_{i\in N} m_i^0(k) \geq \sum_{k\in\overline{S}}\sum_{i\in N} \sigma_i(k) \geq \sum_{k\in\overline{S}} \overline{g}(k)$$

However, by assumption:

$$\overline{g}(\widehat{k}) > \sum_{i\in N} m_i^0(\widehat{k})$$

and also by assumption, $\widehat{k} \in \overline{S}$. Thus, there must exist at least one $k \in \overline{S}$ such that:

$$\overline{g}(k) < \sum_{i\in N} m_i^0(k).$$

This contradicts (8) and completes the proof.■

We introduce some additional 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 $\lceil 9 \rceil = 10$).

Theorem 1: For any strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ there exists a a degenerate strategy profile $m = (m_1, ..., m_n)$ such that:

$$support(m_i) \subset support(\sigma_i) \tag{11}$$

for all i and:

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

for all $k \in S$.

Proof: Denote by $M^*(\sigma)$ the set of vectors $m = (m_1, ..., m_n) \in M(\sigma)$ such that $\sum_i m_i(k) \geq \lfloor \sum_i \sigma_i(k) \rfloor$ for all k. By Lemma 1 this set is non-empty. Proving the Lemma thus amounts to showing that there exists a vector $m \in M^*(\sigma)$ such that $\lceil \sum_i \sigma_i(k) \rceil \geq \sum_i m_i(k)$ for all $s_k \in S$. Suppose not. Then, for every vector $m \in M^*(\sigma)$ there exists some strategy $k \in S$ such that $\sum_i m_i(k) > \lceil \sum_i \sigma_i(k) \rceil$. For any strategy profile $m \in M^*(\sigma)$ define $L(m)$ by:

$$L(m) \equiv \sum_{k: \sum_i m_i(k) > \lceil \sum_i \sigma_i(k) \rceil} \left(\sum_{i=1}^n m_i(k) - \left\lceil \sum_{i=1}^n \sigma_i(k) \right\rceil \right).$$

We note that $L(m)$ is always positive and finite. Pick strategy profile $m^0 \in M^*(\sigma)$ where the value of $L(m)$ is minimized. We note that m^0 comes as close as any profile to satisfying the statement of the Lemma.

Denote by $\widehat{k}$ a pure strategy such that:

$$\sum_{i=1}^n m_i^0(\widehat{k}) > \left\lceil \sum_{i=1}^n \sigma_i(\widehat{k}) \right\rceil.$$

We introduce sets S^t and N^t, $t = 0, 1, 2, ...$, where:

$$N^0 = \{i : m_i^0(\widehat{k}) = 1\} \text{ and for } t > 0$$

and for $t > 0$,

$$\begin{aligned} S^t &= \{k : \sigma_i(k) > 0 \text{ for some } i \in N^{t-1}\} \\ N^t &= \{i : m_i^0(k) = 1 \text{ for some } k \in S^t\}. \end{aligned}$$

For some t^*, $N^{t^*} = N^{t^*+1} \equiv \overline{N}$ and $S^{t^*} = S^{t^*+1} \equiv \overline{S}$. The construction of S^t and N^t imply that for any $k^* \in \overline{S}$ there must exist a set of players

$\{i_0, i_1, ..., i_{\overline{t}}\} \in \overline{N}$ such that:

$$\begin{aligned} m^0_{i_0}(\widehat{k}) &= 1 \text{ and } \sigma_{i_0}(k_1) > 0, \\ m^0_{i_r}(k_r) &= 1 \text{ and } \sigma_{i_r}(k_{r+1}) > 0 \text{ for all } r = 1, .., \overline{t} - 1, \\ m^0_{i_{\overline{t}}}(k_{\overline{t}}) &= 1 \text{ and } \sigma_{i_t}(k^*) > 0, \end{aligned}$$

Suppose there exists $k^* \in \overline{S}$ such that:

$$\sum_{i=1}^{n} m^0_i(k^*) \leq \sum_{i=1}^{n} \sigma_i(k^*).$$

Given the chain of players $\{i_0, i_1, ..., i_{\overline{t}}\} \in \overline{N}$ as introduced above, consider the vector m^* constructed as follows:

$$\begin{aligned} m^*_{i_0}(\widehat{k}) &= 0 \text{ and } m^*_{i_0}(k_1) = 1, \\ m^*_{i_r}(k_r) &= 0 \text{ and } m^*_{i_r}(k_{r+1}) = 1 \text{ for all } r = 1, ..., \overline{t} - 1, \\ m^*_{i_{\overline{t}}}(k_{\overline{t}}) &= 0 \text{ and } m^*_{i_{\overline{t}}}(k^*) = 1, \\ m^*_i(k) &= m^0_i(k) \text{ for all other } k \in S \text{ and } i \in N. \end{aligned}$$

It is easily checked that the vector $m^* \in M(\sigma)$ leads to the desired contradiction given that $L(M^*) = L(m^0) - 1$. We note, however, that:

$$\sum_{i=1}^{n} \sum_{k \in \overline{S}} m^0_i(k) = |\overline{N}| = \sum_{i \in \overline{N}} \sum_{k \in \overline{S}} \sigma_i(k).$$

Thus, if:

$$\sum_{i=1}^{n} m^0_i(\widehat{k}) > \sum_{i=1}^{n} \sigma_i(\widehat{k}) \geq \sum_{i \in \overline{N}} \sigma_i(\widehat{k})$$

there must exist some $k^* \in \overline{S}$ such that:

$$\sum_{i=1}^{n} m_i(k^*) \leq \sum_{i \in \overline{N}} \sigma_i(k^*) \leq \sum_{i=1}^{n} \sigma_i(k^*)$$

giving the desired contradiction.■

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] Cartwright, E. and M. Wooders (2002) "Social conformity in arbitrary games with incomplete information," University of Warwick Working Paper no. 672.

[3] 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.

[4] Conley, J. and M.H. Wooders (1997) "Equivalence of the core and competitive equilibrium in a Tiebout economy with crowding types," *Journal of Urban Economics* 41 (1997) 421-440.

[5] Conley, J. and M.H. Wooders (1996) "Taste homogeneity of optimal jurisdictions in a Tiebout economy with crowding types and endogenous educational investment choices," *Richerche Economiche* 50, 367-387. On-line at http://www2.warwick.ac.uk/fac/soc/economics/staff/faculty/wooders/.

[6] Cripps, M.W, G. Keller, and S. Rady (2002) "Strategic experimentation: The case of poisson bandits," May 2002 CES Working Paper #737.

[7] Glicksberg, I. (1952) "A further generalization of the Kakutani Fixed Point Theorem, with application to Nash equilibrium points," *Proceedings of the American Mathematical Society*, 3: 170-174,

[8] Green, E.J. (1984) "Continuum and finite-player noncooperative models of competition," *Econometrica* vol.52, no. 4: 975-993.

[9] Green, J. and W.P. Heller (1991) "Mathematical analysis and convexity with applications to economics," in *Handbook of Mathematical Economics,* Vol.1, K.J. Arrow and W. Intriligator, eds. North Holland: Amsterdam, New York, Oxford.

[10] Gross, R. (1996) *Psychology. The Science of Mind and Behaviour*, Hodder and Stoughton.

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

[12] Kalai, E. (2002) "Ex-post stability in large games," Northwestern University Discussion Paper 1351.

[13] Khan, A. (1989) "On Cournot-Nash equilibrium distributions for games with a nonmetrizable action space and upper semi continuous payoffs," *Transactions of the American Mathematical society* 293: 737-749.

[14] Khan, A. and Y. Sun (2002) "Noncooperative games with many players," *Handbook of Game Theory*, R. Auman and S. Hart, eds. North Holland.

[15] Khan, A. and Y. Sun (1999) "Non-cooperative games on hyperfinite Loeb spaces," *Journal of Mathematical Economics* 31, 455-492.

[16] 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.

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

[18] Offerman, T., J. Potters and J. Sonnemans, 2002, Imitation and belief learning in an oligopoly experiment, *Review of Economic Studies*, 69: 973-997.

[19] Pascoa, M. (1998) "Nash equilibrium and the law of large numbers," *International Journal of Game Theory* 27: 83-92.

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

[21] Pascoa, M. (1993b) "Noncooperative equilibrium and Chamberlinian monopolistic competition," *Journal of Economic Theory*, 69: 335-353.

[22] Radner, R and R.W. Rosenthal (1982) "Private information and pure-strategy equilibria," *Mathematics of Operations Research* 7: 401-409

[23] Rashid, S. (1983) "Equilibrium points of nonatomic games; Asymptotic results," *Economics Letters* 12: 7-10.

[24] Rath, K.P., Y. Sun, S. Yamashige (1995) "The nonexistence of symmetric equilibria in anonymous games with compact action spaces," *Journal of Mathematical Economics* 24: 331-346.

[25] Rosenthal, R.W (1973) "A class of games possessing pure-strategy Nash equilibria," *International Journal of Game Theory*, 2:65–67, 1973.

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

[27] Wooders, M., E. Cartwright and R. Selten (2001) "Social conformity and social conformity in games with many players," University of Warwick Department of Economics Working Paper 589 (revised as 636).