Categorization and correlation in a random-matching game

Azrieli, Yaron

Tel-Aviv University

25 October 2007

CATEGORIZATION AND CORRELATION IN A RANDOM-MATCHING GAME

YARON AZRIELI[†]

October 25, 2007

ABSTRACT. We consider a random-matching model in which every agent has a categorization (partition) of his potential opponents. In equilibrium, the strategy of each player is a best response to the distribution of actions of his opponents in each category of his categorization. We provide equivalence theorems between distributions generated by equilibrium profiles and correlated equilibria of the underlying game.

JEL classification: C72, D82.

Keywords: Random-matching game, Categorization, Correlated equilibrium.

[†] School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: azrieliy@post.tau.ac.il
I thank Philippe Jehiel and Ehud Lehrer for helpful references.

1. Introduction

Models in which agents are randomly matched to play a fixed game have been studied in game theory, economics and biology for many years. Excluding some exceptions[1], it is usually assumed that the action that each player chooses is independent of his matched opponent. In other words, one action is assigned to each player and a player's success is measured by the performance of his action against the distribution of actions in the population of his potential opponents.

It is quite common, however, that an agent possesses some information about his realized opponent. For instance, a prospective employer can see the skin color of a job candidate; and a seller of some product may know the gender of a potential buyer. The information may be completely payoff irrelevant, meaning that the underlying game is unchanged. But the very fact that such information is available allows agents to use opponent-contingent strategies.

To address this issue we consider a random-matching model where each agent is equipped with an exogeneously given *categorization*, which is a partition of his potential opponents. A player must use the same action against all members of each category in his categorization, but may use different actions against opponents from different categories. In other words, each player's strategy is a measurable (with respect to his categorization) function from the set of his potential opponents to his action set. This reflects a situation where opponents within each category are indistinguishable in the eyes of the categorizer.

To define an equilibrium[2] in such an environment consider a realized match of two agents. Each of the two agents only knows the category to which the other agent belongs. Thus, assuming that the strategies of all other agents are known, each agent can compute the expected payoff he'll obtain from each possible action. In equilibrium every agent plays a (measurable) strategy that maximizes this expected payoff in every possible encounter. A slightly different interpretation[3] of an equilibrium profile is that each player i has in his mind a 'representative agent' for each category in his categorization. The (possibly mixed) action of this prototype is the distribution of actions of the players in that category when they face i. In equilibrium, the strategy of i is a best response to this distribution in every category.

We restrict attention to the case where the underlying game is a two-person normal-form game. We consider, however, both cases of random-matching from two populations and from a single population. Our main results are equivalence theorems between the distributions generated by equilibrium profiles and correlated equilibria distributions of the underlying game. The proof that the distribution induced by an equilibrium profile must be a correlated equilibrium distribution is straightforward. The converse is a little more demanding.

[1]See Section 6 for references.

[2]Since there are many equilibrium concepts closely related to the one defined here, and to avoid confusion, we prefer not to give our solution a new name. We therefore refer to it simply as equilibrium.

[3]This interpretation is in the spirit of Jehiel's (2005) Analogy Based Expectation Equilibrium.

To illustrate our results consider the following version of "chicken".

	L	R
T	$4,4$	$1,5$
B	$5,1$	$0,0$

It is well known that the probability distribution which places a weight of 1/3 on each of the outcomes $(B, L), (T, L)$ and (T, R) is a correlated equilibrium. Now, assume that there are two populations $N_1 = \{1, 2, 3\}$ and $N_2 = \{1', 2', 3'\}$, and that an agent from each population is randomly selected to play the game (probability of 1/9 to each encounter).

Each player categorizes the set of his potential opponents as follows. The partition of player 1 is $C_1 = \{(1', 2'), (3')\}$. Similarly, $C_2 = \{(1'), (2', 3')\}$, $C_3 = \{(1', 3'), (2')\}$, $C_{1'} = \{(1, 2), (3)\}$, $C_{2'} = \{(1), (2, 3)\}$ and $C_{3'} = \{(1, 3), (2)\}$. The strategies of the players are given in the following matrix.

	$1'$	$2'$	$3'$
1	T, L	T, R	B, L
2	B, L	T, L	T, R
3	T, R	B, L	T, L

Notice that each agent's strategy is measurable with respect to his categorization. Moreover, the strategy of each agent maximizes (among all measurable strategies) his expected payoff. Therefore, this strategy profile is an equilibrium in the above specified environment.

Now, given that player $2'$ was chosen from N_2, the induced distribution on the set of action profiles places probability 1/3 on each of the outcomes $(T, R), (T, L)$ and (B, L). The same is true for any one of the 9 agents. In particular, this implies that the total distribution of action profiles induced by this equilibrium is the same as the above mentioned correlated equilibrium distribution.

In Theorem 1 below we prove that, for any correlated equilibrium with rational probabilities in any (finite) two-player normal-form game, there is a two-populations random-matching environment and an equilibrium in this environment which induces the given correlated equilibrium. In Theorem 2 we show that the same holds for symmetric correlated equilibria when ordered pairs are randomly drawn from a single population to play a symmetric game. Thus, our results can be seen as providing another interpretation for the correlated equilibrium concept.

The rest of the paper is organized as follows. Section 2 formally defines the random-matching environments and the equilibrium concept. In Section 3 we state our main theorems. The proofs of the theorems are in Sections 4 and 5. Related literature is surveyed in Section 6 while Section 7 concludes and suggests possible extensions.

2. The model

Let $G = (A_1, A_2, u_1, u_2)$ be a 2-player normal-form game, where A_i is player's i finite set of (pure) actions and $u_i : A_1 \times A_2 \to \mathbb{R}$ is player's i utility function, $i = 1, 2$. Throughout this paper a subscript $-i$ should be read $3 - i$. As usual, we denote $A = A_1 \times A_2$ and, with abuse of notation, $a = (a_i, a_{-i})$ denotes the element of A in which $a_i \in A_i$ and $a_{-i} \in A_{-i}$ (for $i = 1, 2$). The set of probability distributions over some finite set X is denoted by $\Delta(X)$. If $\mu \in \Delta(A)$ then $\mu(a_i) = \sum_{a_{-i} \in A_{-i}} \mu(a_i, a_{-i})$ is the marginal probability of $a_i \in A_i$ according to μ and, for every $a_i \in A_i$ with $\mu(a_i) > 0$, $\mu(\cdot|a_i)$ is the conditional of μ on A_{-i} given a_i.

2.1. **The two-populations case.** We consider a random-matching environment with two finite and non-empty populations N_1 and N_2, where $|N_1| = n_1$ and $|N_2| = n_2$. Let $N = N_1 \times N_2$. Following Mailath et al. (1997), we call a *meeting* to a selection of an agent k_i from each population, who then play the game G. The vector $k = (k_1, k_2) \in N$ is called the *cast* of the meeting. The cast of the meeting is drawn uniformly within each population and independently across populations. That is, the probability that the cast of the meeting will be k is $\frac{1}{n_1 \cdot n_2}$, for every $k \in N$.

To the above standard description we now add a new ingredient. This is the exogenously given categorization profile which specifies how every agent partitions the set of his potential opponents. Formally, for every $k_i \in N_i$ let C_{k_i} be a partition of N_{-i}. For $k_{-i} \in N_{-i}$ we denote by $C_{k_i}(k_{-i})$ the element of C_{k_i} which contains k_{-i}. For $i = 1, 2$, let $C_i = (C_{k_i})_{k_i \in N_i}$ be the profile of categorizations used by agents in the population N_i, and let $C = (C_1, C_2)$. A complete description of the environment, therefore, is given by (G, N_1, N_2, C).

A (pure) strategy for agent $k_i \in N_i$ is a function $\sigma_{k_i} : N_{-i} \to A_i$ specifying the action that k_i plays when confronted with each agent from the other population. For $i = 1, 2$ we denote $\sigma_i = (\sigma_{k_i})_{k_i \in N_i}$ and $\sigma = (\sigma_1, \sigma_2)$.

Definition 1. *A strategy σ_{k_i} of player k_i is **adapted** to the categorization C_{k_i} if σ_{k_i} is C_{k_i}-measurable, that is $\sigma_{k_i}(k_{-i}) = \sigma_{k_i}(k'_{-i})$ whenever $C_{k_i}(k_{-i}) = C_{k_i}(k'_{-i})$. The strategy profile σ is adapted to the categorization profile C if σ_{k_i} is adapted to C_{k_i} for every $k_i \in N_i$ and for $i = 1, 2$.*

Let σ_i be a strategy profile of population N_i and let $D \subseteq N_i$ be a non-empty set. We denote by σ_D the *average strategy* of the agents in D. That is, for every $k_{-i} \in N_{-i}$ and $a_i \in A_i$, $\sigma_D(k_{-i}) \in \Delta(A_i)$ is defined by $\sigma_D(k_{-i})(a_i) = \frac{\left|\left\{k_i \in D \,:\, \sigma_{k_i}(k_{-i}) = a_i\right\}\right|}{|D|}$. We are now ready to define our equilibrium concept.

Definition 2. *A strategy profile σ is an equilibrium in (G, N_1, N_2, C) if the following two conditions hold:*
(i) σ is adapted to C; and
(ii) For $i = 1, 2$, for every $k_i \in N_i$ and for every $k_{-i} \in N_{-i}$, $\sigma_{k_i}(k_{-i})$ is a best response to $\sigma_{C_{k_i}(k_{-i})}(k_i)$, that is $u_i\left(\sigma_{k_i}(k_{-i}), \sigma_{C_{k_i}(k_{-i})}(k_i)\right) \geq u_i\left(a_i, \sigma_{C_{k_i}(k_{-i})}(k_i)\right)$ for every $a_i \in A_i$.

2.2. **The single-population case.** For this subsection we assume that G is a symmetric game. It will be convenient to denote the (common) set of actions by B. Thus, $A_1 = A_2 = B$, $A = B^2$ and $u_1(b, b') = u_2(b', b)$ for every $(b, b') \in A$. There is only one population of agents denoted by $N = \{1, 2, \dots, n\}$ where $n \geq 2$. In this case, a meeting involves a drawing of an *ordered* pair from N. The first player takes the role of player 1 while the second that of player 2. The drawing is uniform among all ordered pairs of agents, meaning that the probability that the cast of the meeting is (i, j) (in this order) is $\frac{1}{n(n-1)}$ for every $(i, j) \in N^2$, $i \neq j$.

Each agent categorizes the rest of the agents in the population. Thus, for every $i \in N$ let C_i be a partition of $N \setminus \{i\}$, and let $C = (C_1, \dots, C_n)$. As before, $C_i(j)$ is the element of C_i which contains agent j. A single population environment is characterized by the triple (G, N, C).

A (pure) strategy for player i is a function $\sigma_i : N \setminus \{i\} \to B$. Notice that, both the categorization and the strategy of a player, are independent of the role of this player in the chosen pair. Let $\sigma = (\sigma_1, \dots, \sigma_n)$ denote the strategy profile of the population.

Definition 3. *A strategy σ_i is **adapted** to the categorization C_i if σ_i is C_i-measurable. The strategy profile σ is adapted to the categorization profile C if σ_i is adapted to C_i for every $i \in N$.*

For $i \in N$ and $D \subseteq N \setminus \{i\}$, let $\sigma_D(i)$ be the average strategy of the agents in the set D when they face i. That is, for every $b \in B$, $\sigma_D(i) \in \Delta(B)$ is defined by $\sigma_D(i)(b) = \frac{|\{j \in D \,:\, \sigma_j(i) = b\}|}{|D|}$.

Definition 4. *A strategy profile σ is an equilibrium in (G, N, C) if the following two conditions hold:*
(i) σ is adapted to C; and
(ii) For every two different agents $i, j \in N$, $\sigma_i(j)$ is a best response to $\sigma_{C_i(j)}(i)$.

3. Main result

A correlated equilibrium in G (Aumann, 1974) is any distribution $\mu \in \Delta(A)$ such that $\sum_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i}) \mu(a_{-i}|a_i) \geq \sum_{a_{-i} \in A_{-i}} u_i(a'_i, a_{-i}) \mu(a_{-i}|a_i)$ for every $a_i \in A_i$ with $\mu(a_i) > 0$ and for every $a'_i \in A_i$, $i = 1, 2$.

3.1. **The two-populations case.** Consider a random-matching environment with two populations (G, N_1, N_2, C). Let $\mathbb{P}_\sigma$ denote the distribution over A induced by the strategy profile σ, and for every $k_i \in N_i$ let $\mathbb{P}_{\sigma|k_i}$ denote the distribution over A induced by σ given that k_i was chosen from the population N_i. That is, for $(a_1, a_2) \in A$,

$$\mathbb{P}_\sigma(a_1, a_2) = \frac{|\{(k_1, k_2) \in N \;:\; \sigma_{k_1}(k_2) = a_1,\; \sigma_{k_2}(k_1) = a_2\}|}{n_1 \cdot n_2}$$

and

$$\mathbb{P}_{\sigma|k_i}(a_1, a_2) = \frac{\left|\{k_{-i} \in N_{-i} \;:\; \sigma_{k_i}(k_{-i}) = a_i,\; \sigma_{k_{-i}}(k_i) = a_{-i}\}\right|}{n_{-i}}.$$

Theorem 1. *(i) If a strategy profile σ in a two-populations environment (G, N_1, N_2, C) is an equilibrium then $\mathbb{P}_\sigma$ is a correlated equilibrium in G.*
(ii) For every correlated equilibrium μ with rational numbers in a game G there exists an environment (G, N_1, N_2, C) and an equilibrium σ in this environment such that $\mathbb{P}_{\sigma|k_i} = \mu$ for every $k_i \in N_i$, $i = 1, 2$.

Remark 1. *Since, $\mathbb{P}_\sigma = \frac{1}{n_i}\sum_{k_i \in N_i} \mathbb{P}_{\sigma|k_i}$ $(i = 1, 2)$, part (ii) of Theorem 1 remains valid if $\mathbb{P}_{\sigma|k_i}$ is replaced with $\mathbb{P}_\sigma$.*

3.2. **The single-population case.** Fix a random-matching environment with one population (G, N, C). For a given strategy profile σ and for every $(b, b') \in B^2$ denote

$$\mathbb{P}_\sigma(b, b') = \frac{\left|\{(i, j) \in N^2 \ : \ i \neq j,\ \sigma_i(j) = b,\ \sigma_j(i) = b'\}\right|}{n(n-1)},$$

and for every $i \in N$ let

$$\mathbb{P}_{\sigma|i}(b, b') = \frac{|\{j \in N \setminus \{i\} \ : \ \sigma_i(j) = b, \sigma_j(i) = b'\}|}{n-1}.$$

Theorem 2. *(i) If a strategy profile σ in a one-population environment (G, N, C) is an equilibrium then $\mathbb{P}_\sigma$ is a symmetric[4] correlated equilibrium in G.*
(ii) For every symmetric correlated equilibrium μ with rational numbers in a game G there exists an environment (G, N, C) and an equilibrium σ in this environment such that $\mathbb{P}_{\sigma|i} = \mu$ for every $i \in N$.

Remark 2. *Since $\mathbb{P}_\sigma = \frac{1}{n}\sum_{i \in N} \mathbb{P}_{\sigma|i}$, part (ii) of Theorem 2 remains valid if $\mathbb{P}_{\sigma|i}$ is replaced with $\mathbb{P}_\sigma$.*

4. Proof of Theorem 1

4.1. **Part** (i)**.** This part is a special case of the main theorem in Aumann (1987)[5]. For completeness we provide a proof. Let σ be an equilibrium in (G, N_1, N_2, C) and fix $a_i \in A_i$ with $\mathbb{P}_\sigma(a_i) > 0$. Now, for every $k_i \in N_i$ that satisfies $\mathbb{P}_{\sigma|k_i}(a_i) > 0$ we have

$$\mathbb{P}_{\sigma|k_i}(a_{-i}|a_i) = \frac{\left|\{k_{-i} \in N_{-i} \ : \ \sigma_{k_i}(k_{-i}) = a_i,\ \sigma_{k_{-i}}(k_i) = a_{-i}\}\right|}{|\{k_{-i} \in N_{-i} \ : \ \sigma_{k_i}(k_{-i}) = a_i\}|}$$
$$= \sum_{D \in C_{k_i}} \frac{|D|}{|\{k_{-i} \in N_{-i} \ : \ \sigma_{k_i}(k_{-i}) = a_i\}|} \cdot \frac{\left|\{k_{-i} \in D \ : \ \sigma_{k_i}(k_{-i}) = a_i,\ \sigma_{k_{-i}}(k_i) = a_{-i}\}\right|}{|D|}.$$

Since σ is adapted to C we know that, for every $D \in C_{k_i}$, either $\sigma_{k_i}(k_{-i}) = a_i$ for every $k_{-i} \in D$ or $\sigma_{k_i}(k_{-i}) \neq a_i$ for every $k_{-i} \in D$. Thus, the last sum can be

[4]A distribution $\mu \in \Delta(A)$ is symmetric if $\mu(b, b') = \mu(b', b)$ for every $(b, b') \in B^2$.
[5]See Section 6 for an explanation.

rewritten as

$$\sum_{\substack{D\in C_{k_i}:\\ D\subseteq\{k_{-i}\in N_{-i}\ :\ \sigma_{k_i}(k_{-i})=a_i\}}} \frac{|D|}{|\{k_{-i}\in N_{-i}\ :\ \sigma_{k_i}(k_{-i})=a_i\}|}\cdot\frac{\big|\{k_{-i}\in D\ :\ \sigma_{k_{-i}}(k_i)=a_{-i}\}\big|}{|D|}$$

$$=\sum_{\substack{D\in C_{k_i}:\\ D\subseteq\{k_{-i}\in N_{-i}\ :\ \sigma_{k_i}(k_{-i})=a_i\}}} \frac{|D|}{|\{k_{-i}\in N_{-i}\ :\ \sigma_{k_i}(k_{-i})=a_i\}|}\cdot\sigma_D(k_i)(a_{-i}).$$

In words, the conditional distribution $\mathbb{P}_{\sigma|k_i}(\cdot|a_i)$ is a convex combination of the distributions $\{\sigma_D(k_i)\}_D$ where D runs over all the sets in C_{k_i} which player k_i plays a_i against their members. Since a_i is a best response to $\sigma_D(k_i)$ for every such D it is also a best response to $\mathbb{P}_{\sigma|k_i}(\cdot|a_i)$.

Using a similar argument we have that

$$\mathbb{P}_\sigma(a_{-i}|a_i)=\frac{\sum_{k_i\in N_i}\mathbb{P}_{\sigma|k_i}(a_i,a_{-i})}{\sum_{k_i\in N_i}\mathbb{P}_{\sigma|k_i}(a_i)}=\sum_{k_i\in N_i}\frac{\mathbb{P}_{\sigma|k_i}(a_i)}{\sum_{k_i'\in N_i}\mathbb{P}_{\sigma|k_i'}(a_i)}\cdot\mathbb{P}_{\sigma|k_i}(a_{-i}|a_i),$$

which means that $\mathbb{P}_\sigma(\cdot|a_i)$ is a convex combination of $\{\mathbb{P}_{\sigma|k_i}(\cdot|a_i)\}_{k_i\in N_i,\ \mathbb{P}_{\sigma|k_i}(a_i)>0}$. This implies that a_i is a best response to $\mathbb{P}_\sigma(\cdot|a_i)$. Since this holds for any $a_i\in A_i$ with $\mathbb{P}_\sigma(a_i)>0$ and for $i=1,2$ the assertion is proved.

4.2. **The result of Lehrer-Sorin.** Before proving part (ii) of the theorem we take a small detour to recall an earlier result of Lehrer and Sorin (1997)[6]. We will then see that (ii) is nothing but a reinterpretation of their result.

A *public mediated talk* is defined by two finite sets of messages S_1, S_2, and an announcement map $f : S = S_1\times S_2 \to X$ where X is a finite set of public announcements. In a *public mediated talk mechanism* each one of two players chooses (independently and possibly randomly) a message $s_i\in S_i$ according to a distribution $\tau_i\in\Delta(S_i)$, and the public announcement $f(s_1,s_2)$ is made. Then, each player chooses an action $a_i\in A_i$ according to a *decoding map* $\theta_i : S_i\times X\to A_i$.

Every public mediated talk mechanism induces a probability distribution $\mathbb{P}_{\tau,\theta}$ over A in an obvious way. For a given $\mu\in\Delta(A)$, say that a public mediated talk mechanism *simulates* μ if

(1) $\mathbb{P}_{\tau,\theta}(a|s_i)=\mu(a)$ for every $a\in A$, every $s_i\in S_i$ and $i=1,2$; and

(2) $\mathbb{P}_{\tau,\theta}(a_{-i}|s_i,x)=\mu(a_{-i}|\theta_i(s_i,x))$ for every $a_{-i}\in A_{-i}$, every $s_i\in S_i$ and $x\in X$ having positive probability under $\mathbb{P}_{\tau,\theta}$ and $i=1,2$.

Theorem 3. *(Lehrer and Sorin, 1997) Let $\mu\in\Delta(A)$ be a distribution with rational probabilities. Then there exists a public mediated talk mechanism that simulates μ. Moreover, the mechanism can be constructed such that both τ_1 and τ_2 are the uniform distributions over S_1 and S_2 respectively.*

[6]We do not present here the Lehrer-Sorin result in its full generality. Rather, we present a simpler version which is enough for our needs.

4.3. **Part** (*ii*)**.** Let μ be a correlated equilibrium of G with rational numbers. By Theorem 3 there is a public mediated talk (S_1, S_2, f, X) and decoding maps θ_1, θ_2 such that if each player i chooses his message uniformly within S_i then the public mediated talk mechanism simulates μ. We now need to define a random-matching environment (G, N_1, N_2, C) and an equilibrium σ in this environment such that $\mathbb{P}_{\sigma|k_i} = \mu$ for every $k_i \in N_i$, $i = 1, 2$.

The two populations will be $N_1 = S_1$ and $N_2 = S_2$. For $i = 1, 2$ and for each $k_i \in N_i$, the partition C_{k_i} will be the one generated by $f(k_i, \cdot)$. That is, k_{-i} and k'_{-i} are in the same category in the partition C_{k_i} iff $f(k_i, k_{-i}) = f(k_i, k'_{-i})$. The strategy profile σ will be defined by $\sigma_{k_i}(k_{-i}) = \theta_i(k_i, f(k_i, k_{-i}))$.

We start by showing that σ is an equilibrium in (G, N_1, N_2, C). First, σ is adapted to C by definition. Second, fix $k_i \in N_i$ and $D \in C_{k_i}$. By the definition of C_{k_i} there is $x \in X$ such that $D = \{k_{-i} \in N_{-i} \ : \ f(k_i, k_{-i}) = x\}$. Denote $a_i = \theta_i(k_i, x)$. Since the mechanism simulates μ (property (2)), and since τ_{-i} is the uniform distribution on N_{-i} it follows that

$$\mu(a_{-i}|a_i) = \mathbb{P}_{\tau,\theta}(a_{-i}|k_i, x) = \frac{|\{k_{-i} \in D \ : \ \sigma_{k_{-i}}(k_i) = a_{-i}\}|}{|D|} = \sigma_D(k_i)(a_{-i})$$

for any $a_{-i} \in A_{-i}$. In other words, the conditional distribution $\mu(\cdot|a_i)$ is equal to the distribution $\sigma_D(k_i)$ for any $D \in C_{k_i}$ that satisfies $\sigma_{k_i}(k_{-i}) = a_i$ for any $k_{-i} \in D$. Since, by assumption, μ is a correlated equilibrium a_i is a best response to $\sigma_D(k_i)$ for any such D. This proves that σ is an equilibrium in (G, N_1, N_2, C).

Finally, we need to show that $\mathbb{P}_{\sigma|k_i} = \mu$ for every $k_i \in N_i$, $i = 1, 2$. This is a straightforward consequence of property (1) in the definition of simulating, and of the uniform distribution of the messages τ. Indeed,

$$\mu(a) = \mathbb{P}_{\tau,\theta}(a|k_i) = \frac{|\{k_{-i} \in N_{-i} \ : \ \sigma_{k_i}(k_{-i}) = a_i, \ \sigma_{k_{-i}}(k_i) = a_{-i}\}|}{n_{-i}} = \mathbb{P}_{\sigma|k_i}(a).$$

5. Proof of Theorem 2

To prove part (i), notice first that $\mathbb{P}_\sigma$ is always a symmetric distribution. To show that $\mathbb{P}_\sigma$ is a correlated equilibrium repeat the arguments of subsection 4.1 with the necessary changes.

In contrast to Theorem 1, here we cannot apply the result of Lehrer-Sorin in order to prove part (ii). Instead, we provide a direct proof which uses similar ideas to those in the proof of Lehrer-Sorin but is somewhat simpler. Let $\mu \in \Delta(A)$ be a symmetric correlated equilibrium with rational probabilities. Assume that the (common) set of actions is $B = \{b_1, b_2, \ldots, b_m\}$. Then μ can be described by the matrix $(c_{kl}/d)_{1\leq k\leq m, \ 1\leq l\leq m}$ where $\sum_{k,l=1}^m c_{kl} = d$ and $c_{kl} = c_{lk}$ for every $1 \leq k, l \leq m$. We denote $c_k = \sum_{l=1}^m c_{kl}$.

We will need the following notation. If $(x_1, x_2, \ldots, x_s)$ is a string of s symbols then the *Latin square* corresponding to this string is the $s \times s$ matrix

$$\begin{pmatrix} x_1 & x_2 & \ldots & x_{s-1} & x_s \\ x_s & x_1 & \ldots & x_{s-2} & x_{s-1} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ x_3 & x_4 & \ldots & x_1 & x_2 \\ x_2 & x_3 & \ldots & x_s & x_1 \end{pmatrix}$$

whose lines are successive right shifts of the string.

Now, consider the string of length $2d+1$ defined as follows (a diagram of this string appears below). The first symbol of the string is 0. The next c_1 symbols are b_1 and following them are c_2 symbols of b_2. We continue in this way until we reach c_m symbols of b_m. This defines the first $d+1$ symbols of the string. The last d symbols of the string are defined by (in this order) c_{mm} symbols of b_m, c_{mm-1} symbols of b_{m-1}, $\ldots$, c_{m1} symbols of b_1, c_{m-1m} symbols of b_m, c_{m-1m-1} symbols of b_{m-1}, $\ldots$, c_{m-11} symbols of b_1, $\ldots$, c_{1m} symbols of b_m, c_{1m-1} symbols of b_{m-1}, $\ldots$, c_{11} symbols of b_1. We call it the string *generated* by μ.

$$0\ \underbrace{b_1,\ldots,b_1}_{c_1}\ \underbrace{b_2,\ldots,b_2}_{c_2}\ \ldots\ \underbrace{b_m,\ldots,b_m}_{c_m}\ \underbrace{b_m}_{c_{mm}}\ \ldots\ \underbrace{b_1}_{c_{m1}}\ldots\underbrace{b_m}_{c_{2m}}\ldots\underbrace{b_1}_{c_{21}}\ \underbrace{b_m}_{c_{1m}}\ldots\underbrace{b_1}_{c_{11}}$$

To illustrate the above construction consider the case where $m = 3$ and the (symmetric) distribution μ is given by

	b_1	b_2	b_3
b_1	0	1/7	0
b_2	1/7	1/7	1/7
b_3	0	1/7	2/7

Thus, $c_{11} = 0,\ c_{12} = 1,\ c_{13} = 0,\ c_{22} = 1,\ c_{23} = 1,\ c_{33} = 2$ and $d = 7$. The generated string of length $2d+1 = 15$ is $(0; b_1, b_2, b_2, b_2, b_3, b_3, b_3; b_3, b_3, b_2, b_3, b_2, b_1, b_2)$.

We now describe the environment (G, N, C) and the strategy profile σ. The population is taken to be $N = \{1, 2, \ldots, 2d+1\}$. The action $\sigma_i(j)$ that player i plays when he meets player j is the action that stands in the ij place of the Latin square corresponding to the string generated by μ. Finally, the categorization C_i is according to the strategy σ_i, that is $C_i(j) = C_i(j')$ iff $\sigma_i(j) = \sigma_i(j')$.

Let us return to the last example and consider the Latin square corresponding to the generated string. To make the reading easier we write $x = b_1,\ y = b_2$ and $z = b_3$.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	x	y	y	y	z	z	z	z	z	y	z	y	x	y
2	y	0	x	y	y	y	z	z	z	z	z	y	z	y	x
3	x	y	0	x	y	y	y	z	z	z	z	z	y	z	y
4	y	x	y	0	x	y	y	y	z	z	z	z	z	y	z
5	z	y	x	y	0	x	y	y	y	z	z	z	z	z	y
6	y	z	y	x	y	0	x	y	y	y	z	z	z	z	z
7	z	y	z	y	x	y	0	x	y	y	y	z	z	z	z
8	z	z	y	z	y	x	y	0	x	y	y	y	z	z	z
9	z	z	z	y	z	y	x	y	0	x	y	y	y	z	z
10	z	z	z	z	y	z	y	x	y	0	x	y	y	y	z
11	z	z	z	z	z	y	z	y	x	y	0	x	y	y	y
12	y	z	z	z	z	z	y	z	y	x	y	0	x	y	y
13	y	y	z	z	z	z	z	y	z	y	x	y	0	x	y
14	y	y	y	z	z	z	z	z	y	z	y	x	y	0	x
15	x	y	y	y	z	z	z	z	z	y	z	y	x	y	0

Thus, when players 3 and 5 are the cast of the meeting, player 3 plays b_2 while 5 plays b_1. The categorization of player 9 is $C_9 = \{\{7,10\},\{4,6,8,11,12,13\},\{1,2,3,5,14,15\}\}$.

We now show that $\mathbb{P}_{\sigma|1} = \mu$. A key observation which follows directly from the definition of a Latin square is that $\sigma_j(1) = \sigma_1(2d+3-j)$ for every $j = 2,\dots,2d+1$. In words, the action that player j plays against 1 is the same as the one that player 1 plays against $2d+3-j$.

Fix two actions $b_k, b_l \in B$ (possibly $b_k = b_l$). By the last observation

$$\{j \in N\backslash\{1\} \,:\, \sigma_1(j) = b_k,\ \sigma_j(1) = b_l\} = \{j \in N\backslash\{1\} \,:\, \sigma_1(j) = b_k,\ \sigma_1(2d{+}3{-}j) = b_l\}.$$

We claim that the construction of the generated string is such that the cardinality of this set is $2c_{kl}$. To avoid complex notation, we leave this fact for the reader to convince herself (the diagram is helpful). Thus,

$$\mathbb{P}_{\sigma|1}(b_k,b_l) = \frac{|\{j \in N\setminus\{1\} \ :\ \sigma_1(j) = b_k,\ \sigma_j(1) = b_l\}|}{2d} = \frac{2c_{kl}}{2d} = \mu(b_k,b_l).$$

Next, we show that $\mathbb{P}_{\sigma|i} = \mu$ for every $i \in N$. Notice that, for every $1 \le i \le 2d$ and every $b_k, b_l \in B$,

$$\{j \in N\setminus\{i\} : \sigma_i(j) = b_k,\ \sigma_j(i) = b_l\}{+}1 = \{j \in N\setminus\{i+1\} : \sigma_{i+1}(j) = b_k,\ \sigma_j(i+1) = b_l\}$$

where the summation means adding 1 to each element of the set, and with the convention that $2d+2 = 1$. Thus,

$$|\{j \in N\setminus\{i\} : \sigma_i(j) = b_k,\ \sigma_j(i) = b_l\}| = |\{j \in N\setminus\{i+1\} : \sigma_{i+1}(j) = b_k,\ \sigma_j(i+1) = b_l\}|\,.$$

Since we already showed that $\mathbb{P}_{\sigma|1} = \mu$, it follows by induction that $\mathbb{P}_{\sigma|i} = \mu$ for every $i \in N$.

To finish the proof it should be shown that σ is an equilibrium in (G,N,C). First, it is clear that σ is adapted to C. Second, fix $i \in N$ and $D \in C_i$. By construction, there is $b \in B$ such that $D = \{j \in N\setminus\{i\} \ :\ \sigma_i(j) = b\}$. Thus,

$\sigma_D(i) = \mathbb{P}_{\sigma|i}(\cdot|b) = \mu(\cdot|b)$. Since μ is a correlated equilibrium, a best response to $\mu(\cdot|b)$ is b. Thus, i is playing a best response to $\sigma_D(i)$.

6. Related literature

We start by describing the relation between our model and Aumann's (1987). Fix a two-populations environment (G, N_1, N_2, C) and consider the two-players incomplete information game where the set of possible states of the world is $\Omega = N$, the common prior is $p(k) = \frac{1}{n_1 \cdot n_2}$ for every $k \in \Omega$ and, for every $k \in \Omega$, the game being played at k is G (states of the world are payoff-irrelevant). For $i = 1, 2$, the information partition of player i is given by $\mathcal{P}_i = \{k_i \times D : k_i \in N_i,\ D \in C_{k_i}\}$.

It is straightforward to see that any Nash-Bayes equilibrium of this incomplete information game is an equilibrium of the random-matching environment and vice-versa. By the main theorem in Aumann (1987), the distribution generated by any equilibrium is a correlated equilibria of G. This provides a proof for part (i) of our theorems. However, part (ii) of the theorems doesn't follow from Aumann's results.

More recently, Jehiel and Koessler (2007) study the notion of Analogy Based Expectation Equilibrium (ABEE) in games with incomplete information. They motivate their solution concept by considering a learning process where agents from two populations are randomly drawn to play a fixed game. ABEE of the incomplete information game correspond to steady states of the learning process. Our model is narrower than theirs since we assume that the pattern of meetings is uniform (the same probability for every match) and since the analogy partitions are equal to the information partitions ($\mathcal{A}_i = \mathcal{P}_i$). We hope that this note will somehow help to construct a natural learning process converging to equilibrium.

Conceptually, the most similar paper to ours is Mailath et al. (1997). There it is shown that equilibrium of a random-matching game with local interactions correspond to correlated equilibria of the underlying game. In their model each agent plays a constant strategy (opponent-contingent strategies are not allowed), but the pattern of interactions is not necessarily uniform. This makes the equivalence proof significantly simpler.

As opposed to our approach, most of the literature on random-matching games study the evolution of actions in the population. Kandori et al. (1993) and Young (1993) are two of the well-known examples. Okuno-Fujiwara and Postlewaite (1995) and Mailath et al. (2000) are closer to the current paper since they allow players to use opponent-contingent strategies to a certain extent. Finally, several recent papers study subjects related to categorical thinking in different contexts. This literature includes Fryer and Jackson (2007), Pęski (2006), Azrieli and Lehrer (2007) and Azrieli (2007).

7. Final remarks and possible extensions

7.1. **Mixed strategies and equilibrium existence.** Throughout the paper we restricted attention to the case where players use pure strategies. That is, a strategy of a player is a function specifying the *pure* action that this player uses when facing each possible opponent. A result of this restriction is that equilibrium may fail to exist in some environments. Indeed, if a two-populations environment consists of one player in each population, and the game G has no pure equilibrium then there is no equilibrium in this environment.

Consider the case of two populations[7] and assume that players can use mixed strategies. Here, we understand mixed strategies in the 'behavioral sense', meaning that a mixed strategy for player $k_i \in N_i$ is a function $\sigma_{k_i} : N_{-i} \to \Delta(A_i)$. The definition of equilibrium can be changed in an obvious way to allow for mixed strategies: First, the strategy of each player should be constant in each of the atoms of his categorization; second, $\sigma_{k_i}(k_{-i})$ should be a best response to the average (or expected) strategy of players in $C_{k_i}(k_{-i})$ when they face k_i.

Notice first that allowing players to use mixed strategies eliminates the possibility of equilibrium non-existence. Indeed, if $\mu_1 \in \Delta(A_1),\ \mu_2 \in \Delta(A_2)$ is a Nash profile of G then the strategy profile σ where $\sigma_{k_i}(k_{-i}) = \mu_i$ for every $k_i \in N_i$ and $k_{-i} \in N_{-i}$ $(i = 1, 2)$ is an equilibrium for every categorization profile C.

Second, we claim that part (i) of Theorem 1 holds even if mixed strategies are allowed. Indeed, fix $k_i \in N_i,\ D \in C_{k_i}$ and $a_i \in A_i$ such that $\sigma_{k_i}(k_{-i})(a_i) > 0$ for $k_{-i} \in D$. Then a_i is a best response to $\sigma_D(k_i)$. If $\bar{D}$ is the union of all such sets D then it is not hard to check that $\mathbb{P}_{\sigma|k_i}(\cdot|a_i)$ is a convex combination of $\{\sigma_D(k_i)\}_{D\subseteq\bar{D}}$, which implies that a_i is a best response also to $\mathbb{P}_{\sigma|k_i}(\cdot|a_i)$. It follows that $\mathbb{P}_\sigma$ is a correlated equilibrium.

Finally, we conjecture that any correlated equilibrium (not just those with rational probabilities) can be obtained as the induced distribution of an equilibrium in a random-matching environment, when players are allowed to use mixed strategies. Whether this is true remains to be determined.

7.2. **Non-adapted strategies.** Let G be a game with payoffs as in the following matrix.

	a_2	a_2'
a_1	$0, 0$	$2, 0$
a_1'	$2, 0$	$0, 0$

Consider the two-populations environment (G, N_1, N_2, C) where $N_1 = \{k_1\},\ N_2 = \{k_2, k_2'\}$ and $C_{k_1} = \{N_2\}$. Assume that $\sigma_{k_2}(k_1) = a_2$ and $\sigma_{k_2'}(k_1) = a_2'$ (notice that players in N_2 are indifferent among all the possible outcomes of the game). In this case, player k_1 is indifferent between a_1 and a_1' since both give an expected payoff of 1.

Now, if player k_1 plays the (non-adapted) strategy $\sigma_{k_1}(k_2) = a_1$ and $\sigma_{k_1}(k_2') = a_1'$ then the resulting $\mathbb{P}_\sigma$ is not a correlated equilibrium even though his actions are optimal against the average strategy of his opponents. This example shows that the

[7] The following discussion, with necessary changes, is valid in the single-population case as well.

requirement of adapted strategies is necessary for our results. A similar example can be constructed for the single-population case.

7.3. **Games with more than 2 players.** Throughout the paper we only considered games G with two players. One can, however, generalize the model to allow for games with any number $m \in \mathbb{N}$ of players. Assume that there are m different populations $N_1, \ldots, N_m$, and that every $k = (k_1, \ldots, k_m) \in N := \prod_{i=1}^m N_i$ is chosen with probability $\frac{1}{\prod_{i=1}^m n_i}$ $(n_i := |N_i|)$.

It is not clear in this case what should be the definition of a categorization of a player. There are two options. The first is that a categorization of player $k_i \in N_i$ is any partition of the product $\prod_{j \neq i} N_j$. The interpretation of such definition is that each player categorizes the set of possible compositions of opponents. The second option is that only product partitions are allowed. That is, player k_i has a partition of every N_j, $j \neq i$ and C_{k_i} is the product of these partitions. Here, player k_i categorize each population separately and categorize the composition of opponents accordingly. In my view, the second option is behaviorally more plausible[8].

If the first option is used to define a categorization then the result of Theorem 1 holds in this generalized setting. The proof of this assertion uses a more general form of the Lehrer-Sorin (1997) result. We do not know if the theorem is true when only product partitions are allowed. It is also not clear how to generalize the definition of categorization in the single-population model.

7.4. **Common categorization.** One may want to consider the case where all agents are using the same categorization. A natural question is weather any correlated equilibrium can be achieved with this additional restriction. The negative answer to this question is a consequence of the following argument.

Assume that in a two-populations environment (G, N_1, N_2, C) the categorization profile C satisfies $C_{k_i} = C_{k'_i}$ for every $k_i, k'_i \in N_i$ and for $i = 1, 2$. Denote the (common) categorization of N_i by R_i $(i = 1, 2)$. Fix any two sets $D_1 \in R_1$, $D_2 \in R_2$. Now, given that $(k_1, k_2) \in D_1 \times D_2$, the induced distribution over A is a product measure and, moreover, the marginal distributions over A_1 and A_2 constitute a Nash equilibrium of G. It follows that the induced distribution P_σ is in the convex hull of the set of Nash equilibria of G. Since in general there may be correlated equilibria outside this convex hull, not any correlated equilibrium can be obtained.

7.5. **A continuum of agents.** Our model can be adapted quite naturally to the case where the population consists of a continuum of agents. Consider the two-populations case where N_1 and N_2 are two copies of the interval $[0, 1]$ endowed with its Borel field. A meeting in this case is a random selection of a point in the square $[0, 1]^2$. For $i = 1, 2$, every $t_i \in N_i$ has a (measurable) partition C_{t_i} of N_{-i}. A strategy

[8]There is also a third intermediate option where elements of C_{k_i} are restricted to be product sets, but C_{k_i} is not necessarily a product partition. It is not clear how to interpret this kind of partitions.

for player t_i is a (measurable) function[9] $\sigma_{t_i} : N_{-i} \to A_i$. One can then define an equilibrium in a similar way to that of Definition 1.

Similar results to those of this paper can be proved for the continuous case as well. If the probabilities of the correlated equilibrium μ are rational then one can simply repeat the proof of the finite case, where individual agents are replaced by equi-length intervals. We conjecture that any correlated equilibrium (not just with rational probabilities) can be induced by some equilibrium in such an environment. This subject is beyond the scope of this paper.

7.6. **Comparing categorization profiles.** An interesting direction of research which we didn't pursue here is the effect that different categorization profiles can have on equilibrium payoffs. Assume that two environments (G, N_1, N_2, C) and (G, N_1, N_2, C') are given. Can anything be said about the relation between the sets of equilibrium payoffs in the two environments?

7.7. **Endogenizing the categorization profile.** In our model the categorization profile is part of the exogeneously given environment. It would be interesting to further develop this type of model by relaxing this assumption. One possible approach[10] is to introduce a first stage to the game where each player chooses his categorization, and to consider Nash equilibria of the extended two-stage game.

8. References

Aumann, R.J., 1974. Subjectivity and Correlation in Randomized Strategies. Journal of Mathematical Economics 1, 67-96.

Aumann, R.J., 1987. Correlated Equilibrium as an Expression of Bayesian Rationality. Econometrica 55, 1-18.

Azrieli, Y., 2007. Thinking Categorically About Others: A Conjectural Equilibrium Approach. manuscript.

Azrieli, Y., Lehrer, E., 2007. Categorization generated by extended prototypes – an axiomatic approach. Journal of Mathematical Psychology 51, 14-28.

Fryer, R.G., Jackson, M.O., 2007. A Categorical Model of Cognition and Biased Decision-Making. Forthcoming in Contributions in Theoretical Economics.

Jehiel, P., 2005. Analogy-Based Expectation Equilibrium. Journal of Economic Theory 123, 81-104.

Jehiel, P., Koessler, F., 2007. Revisiting Games of Incomplete Information with Analogy-Based Expectations. Games and Economic Behavior, forthcoming.

Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, Mutation, and Long Run Equilibria in Games. Econometrica 61, 29-56.

Lehrer, E., Sorin, S., 1997. One-Shot Public Mediated Talk. Games and Economic Behavior 20, 131-148.

[9] In order to be able to compute the average strategy in some element of player t_i's categorization, we should also require that $\sigma_{t_{-i}}(t_i)$ is measurable as a function of t_{-i}.

[10] This idea is taken from Jehiel (2005, Section 6).

Mailath, G.J., Samuelson, L., Shaked, A., 1997. Correlated Equilibria and Local Interactions. Economic Theory 9, 551-556.

Mailath, G.J., Samuelson, L., Shaked, A., 2000. Endogenous Inequality in Integrated Labor Markets with Two-Sided Search. American Economic Review 90, 46-72.

Okuno-Fujiwara, M., Postlewaite, A, 1995. Social Norms and Random Matching Games. Games and Economic Behavior 9, 79-109.

Pęski, M., 2006. Prior symmetry, categorization and similarity-based reasoning. mimeo.

Young, H.P., 1993. The Evolution of Conventions. Econometrica 61, 57-84.