EVOLUTION AND ENDOGENOUS INTERACTIONS[1]

George J. Mailath
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104 USA

Larry Samuelson
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, Wisconsin 53706 USA

Avner Shaked
Department of Economics
University of Bonn
Adenauerallee 24-26
D 53113 Bonn, Germany

October 31, 1994

[1]First draft, December 10, 1992. Part of this work was done while George Mailath and Larry Samuelson were visiting the University of Bonn and while all three authors were visiting the Institute for Advanced Studies at the Hebrew University of Jerusalem. We are grateful for the hospitality of both. We thank Ken Binmore for helpful discussions. This work has been presented under the titles "The Evolution of Heterogeneity," "Evolution with Endogenous Interactions," and "Evolution and Correlated Equilibria." Financial support from the National Science Foundation and the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged. The third author apologizes for the Americanisms in the paper.

Abstract

We examine an evolutionary model with "local interactions," so that some agents may be more likely to interact than others. We show that equilibrium strategy choices with given local interactions correspond to correlated equilibria of the underlying game, suggesting a new interpretation for correlated equilibrium. We then allow the pattern of interactions itself to be shaped by evolutionary pressures. If agents do not have the ability to avoid unwanted interactions, then heterogeneous outcomes can appear, including outcomes in which different groups play different Pareto ranked equilibria. If agents do have the ability to avoid undesired interactions, then we derive conditions under which outcomes must be not only homogeneous but efficient.

Keywords: Local interactions, Correlated equilibrium, Evolutionary stability, Networks

Journal of Economic Literature Classification Number C70.

EVOLUTION AND ENDOGENOUS INTERACTIONS

by George J. Mailath, Larry Samuelson, and Avner Shaked

1 Introduction

"It's not what you know, it's who you know that counts." We have all encountered assertions of this type, stressing the importance of interacting with the right people. In a similar vein, how does one explain the tremendous premium prospective MBAs put on attending a top business school if one is unconvinced that such schools differ greatly in their value-added or in the ability of their students? A common response is that the real reason for obtaining an MBA from the right school is the chance to "network" with future captains of industry.

In this paper, we examine groups or populations of players who interact in pairs to play a game. However, a player may not be equally likely to meet each member of the opposing population. Instead, patterns can arise in which some pairs of agents are more likely to interact than others. In addition, players may have both the desire and the ability to affect the pattern of such interactions. Hence, the game involves choosing a strategy that one plays whenever matched with an opponent and also choosing an *activity* that will affect which opponents one is matched with to play the game.

The existence of large populations of players who are matched to play the game suggests an evolutionary model. We follow conventional models in assuming that a player is characterized by a single pure strategy that is played against all opponents the player happens to meet, and in allowing that strategy to be adjusted via an evolutionary process. We depart from conventional models both in assuming that agents can be more likely to meet some opponents than others and in assuming that the activities that affect these meeting patterns are themselves subject to evolutionary pressures. We say that an evolutionary process exhibiting the first characteristic has *local interactions* and a process that also has the second characteristics exhibits endogenous local interactions or simply *endogenous interactions.*

We are interested in local interactions because we believe such interactions to be the rule rather than the exception, and also because such interactions can have important stability implications in an evolutionary model. In particular, local interactions may cause a mutant to face a local environment that differs considerably from the aggregate population. By supplying “bridgeheads” for invasion, primarily by allowing mutants to initially face a relatively high concentration of other mutants, local interactions may allow mutants to succeed that would otherwise be doomed to extinction.[1] If local interactions are important, however, then our attention is naturally directed to the forces that determine these interactions, and hence to endogenous interactions.

We are interested in three questions. What can we say about the rest points of an evolutionary process with local interactions? Can an evolutionary process with endogenous interactions lead to heterogeneous outcomes, in which different groups of agents play different strategies and receive different payoffs? Under what conditions will an evolutionary process with endogenous interactions yield efficient outcomes?

We first examine the case where interactions are local but agents' activities are fixed, so that agents cannot alter the opponents with whom they play. We also hold evolutionary forces at bay and examine Nash equilibria in strategy choices given the local interaction pattern. We find that Nash equilibria with fixed (local) interactions correspond to correlated equilibria of the underlying game. Conversely, for any correlated equilibrium of the underlying game, we can find a pattern of local interactions such that, fixing this pattern, there is a Nash equilibrium in strategy choices that gives the correlated equilibrium outcome.

This result is important for two reasons. First, it suggests a new interpretation for correlated equilibrium, where the differing information partitions with which agents observe a public signal in the conventional interpretation are replaced by differing probabilities of meeting various opponents. Second, we want to describe the evolution of a population with local interactions. What is the state of such a population? A complete description of such a state involves listing the strategy chosen by each agent as well as his pattern of interactions. This is a complicated object, and it is helpful to have a simpler statistic that captures the important features of a state. In conventional evolutionary models, this simple statistic is the mixed strategy profile played by the total population. In our model with local interactions, the simple statistic is the correlated equilibrium induced by the state.

[1] The viscosity model of Myerson, Pollock and Swinkels [**?**] is in this vein.

Our interest then turns to the case in which interactions are endogenous, with both strategies and activities adjusted by an evolutionary process. What properties will stable equilibria of the evolutionary process have, and in particular will they be either homogeneous or efficient? The answers to these questions depend upon the extent to which agents can control their interactions. The key factor here is whether agents can ensure that they do *not* meet certain opponents. If agents do not have the ability to seclude themselves, then heterogeneous stable outcomes are possible, including outcomes in which agents are separated into groups, some of which play a "good" equilibrium and others of which play an inefficient, Pareto inferior "bad" equilibrium. If agents have the ability to seclude themselves without meeting undesired opponents, then stable outcomes must be efficient as well as homogeneous.

These results may initially appear to be counterintuitive. Heterogeneity would appear to be most likely when secluded groups can form; while the coexistence of good and bad equilibrium outcomes would appear to be problematic when those playing the bad outcome cannot be excluded from interacting with those players enjoying the good outcome. However, heterogeneity persists in the latter case, even though agents playing the bad equilibrium can seek interactions with agents playing the good outcome, because the former cannot avoid other agents playing the bad equilibrium. Because some agents cannot avoid others playing the bad equilibrium, it can be a best response to acquiesce in playing the bad equilibrium; rather than seeking other partners with whom to play the good equilibrium and in the process miscoordinating with those who play the bad one. Conversely, efficient outcomes are ensured if groups of agents have the ability to segregate themselves from others. An outcome that is not efficient can be displaced by a small group of agents who seclude themselves and play an efficient outcome. This group will attract other agents who then find it optimal to switch to the efficient outcome because they are ensured of meeting only opponents playing that outcome.

These results depend heavily on the assumption that each agent must play the same strategy against all opponents he happens to meet. In particular, an agent could always (at least weakly) improve his outcome by making his strategy contingent on the strategy played by the opponent he meets.[2] In many cases, however, such discrimination is impossible, either because it is impossible to discern the opponent's strategy before choosing one's own or because the technology of choosing strategies allows only one choice. This inability to make strategies con-

[2]See Banerjee and Weibull [**?**] for a model of "discriminating" players.

tingent on opponents is the essence of the local interaction problem. If contingent strategies could be played, then each interaction could be treated as a separate game and local interaction issues would be uninteresting.

To help in interpreting the process by which agents affect their interactions, we use the analogy of economists at the annual winter meetings of the Allied Social Sciences Associations. Interactions or networking are the heart of these meetings. Those who attend the meetings often devote great attention to arranging their activities so as to achieve desired contacts. Participants interact in many different ways, including interviews, paper sessions, cocktail parties and encounters in hotel lobbies. These activities provide quite different opportunities to control one's pattern of meetings.

The following section considers fixed interaction patterns. Our analogy here is with interviews at the winter meetings, where there is very little ability, once the meetings have started, to control one's interactions. We show that Nash equilibrium strategy choices with fixed local interactions correspond to correlated equilibria of the underlying game. Section 3 introduces the dynamics by which interactions evolve. Sections 4 and 5 examine endogenous interactions. Section 4 shows that stable, heterogeneous outcomes can arise if agents do not have the ability to avoid other agents. Our analogy here is with paper sessions at the meetings, where one can choose which sessions to attend but cannot help but interact with those who attend one's own sessions. Section 5 shows that stable outcomes must be efficient if agents have the ability to interact in isolated groups. Our analogy here is with cocktail parties and hotel lobbies at the winter meetings, where groups can always steal off into seclusion. Indeed, the number of societies and associations (and corresponding cocktail parties) at the meetings has been growing steadily over the years, presumably reflecting a desire for like-minded participants to interact with each other (though we would hesitate to claim that the meetings are efficient).

2 Fixed Interactions and Correlated Equilibria: Interviews

In this section, we explore the implications of equilibrium strategy choices in the presence of local interactions. The pattern of interactions between agents is fixed, so that agents have no freedom to affect the distribution of their opponents.

In terms of the winter meetings, our analogy here is with the interviewing process. Agents interact according to a fixed interview schedule, with very little ability to affect this pattern by arranging interviews at the meetings. The interaction pattern is local, however, with some pairs of agents much more likely to meet than others. In addition, certain aspects of one's strategy are constrained to be the same against all opponents. For example, graduate students typically have only a single paper to present to all interviewers.

The interaction between a pair of agents is described by a finite, two–player normal form game, denoted $G = (S, \pi)$, where $S = S_1 \times S_2$ is the joint action set and $\pi = (\pi_1, \pi_2)$ is the reward function for the pairwise interaction.

We assume that there is a finite population N_1 of player 1's and a finite population N_2 of player 2's. We think of these populations as being large, and require each population to have at least as many members as there are pure strategies for that player, i.e., $|N_1| \geq |S_1|$ and $|N_2| \geq |S_2|$. Each member of each population is associated with a pure strategy, so that $s_1 : N_1 \to S_1$ and $s_2 : N_2 \to S_2$ are functions specifying the strategies of the members of populations 1 and 2, with player $i \in N_1$ playing $s_1(i)$ and player $j \in N_2$ playing $s_2(j)$.

Players from these populations are drawn in pairs, one from each population, to "meet" and play the game.[3] These meetings are described by a list of the number of times that each pair of agents meets. Without loss of generality, we normalize the total number of meetings to one. The proportion of meetings between players i and j can then also be described as the probability that, given a meeting, it involves players i and j. The interactions between the two populations are thus described by a probability distribution μ on the finite space $N_1 \times N_2$, with $\mu(i, j)$ interpreted as either the proportion of all matches that are matches between i and j or as the probability that, given a match, it involves player i from population 1 and player j from population 2.

Definition 1 *A triple $(s_1, s_2; \mu)$, consisting of an assignment of strategies to agents (s_1, s_2) and an interaction pattern μ, is an **equilibrium with fixed interactions** if, for all $i \in N_1$,*

$$\sum_{j \in N_2} \pi_1(s_1(i), s_2(j))\mu(i, j) \geq \sum_{j \in N_2} \pi_1(\hat{s}_1, s_2(j))\mu(i, j) \quad \forall \hat{s}_1 \in S_1. \tag{1}$$

and for all $j \in N_2$,

$$\sum_{i \in N_1} \pi_j(s_1(i), s_2(j))\mu(i, j) \geq \sum_{i \in N_1} \pi_2(s_1(i), \hat{s}_2)\mu(i, j) \quad \forall \hat{s}_2 \in S_2. \tag{2}$$

[3] Notice that we explicitly assume asymmetric interactions, i.e., there is role identification. We can allow for no role identification at the cost of additional notational complexity.

In equilibrium, different members of a population may face different opponents, and hence find different strategies optimal. This has the flavor of a correlated equilibrium, where different signals from a referee also result in different actions being optimal. We now compare the notion of an equilibrium with fixed activities with correlated equilibrium.

Our definition of correlated equilibrium is motivated by the following interpretation.[4] A referee randomly determines an action profile (s_1, s_2) according to some distribution ξ, and then privately recommends the action s_k to player k. The distribution ξ is a correlated equilibrium if it is a best reply for each player to follow the recommendation. More formally:

Definition 2 *A* ***correlated equilibrium*** *is a probability distribution ξ on $S_1 \times S_2$, such that, for $i = 1,2, i \neq k = 1,2$, if $\xi(s_i) > 0$, where $\xi(s_i)$ is the probability that the referee recommends action s_i to player i, then*

$$\sum_{s_k \in S_k} \pi_i(s_i, s_k)\xi(s_k|s_i) \geq \sum_{s_k \in S_k} \pi_i(\hat{s}_i, s_k)\xi(s_k|s_i), \quad \forall \hat{s}_i \in S_i, \tag{3}$$

where $\xi(s_k|s_i)$ is the probability that the referee recommends s_k to player k conditional on s_i being recommended to player i.

A correlated equilibrium allows a player to contemplate changing his strategy (and requires such a change to be suboptimal), but does not allow the player to alter the information that is conveyed by the recommendation received by the referee. This is important, because different recommendations may give different equilibrium payoffs.[5] The counterpart of this in the model with fixed interactions is that the player can choose a strategy, but cannot affect the mix of opponents with whom he plays the game. This is again important, as the player may well prefer some different mix of opponents.

Let strategies be given by (s_1, s_2) and the pattern of interactions by μ. Then the probability that the strategy pair (s_1^*, s_2^*) is played in a meeting between two agents is denoted $\xi(s_1^*, s_2^*; \mu)$ and is given by:

[4] There are several equivalent definitions of correlated equilibrium. An alternative involves specifying an information structure for the players and action choices as functions of signals received. The equilibrium condition is that prescribed choices are optimal given beliefs conditional on the players' information (i.e., signals).

[5] Different recommendations may correspond to different conditional distributions of the opponent's recommendations and hence to different behavior.

Definition 3 *The distribution over strategy pairs generated by* $(s_1, s_2; \mu)$ *is*

$$\xi_{(s_1,s_2;\mu)}(s_1^*, s_2^*) = \sum_{\{i:s_1(i)=s_1^*\}} \sum_{\{j:s_2(j)=s_2^*\}} \mu(i,j). \tag{4}$$

It is straightforward to support *some* outcomes that are correlated but not Nash equilibria as equilibria with fixed interactions: Suppose each population has two players (or two groups of players), called α and α' in population 1 and β and β' in population 2. Let the matching be such that players α and β meet and players α' and β' meet. Moreover, suppose the game G is a battle of the sexes. One equilibrium in fixed activities is then for agents α and β to play one of the pure strategy equilibria of the game and for agents α' and β' to play the other. This equilibrium coincides with a correlated, non-Nash equilibrium of the game G.

Of more interest is the finding that we can support *any* correlated equilibrium as an equilibrium with fixed activities:

Theorem 1 *(1.1) If a triple* $(s_1, s_2; \mu)$ *is an equilibrium with fixed interactions, then* $\xi(s_1, s_2; \mu)$ *is a correlated equilibrium.*

(1.2) If ξ *is a correlated equilibrium, then there exists an equilibrium with fixed interactions* $(s_1, s_2; \mu)$ *such that* $\xi(s_1, s_2; \mu) = \xi$.

Proof (1.1) Fix an equilibrium with fixed interactions $(s_1, s_2; \mu)$. From the definition of $\mu(i,j)$, we have

$$\sum_{j\in N_2} \pi_1(\hat{s}_1, s_2(j))\mu(i,j) = \sum_{s_2\in S_2} \pi_1(\hat{s}_1, s_2) \sum_{\{j:s_2(j)=s_2\}} \mu(i,j).$$

Equation (1) can then be rewritten as, for all $i \in N_1$,

$$\sum_{s_2\in S_2} \pi_1(s_1(i), s_2) \sum_{\{j:s_2(j)=s_2\}} \mu(i,j) \geq \sum_{s_2\in S_2} \pi_1(\hat{s}_1, s_2) \sum_{\{j:s_2(j)=s_2\}} \mu(i,j). \tag{5}$$

Fix a strategy s_1 that is played by some player and sum (5) over all agents playing s_1. This gives:

$$\sum_{s_2\in S_2} \pi_1(s_1, s_2) \sum_{\{i:s_1(i)=s_1\}} \sum_{\{j:s_2(j)=s_2\}} \mu(i,j) \geq$$
$$\sum_{s_2\in S_2} \pi_1(\hat{s}_1, s_2) \sum_{\{i:s_1(i)=s_1\}} \sum_{\{j:s_2(j)=s_2\}} \mu(i,j).$$

Dividing both sides by $\sum_{\{i:s_1(i)=s_1\}} \sum_j \mu(i,j)$ (> 0) yields equation (3), since by equation (4),

$$\xi(s_2 \mid s_1) = \left(\sum_{\{i:s_1(i)=s_1\}} \sum_{\{j:s_2(j)=s_2\}} \mu(i,j) \right) / \left(\sum_{\{i:s_1(i)=s_1\}} \sum_j \mu(i,j) \right).$$

(1.2) Next, suppose we have a correlated equilibrium ξ. First, we select $|S_1|$ players from population 1 and $|S_2|$ players from population 2. We refer to these as the "active" players. Next, construct the functions s_1 and s_2 by assigning each of the pure strategies in S_1 to one of the $|S_1|$ players we have selected from population 1 and assigning each of the pure strategies in S_2 to one of the $|S_2|$ players we have selected from population 2. Then let $\mu(i,j)$ (the probability that agents i and j meet) be equal to zero unless both agents are in the groups of active players that we have selected, and equal to $\xi(s_1(i), s_2(j))$ if they are both in their respective groups of active players. Then the inequalities contained in (3) coincide with those of (1)-(2), and the fact that ξ is a correlated equilibrium implies that we have constructed an equilibrium with fixed activities. □

What is the intuition behind this result? Beginning with an equilibrium with fixed interactions, we construct a correlated equilibrium by recommending the pure strategy combination (s_1, s_2) with the same probability that the equilibrium with fixed interactions produces a match in which (s_1, s_2) is played. Now suppose player 1 receives a recommendation to play strategy s_1^*. The distribution over S_2, describing the opponent's strategy conditional on receiving a recommendation to play s_1^*, is a weighted average of the distributions over opponent strategies faced by all of the population-1 agents in the equilibrium-with-fixed-interactions who play s_1^*. But if s_1^* is a best reply to the distribution over opponent strategies facing each of these agents, then it is a best reply to the weighted average of these distributions. Hence, it is a best reply to play s_1^* when it is recommended, ensuring that we have a correlated equilibrium.

The converse is demonstrated by noting that, given a correlated equilibrium, we can simply assign each pure strategy to a single player and then construct interactions between these players so that the probabilities with which any two players meet matches the probability with which the correlated equilibrium pairs the strategies played by these players. This construction potentially leaves large numbers of players with no possibility for meeting others and playing the game. As the following example shows, however, it is straightforward to bring these players into the game by replacing the individual players in our construction with groups of players.

Example 1 Consider the following version of the game "chicken":

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

Game G1

Nash equilibria of this game include (B, L) and (T, R), as well as the mixed strategy $(.5T + .5B, .5L + .5R)$, for payoffs of $(5, 1)$, $(1, 5)$, and $(5/2, 5/2)$. There are also a number of correlated equilibria that are not Nash equilibria. The symmetric efficient one places probability 1/3 on each of the outcomes (B, L), (T, L), and (T, R).

We construct two models of local interactions with equilibria that coincide with the symmetric efficient correlated equilibrium. First, let each population consist of two players (or select two players from each population). In population 1, one of these players plays T and one plays B. In population 2, one of these players plays L and one plays R. The selected players in population 1 we name "player 1_a" and "player 1_b" and those in population 2 we name " player 2_a" and "player 2_b". Then let μ be given by the following, where the number in each cell is the probability that the corresponding row and column player meet and where each player's strategy is shown in parentheses:

	Player 2_a (L)	*Player* 2_b (R)
Player 1_a (T)	$\frac{1}{3}$	$\frac{1}{3}$
Player 1_b (B)	$\frac{1}{3}$	0

This yields an equilibrium with fixed interactions that coincides with the symmetric efficient correlated equilibrium.

Since each population may contain more than two players, this construction may leave most players with no opportunity for playing the game. We can avoid this by dividing the matching probabilities over multiple agents. For example, let there be three active players in each population with strategies and interactions be given by:

	Player 2_a (L)	*Player* 2_b (L)	*Player* 2_c (R)
Player 1_a (T)	$\frac{1}{6}$	0	$\frac{1}{6}$
Player 1_b (T)	0	$\frac{1}{6}$	$\frac{1}{6}$
Player 1_c (B)	$\frac{1}{6}$	$\frac{1}{6}$	0

This again yields an equilibrium with fixed interactions that corresponds to the symmetric efficient correlated equilibrium of game (G1). Similar constructions allow all players to be active. □

Correlated equilibria require that players make their strategy choices contingent upon the signals they receive. How are these signals generated? While the existence of a publicly observed randomizing device is plausible in many contexts (such as a coin toss to resolve the battle of the sexes), many correlated equilibria (such as in Example 1) require that the randomization device produce *private* signals. Three interpretations of the randomizing device have appeared.

First, this correlation device may represent a referee who recommends strategies to players. This referee has the option of making recommendations contingent on the realization of a random event, and can make recommendations with the property that different players may infer different information from a given recommendation. Second, the players may be basing their choices on the observation of a random event, with the circumstances under which the players observe this random event causing them to have different partitions. A common analogy is that of two agents who base their strategies on the state of the weather and who look out of windows on different sides of a building to collect this information. Third, it may be that the players use pre-play communication to explicitly construct a correlation device (see, for example, Barany [**?**]).

The equivalence between equilibria with fixed activities and correlated equilibria provides another interpretation of correlated equilibrium. Here, the differing partitions of the two players reflect the differing possibilities for meeting other players that arise out of the local nature of the interactions.

In some cases, the pattern of interactions may be fixed by restrictions inherent in the environment. For example, population 1 may be buyers and population 2 sellers of a durable good. Population 1 includes dealers who buy for later resale as well as consumers who buy for private use. Population 2 also includes dealers and private consumers. Examples are the markets for antiques, works of art, used cars, or financial assets. Dealers buy from and sell to both consumers and dealers, but consumers typically interact only with dealers, yielding a local interaction pattern.

In other cases, the interaction pattern may arise endogenously out of the actions taken by agents. We turn to this in the next section.

3 Evolution

The equivalence between equilibria with fixed activities and correlated equilibria depends upon the fact the pattern of interactions is indeed *fixed*, so that agents cannot choose or affect the distribution of their opponents. This restriction is important, because some opponents may provide higher payoffs than others, so that switching between opponents would occur if it were possible. We now expand the model to allow agents to affect their pattern of interactions.

We are interested in two types of questions. First, what patterns of interactions evolve? Second, we know from Theorem 1 that correlated equilibria are likely candidates for rest points of the evolutionary process. Which correlated equilibria will be selected?

We now find it convenient to assume that the players are drawn from infinite populations, so that there is an infinite population of row players (population 1) and a population of column players (population 2), each of which has measure 1. We allow population 1 to consist of J_1 distinct subpopulations or groups. Similarly, population 2 consists of J_2 distinct subpopulations. We think of these groups as being observable labels that are irrelevant to the payoffs of the game and are exogenously attached to players. For example, the groups may be characterized by ethnicity or culture. At the winter meetings, these groups may be fields of specialization among economists or may be departmental affiliations. The role played in the analysis by these groups will depend upon the details of the process by which interactions evolve.

We assume that interactions are determined by some features, characteristics, or choices of the players, which we refer to as *activities*. At the winter meetings, these activities may include attending particular paper sessions or cocktail parties. Denote by A_i^j the finite set of activities for group $j \in J_i$ in population i ($i = 1, 2$) and define $A_i \equiv \cup_j A_i^j$. To simplify notation, we assume $A_i^j \cap A_i^{j'} = \emptyset$ for all $j \neq j'$ and $i = 1, 2$.[6] This includes the case in which different groups can choose the same activity, since we can simply give these activities different names.

For any activity $\alpha_i \in A_i$ and action $s_i \in S_i$, we denote by $q_{\alpha_i s_i}$ the proportion

[6]We now use the index i and k for populations and j for groups.

of agents of population i with activity α_i and action s_i. A specification of $q_{\alpha_i s_i}$ for each activity, action, and population comprises a *state* of the system. We denote a state by θ and the state space by Θ. It is also convenient to define

$$Q_{\alpha_i} \equiv \sum_{s_i \in S_i} q_{\alpha_i s_i},$$

so that Q_{α_i} is the proportion of population i with activity α_i.

The activities of players affect the identities of their opponents. In particular, interactions between players are described by a function $\zeta \colon \Theta \times A_1 \times A_2 \rightarrow [0,1]$. We interpret $\zeta(\theta, \alpha_1, \alpha_2)$ as the probability that, given state θ, a match occurs and is between agents with activities α_1 and α_2. Hence, for all θ,

$$\sum_{(\alpha_1,\alpha_2)\in A_1\times A_2} \zeta(\theta, \alpha_1, \alpha_2) \leq 1. \tag{6}$$

We assume that the matching configuration depends only on the proportion of agents with each *activity*, and not on their actions. Hence, for all $(\alpha_1, \alpha_2) \in A_1 \times A_2$,

$$Q_{\alpha_1}(\theta) = Q_{\alpha_1}(\theta'),\ Q_{\alpha_2}(\theta) = Q_{\alpha_2}(\theta') \ \Rightarrow\ \zeta(\theta, \alpha_1, \alpha_2) = \zeta(\theta', \alpha_1, \alpha_2). \tag{7}$$

Finally, we assume that agents choosing a given activity are randomly selected to participate in matches, so that the probability that an agent of activity α_i is matched is $\sum_{\alpha_k \in A_k} \zeta(\theta, \alpha_i, \alpha_k)/Q_{\alpha_i}$. Hence, for $k \neq i$ we must have

$$\sum_{\alpha_k \in A_k} \zeta(\theta, \alpha_i, \alpha_k) \leq Q_{\alpha_i}. \tag{8}$$

We view $\zeta(\theta, \alpha_1, \alpha_2)$ as a technological phenomenon, determined by the society or environment in which the agents interact. The function $\zeta(\theta, \alpha_1, \alpha_2)$ is thus fixed. However, agents can potentially alter the opponents with whom they interact by altering their choice of activity, and the pattern of interactions may evolve as agents' choices of activities evolve.

The evolutionary process by which agents change both actions and activities is given by $\theta(t, \theta_0)$, a differentiable dynamic system on the state space Θ, with $\theta(t, \theta_0)$ identifying the state of the system at time t given that the initial condition was θ_0. We often write $\theta(t, \theta_0)$ as $\theta(t)$. We take $\theta(t, \theta_0)$ to be a continuous–time system.

Let $\pi_{\alpha_i s_i}$ be the expected payoff to an agent who chooses action s_i and activity α_i (we suppress the dependence of this payoff on the state). We assume that $\theta(t, \theta_0)$ is *monotonic*, by which we mean:

$$\pi_{\alpha_i s_i}(\theta(t)) >\ (=)\ \pi_{\hat{\alpha}_i s_i'}(\theta(t)) \Rightarrow \frac{dq_{\alpha_i s_i}(\theta(t))}{dt}\frac{1}{q_{\alpha_i s_i}} >\ (=)\ \frac{dq_{\hat{\alpha}_i s_i'}(\theta(t))}{dt}\frac{1}{q_{\hat{\alpha}_i s_i'}} \tag{9}$$

for all $\alpha_i, \hat{\alpha}_i \in A_i^j$, $s_i, s_i' \in S_i$, $j \in J_i$, $i = 1, 2$.[7] The key feature of these dynamics is that action/activity pairs that currently earn higher payoffs grow faster. The best-known dynamic process satisfying this condition is the replicator dynamics.

There are several possible interpretations of these dynamics. It may be that individual agents play the game only once, with new agents constantly replacing old ones, but with new agents choosing their strategies on the basis of the payoffs in the previous generation. This may be especially appropriate if the matches are taken to be choices of mates. Alternatively, it may be that agents play the game repeatedly, adjusting their payoffs over time in response to their experience. This may be the appropriate model if the matchings represent social interactions. For more discussion of the various justifications of this type of dynamic analysis, see Mailath [**?**] and Selten [**?**].

We define an **evolutionary rest point** to be a state θ with $d\theta(t)/dt = 0$. This definition leads immediately to a straightforward but important observation concerning evolutionary rest points:

Remark 1 *If all agents in a population have access to the same set of activities, then in an evolutionary rest point, all agents (of the same population) must receive the same payoff (otherwise at least one agent would change activity or action).*

If activities are fixed, then every equilibrium with fixed activities is an evolutionary rest point. The converse need not hold. Every pure state, meaning every state in which all of the agents in a given population play the same strategy, is an evolutionary rest point for the replicator dynamics (for example), but not all of these will be equilibria with fixed activities. The difficulty is that there are pure states in which some agents are not playing best replies (because not all pure strategy combinations are Nash equilibria of the underlying game). These

states are not equilibria with fixed activities, but they are stationary points and hence evolutionary rest points under the replicator dynamics. In particular, the replicator dynamics induces no movement toward best replies as long as these best replies are currently played by a zero proportion of the population. One suspects that these evolutionary rest points will have poor stability properties, with a small perturbation in the direction of a best reply prompting dynamics that lead away from the evolutionary rest points.

[7]Notice that α_i and $\hat{\alpha}_i$ are both taken to be elements of A_i^j in (9). Hence, we are explicitly restricting each group of agents to playing only activities that are available to that group.

We accordingly require that rest points be asymptotically stable. An asymptotically stable state is a state with the property that if the process starts nearby, then it stays nearby (so that the state is stable) and the process converges to the state in question. There may be many states that are stable, meaning simlply that if the system starts nearby, then it stays nearby, (though it may not converge to the state in question). *Asymptotically stable* states are unlikely to exist because the multitude of activities allows ample opportunity for the system to drift between states. We therefore consider a set–valued notion, which is that a set of states, each of which is individually stable, be collectively asymptotically stable:[8]

Definition 4 *The set $\Theta' \subseteq \Theta$ is* ***asymptotically stable*** *if Θ' is closed and connected, and if*

(4.1) for every $\theta \in \Theta'$, $d\theta/dt = 0$;

(4.2) for every $\theta \in \Theta'$ and every open set V containing θ, there exists an open set U with $\theta \in U$ such that $\theta_0 \in U \Rightarrow \theta(t, \theta_0) \in V \ \ \forall t$; and

(4.3) there exists an open set U' with $\Theta' \subset U'$ such that $\theta_0 \in U' \Rightarrow \theta(t, \theta_0) \rightarrow \Theta'$ as $t \rightarrow \infty$.

A stable set is thus a collection of states with the property that each of the states is an evolutionary equilibrium (4.1), the dynamics surrounding each state in the set cannot lead the system too far away from that state (4.2), the dynamics take states near the set to the set (4.3), and points within the set can be viewed as connected via a drift process (the latter is the connectedness assumption). It is the possibility of drift to move the system along a connected set of states, without creating any dynamic forces, that forces us to use a set–valued rather than singleton valued stability notion. A number of papers have recently used similar stability conditions, including Kim and Sobel [**?**], Matsui [**?**], Nöldeke and Samuelson [**?**], Sobel [**?**], and Swinkels [**?**, **?**, **?**].[9]

[8] An alternative formulation would be to replace (4.2) with a minimality requirement.

[9] Gilboa and Matsui [**?**] consider a stability concept that shares a similar spirit of being set–valued, but allows movements between members of the set to arise out of forces other than genetic drift.

4 Endogenous Interactions: Paper Sessions

4.1 Interactions

We now examine a model with endogenous interactions. Our first model of interactions is based on the idea that agents may have preferences over the groups from which their opponents are drawn, prompting us to refer to this as the "preferences" model. These preferences are induced by the possibility that agents in different groups may choose different actions.

There are many ways that a person's actions may affect the set of opponents with whom the person is likely to interact. We examine a particularly simple model of such activities: Each activity for an agent designates one group from the opposing population with whom the agent seeks to interact. This choice will make it more likely that one's opponents are drawn from the preferred group. However, each agent i belongs to a group that opponents may be seeking, and agent i cannot be assured of avoiding meetings with opponents who are seeking i's group, even if i prefers to avoid such meetings.

Continuing with our analogy, we think of discussions held at paper sessions at the winter meetings. The groups consist of those working in the various fields of economics, such as game theory, macroeconomics, labor economics, and so on. Activities consist of choices of paper sessions. Hence, a game theorist may prefer to talk with other game theorists. This objective can be advanced by attending and asking questions at game theory sessions. However, the game theorist must present a paper at his own session, and cannot preclude the possibility of questions from macroeconomists at that session.

To model this type of interaction, we associate with each group of agents a distinct "location". We will often speak of these as if they are physical locations, though other interpretations are possible. An activity consists of an attempt to visit another location to meet members of the group associated with that location. However, there is a positive probability that any agent will be matched at his "home" location with agents from the other population who have chosen to visit that location.

More formally, each agent in group j of population 1 has available J_2 activities (corresponding to the locations or sessions of the other population). We can then denote an activity for population 1 agents by $\alpha_{j\ell}$, where $\alpha_{j\ell}$ represents an attempt by a member of group j to be matched with a member of group ℓ. Similarly, each member of group ℓ in population 2 has available J_1 activities, denoted $\beta_{\ell j}$, where $\beta_{\ell j}$ represents an attempt by a member of group ℓ in population 2 to meet an opponent from group j in population 1. We let $|J_j|$ denote the size of the jth group in population 1 and $|J_\ell|$ the size of the ℓth group in population 2.

We assume that $\alpha_{j\ell}$ causes an agent from group j in population 1 to be assigned to location j in population 1 with probability 1/2, and to location ℓ in population 2 with probability 1/2. The agent then interacts with the agents from the other population who are also assigned to the location to which he has been assigned. This specification is designed to capture the fact that agents can pursue their preferences for matching partners, but cannot completely escape potential matches who are seeking them. Hence, the rich may prefer to meet the rich, but cannot completely escape the poor. The probability 1/2 that a population 1 agent from group j is assigned to location j in population 1 is then interpreted as stating that with probability 1/2, the agent is matched as a result of a population 2 agent seeking a match with group j. The 1/2 probability of being assigned to a population 2 group represents the probability of being matched as a result of one's own efforts to meet a certain opponent.[10]

Notice that we *force* agents to seek a match with some group from the opposing population, i.e., we do not allow players the option of endeavoring to not play the game at all. In some cases, this restriction may be undesirable from the player's point of view, because every possible match may yield a lower payoff than not playing. Our opinion is that if players are to have the choice of hiding from the world and not playing the game, then this should be modeled as one of the choices in the game. We return to the importance of the payoff of not playing the game below.

Given a particular pattern of activities, each location has a collection of population 1 agents, of size H_1, and of population 2 agents, of size H_2. The probability that an agent from population i is matched is given by the function $\rho^i(H_1, H_2)$, $i = 1, 2$. One important example is the *proportional matching rule*, which has the property that all agents on the "short side" of the market are matched:

$$\rho^i(H_1, H_2) = \min\{1, H_k/H_i\} . \tag{10}$$

[10]The 1/2's are motivated by the equal population sizes. A number of alternatives readily suggest themselves. For example, agents may be able to identify a number of groups from the opposing population with whom they would like to play the game, possibly with weights. Agents might also be able to choose the mix of probabilities between being assigned to one's own location and being assigned to a location in the other population. A higher probability of being assigned to the other population would be interpreted as an increased ability to avoid matches with unsought partners.

Note that the same matching function is used at all locations, with the matching probabilities at a location determined only by the relative sizes of the collections of population 1 agents and population 2 agents at that location.

We assume that matching is anonymous within groups. Hence, an agent in population 1 can attempt to meet a member of group ℓ in population 2, but has no control over which member of that group he will meet. Hence, the distribution of the strategies of the opponents he meets is given by the distribution of opponents' strategies in the group to which he is assigned. This reflects our assumption that groups represent observable labels, so that agents can readily seek certain groups from the opposing population, but that agents cannot distinguish between different members of a group and cannot observe their opponents' strategies before play. This is important because the members of that group may play different actions.

Consider an agent of population 1 who chooses activity $\alpha_{j\ell}$, meaning that the agent belongs to group j and would like to meet an opponent from group ℓ. The probability that this agent is matched is given by:

$$\frac{1}{2}\left(\rho^1(\frac{1}{2}\sum_{\kappa=1}^{J_1} Q_{\alpha_{\kappa\ell}}, \frac{1}{2}|J_\ell|)\right) + \frac{1}{2}\left(\rho^1(\frac{1}{2}|J_j|, \frac{1}{2}\sum_{\kappa=1}^{J_2} Q_{\beta_{\kappa j}})\right).$$

The first term represents the probability of being matched as a result of being assigned to location ℓ in population 2; the second term as a result of being assigned to location j in population 1. The arguments of ρ^1 in the first term are the measure of the population-1 agents who are assigned to location ℓ, which is 1/2 of all the agents who attempt to meet members of group ℓ (which is given by $\sum_{\kappa=1}^{J_1} Q_{\alpha_{\kappa\ell}}$); and the number of population-2 agents assigned to that location, which is 1/2 of the population-2 agents who are members of group ℓ (which is $|J_\ell|$). The arguments of ρ^1 in the second term are analogous.

The probability that a match with an opponent from the desired group ℓ of population 2 occurs is given by:

$$\frac{1}{2}\left(\rho^1(\frac{1}{2}\sum_{\kappa=1}^{J_1} Q_{\alpha_{\kappa\ell}}, \frac{1}{2}|J_\ell|)\right) + \frac{1}{2}\left(\rho^1(\frac{1}{2}|J_j|, \frac{1}{2}\sum_{\kappa=1}^{J_2} Q_{\beta_{\kappa j}})\right)\frac{Q_{\beta_{\ell j}}}{\sum_{\kappa=1}^{J_2} Q_{\beta_{\kappa j}}}. \tag{11}$$

The probability that a match with an opponent from group $h \neq \ell$ occurs is:

$$\frac{1}{2}\left(\rho^1(\frac{1}{2}|J_j|, \frac{1}{2}\sum_{\kappa=1}^{J_2} Q_{\beta_{\kappa j}})\right)\frac{Q_{\beta_{hj}}}{\sum_{\kappa=1}^{J_2} Q_{\beta_{\kappa j}}}. \tag{12}$$

An extra term appears in (11) that does not appear in (12) because the population-1 agent in question is seeking a meeting with an opponent from group ℓ, and hence has two possibilities for such a meeting, arising out of his attempt to meet a member of group ℓ or an attempt by a group-ℓ member to meet the population-1 agent. In contrast, the population-1 agent does not seek a meeting with group $h \neq \ell$, and such a meeting potentially occurs only out of the efforts of the group h agent.

4.2 Heterogeneous Outcomes

In this section, we discuss three examples showing that given the matching technology described in the previous section, asymptotically stable states can give heterogeneous outcomes.

Example 2 Consider the following game:

	L	R
T	$4, 4$	$-1, -1$
B	$-1, -1$	$2, 2$

Let the matching rule be proportional. Each population is divided into two groups of equal sizes, denoted groups a and b. Suppose group a agents play T (in population 1) and L (in population 2), and they seek each other. Suppose group b agents play B (in population 1) and R (in population 2) and they seek each other.[11] We can represent this schematically as:

$$1_a\ (T) \Longleftrightarrow 2_a\ (L)$$

$$1_b\ (B) \Longleftrightarrow 2_b\ (R)$$

As a result, half of the agents earn a payoff of 4 (the "rich", or group-a agents from both populations), while half earn a payoff of 2 (the "poor", or group-b agents from both populations). Rich agents meet other rich agents (and play the good equilibrium—this is why they are "rich"), while poor agents meet other poor agents (who, in turn, play the bad equilibrium). This state, considered as a singleton, is asymptotically stable. □

[11]Formally, $q_{\alpha_{aa}T} = q_{\beta_{aa}L} = q_{\alpha_{bb}B} = q_{\beta_{bb}R} = \frac{1}{2}$.

How can the equilibrium in Example 2 be stable when poor agents receive an inferior payoff? Given that poor agents are most likely to meet other poor agents, it is in their best interests to continue to play the bad equilibrium. But why do the poor agents not simply seek rich agents with which to play the good equilibrium of the game? Equivalently, why isn't the movement in the direction of some group b agents in population 1 choosing $(\alpha_{ba}T)$ rather than $(\alpha_{bb}B)$ destabilizing? The rich agents are indifferent between playing against group a agents who choose T and group b agents who choose T, and hence are happy to meet the poor. The difficulty, however, is that the rich are not seeking the poor, and other poor *are* seeking the poor. No matter how much a poor agent desires to play the game with rich agents, he cannot escape the fact that nearly half of his matches are going to be with poor agents who play the bad equilibrium, and it is then not a best response to switch to seeking rich agents and playing the good equilibrium.

In particular, the payoff to a poor agent i in population 1 from playing T and seeking rich agents is $\frac{1}{2}4 + \frac{1}{2}(-1)$, which falls short of the payoff of 2 garnered from playing B and seeking other poor agents. The key here is that agent i can be identified as belonging to group b. Other agents expect group b agents to play B. Hence, population 2 agents who play R attempt to meet the agent i, and i cannot avoid such meetings. These meetings occur often enough that i's best response is to surrender to playing B and receiving the bad-equilibrium payoff.

This example provides some insight into the questions of equality and income distribution. Attention has recently been devoted to the question of how an economy consisting of agents who are *ex ante* identical can give rise to persistent income inequality.[12] Existing theories have shown that as long as economic outcomes are stochastic, different outcome realizations can lead initially identical agents to different incomes. But why do these differences persist rather than being eliminated by the propensity of independent random processes to regress to the mean? The conventional explanation invokes externalities to create a link between current income and future prospects (e.g., Durlauf [**?**]). For example, low incomes can lead to low investment in children's education which leads to low future incomes. If there is also a link between community income levels and the effectiveness of education (perhaps because different income levels lead to different academic habits and norms of achievement), then this cycle can be reinforced and can create an absorbing "poverty trap".

Our example suggests that a poverty trap can arise without externalities (or,

[12] See Durlauf [**?**] for a discussion and references.

depending upon one's interpretation, from externalities in the matching process). Poor agents are poor, and will remain poor, simply because most of their interactions are with other poor people who happen to have coordinated on a low–payoff equilibrium; and because one's wealth or social status is sufficiently correlated with observable characteristics that one cannot simply decide to hereafter be taken for a rich person and interact only with other rich people. This example illustrates the argument that a poverty program serving only a few of the poor and leaving them in their current environment ignores potentially important factors. In particular, to the extent that our example is capturing a crucial feature, it suggests that breaking the "cycle of poverty" will require either a massive intervention to switch the poor groups to the good equilibrium or will require removing poor agents from their environments and putting them into environments where they will be sought by (as well as being able to seek) rich agents.

We need only attach labels such as "black" and "white" to the groups in Example 2 to obtain an outcome in which one group appears to be the victim of discrimination. "Statistical" theories of discrimination have produced models with equilibria in which blacks fare less well then whites even though the latter entertain no antipathy for the former.[13] Example 2 suggests yet another such theory, with the statistical discrimination taking the form of white agents not seeking interactions with blacks because the latter play the bad equilibria, while blacks play the bad equilibria because whites do not seek interactions with them. Once again, the key is that certain behaviors have come to be associated with the labels black and white. For a black agent i, switching behavior does not alter the fact that others still observe simply that i is black, and play against i as they would against other blacks.

One's first impression is that affirmative action policies, by vitiating the statistical profile of the oppressed group, should be able to eliminate such discrimination. Coate and Loury [**?**] show that the effectiveness of affirmative actions in such situations, and especially the question of whether a temporary affirmative action policy can have permanent effects, depends critically upon how the policy affects whites' beliefs about blacks. They find that in some cases only permanent affirmative action policies will be effective. In our case, the affirmative action would have to involve ensuring blacks sufficient access to interactions with whites to allow the good equilibrium to become a best reply for the former. Once blacks were induced to play the good equilibrium, the policy would be unnecessary.

[13] See Coate and Loury [**?**] for a discussion and references.

Example 3 Consider Game G1. Again we suppose that the each populations is divided into two groups, denoted groups a and b. Consider the state in which group a agents play T (population 1) and L (population 2) and seek each other, while group b agents play B (population 1) and R (population 2) and seek group a from the opposing population. This is described schematically by:

$$\begin{array}{ccc} 1_a\ (T) & \Longleftrightarrow & 2_a\ (L) \\ \Uparrow & & \Uparrow \\ 2_b\ (R) & & 1_b\ (B) \end{array}$$

Assume, however, that group a in each population is of size x while group b is of size $1 - x$.[14] It is straightforward to verify that if $1/4 < x < 1/3$, then each agent's activity-action choice is a strict best reply, and so the state is asymptotically stable.[15] *Conditional* on being matched, a group a agent in population 1 is matched with a population 2 agent choosing L with probability $2x/(1+x)$. Also, conditional on being matched, a group b agent in population 1 is matched with a population 2 agent choosing L with probability 1. Thus, for x close to (but less than) 1/3, the conditional distributions are close to those in the correlated distribution that places equal probabilities on the three profiles TL, TR, and BL. □

In contrast to the fixed activities case of Example 1, there is a positive probability that an agent will not be matched in the equilibrium described in Example 3. For example, group-b agents are not matched when the matching process allocates them to their own location (which occurs with probability 1/2). In addition, the entire population is attempting to match with group-a agents of the opposing population, but only a fraction x can succeed at that location. Thus, group b agents in each population are not matched with a probability of $1/2 + (1-x)/2 = 1 - x/2$. An analogous calculation shows that group a agents are not matched with probability $(1 - x)/2$.

As long as interactions are endogenous, so that agents have some ability to affect the identity of their opponents, then we must take seriously the possibility that some agents are not matched or, equivalently, that some agents are matched more often than others. The desired matching plans of all agents may not be compatible or feasible, and the result may be that some agents are frustrated in

[14]Formally, $q_{\alpha_{aa}T} = q_{\beta_{aa}L} = x$, and $q_{\alpha_{ba}B} = q_{\beta_{ba}R} = 1 - x$.

[15]If $x \geq 1/3$, then $\alpha_{aa}B$ is a best reply for group a agents and if $x \leq 1/4$, then $\alpha_{ab}T$ is a best reply for group a agents of population 1 (and similarly for population 2). Moreover, if $x \leq 1/5$, $\alpha_{bb}T$ is a best reply for group b agents in population 1 (and similarly for population 2).

their efforts to meet others. In calculating the payoff to an activity, agents must then include the possibility that they are not matched, with its attendant payoff.

As a result, the payoffs in a game do not provide a complete description of the strategic situation until the payoff to not being matched is specified. While it is tempting to set this payoff to zero, this is not the only possibility. The following example shows how the magnitude of this payoff can affect the outcome of the evolutionary process:

Example 4 Suppose that we alter the payoff to not being matched in Game G1 to be 1 rather than 0. We can then normalize the payoffs in the new game by subtracting 1 from each payoff, *including* the payoff of not being matched, to obtain a game in which the payoff to not being matched is zero and the other payoffs are:[16]

	L	R
T	$3,3$	$0,4$
B	$4,0$	$-1,-1$

As in Example 3, suppose population 1 is divided into two groups, as is population 2. Let x denote the size of the first group (assume that it is the same size for each population). Again suppose all groups wish to meet the opposing group a's; and the row group a's choose T and group b's choose B; the column group a's choose L and the group b's choose R.

Is this an asymptotically stable state when $1/4 < x < 1/3$, as it was in Example 3? Consider a group a player from population 1. Choosing $\alpha_{aa}T$ yields payoffs:[17]

$$\frac{1}{2}\left(x \times \pi_1(T,L) + (1-x) \times \pi_1(T,R)\right) + \frac{1}{2}\left(x \times \pi_1(T,L)\right)$$

$$= \frac{1}{2}\left(x \times 3 + (1-x) \times 0\right) + \frac{1}{2}\left(x \times 3\right) = 3x.$$

The pair $\alpha_{aa}B$ yields:

$$\frac{1}{2}\left(x \times 4 + (1-x) \times (-1)\right) + \frac{1}{2}\left(x \times 4\right) = \frac{1}{2}\left(9x - 1\right).$$

[16]This example is thus equivalent to $G1$ with the payoff to not being matched set at 1.

[17]Since with probability 1/2, the player is at his own location, is not rationed and faces the distribution x of group a and $(1-x)$ of group b; and with probability 1/2, the player is at the opponent group a's location and only matches with probability x.

The pair $\alpha_{ab}T$ yields (note that the player is now never rationed):

$$\frac{1}{2}(x \times 3 + (1 - x) \times 0) + \frac{1}{2}(0) = \frac{3}{2}x.$$

Finally, $\alpha_{ab}B$ yields:

$$\frac{1}{2}(x \times 4 + (1 - x) \times (-1)) + \frac{1}{2}(-1) = \frac{1}{2}(5x - 2).$$

The pair $\alpha_{aa}T$ yields a strictly higher payoff than the other choices if $x < 1/3$. It is also straightforward to verify that the group b's strictly prefer $\alpha_{ba}B$ to the other choices (and similarly for population 2) if $x < 1/3$. We thus have an equilibrium that resembles that of Example 3, but this equilibrium is an asymptotically stable state as long as $x < 1/3$ (rather than also requiring $1/4 < x$, as in Example 3). The difference is that the value of not being matched is higher in Example 4 than in Example 3. This makes it more attractive to endure the rationing associated with $\alpha_{aa}T$ in quest of the relatively high payoff $\pi_1(T, L)$ (rather than avoiding rationing by choosing $\alpha_{ab}T$ and settling for the relatively low payoff $\pi_1(T, R)$). As a result, the constraint $1/4 < x$, needed in Example 3 to ensure that agents choosing $\alpha_{aa}T$ are not too severely rationed, is not needed here.[18] □

5 Endogenous Interactions: Cocktail Parties and Hotel Lobbies

5.1 Interactions

In the preferences model of the previous section, agents could express preferences over desired matches but could not avoid undesired matches. We now examine a model in which agents can sometimes be assured of avoiding certain other agents. We will find it convenient to again describe the interaction technology in terms of locations, and will refer to this as the "location" model.

In our winter meetings parable, the locations are cocktail parties and hotel lobbies. Agents frequent a location in an attempt to meet other agents, and interact

[18] Notice that the payoff to not playing the game does not affect the relative payoffs to strategies $\alpha_{aa}T$ and $\alpha_{aa}B$, since these involved identical rationing frequencies, so that the equilibria in Examples 3 and 4 share the constraint $x < 1/3$, needed to ensure that the payoff of $\alpha_{aa}T$ exceeds that of $\alpha_{ab}B$.

with others who appear at that location. By finding a suitably secluded corner or hotel suite, agents can always ensure that they are not found by those they prefer not to meet. However, this ability to avoid others carries with it a lessened ability to ensure any meeting. There is now no guarantee that any agents of the other population will be at any particular location. One need not attend even the cocktail party of one's own department if one prefers not to, unlike paper sessions.

More formally, suppose there are a finite number of locations, denoted $\lambda \equiv \{\ell_1, \dots, \ell_L\}$. In its simplest form, a choice of an activity consists of a choice of location. In general, however, this choice of location may be random, with an activity inducing a probability distribution over locations rather than corresponding to a single location. For example, an activity might be a decision to frequent one of the convention hotels on a certain day. Once there, this may induce a distribution concerning the likelihood that one is to be found in various locations, including cocktail parties, book displays, bars and the hotel lobby. At each location, other agents are likely to be encountered. Finally, notice that a location need not be a physical location in the usual sense of the word, though the latter provides a convenient interpretation that we shall adopt when describing the model.

The meeting technology is specified by a collection of probabilities of the form $\gamma_{\alpha\ell} \in [0,1]$, where $\gamma_{\alpha\ell}$ is the probability that an agent choosing activity α arrives at location ℓ, so that $\sum_{\ell\in\lambda} \gamma_{\alpha\ell} = 1$. Let $H_{\ell s_1}$ denote the proportion of population 1 agents who play action s_1 and arrive at location ℓ:

$$H_{\ell s_1} = \sum_{\alpha\in A_1} q^1_{\alpha s_1}\gamma_{\alpha\ell}.$$

The measure of population 1 agents arriving at location ℓ is $\sum_{s_1\in S_1} H_{\ell s_1}$. Normalizing to obtain a probability distribution, we have $h_{\ell s_1}$ as the proportion of player 1 agents at location ℓ who play action s_1:

$$h_{\ell s_1} = \frac{H_{\ell s_1}}{\sum_{s_1'\in S_1} H_{\ell s_1'}}.$$

The agents at each location are randomly matched to play, with the possibility arising that not all agents are matched. We let ρ^i_ℓ be the probability that an agent from population i who arrives at location ℓ is matched. We assume, as in the previous section, that this probability is a function only of the total numbers of agents from populations 1 and 2 who arrive at location ℓ:

$$\rho^i_\ell = \rho^i\left(\sum_{s_1\in S_1} H_{\ell s_1}, \sum_{s_2\in S_2} H_{\ell s_2}\right),$$

where $\rho^i_\ell : [0,1]^2 \to [0,1]$. We require ρ^i_ℓ to be continuous on $[0,1]^2\setminus\{(0,0)\}$,[19] and that the number of matches by population 1 agents equal the number of population 2 matches:

$$\rho^1_\ell\left(\sum_{s_1\in S_1} H_{\ell s_1}, \sum_{s_2\in S_2} H_{\ell s_2}\right)\sum_{s_1\in S_1} H_{\ell s_1} = \rho^2_\ell\left(\sum_{s_1\in S_1} H_{\ell s_1}, \sum_{s_2\in S_2} H_{\ell s_2}\right)\sum_{s_2\in S_2} H_{\ell s_2}.$$

We also assume that ρ^1_ℓ is increasing in its first argument and decreasing in its second, while ρ^2_ℓ is decreasing in its first argument and increasing in its second. This implies that an agent is more likely to be matched the fewer agents from his own population at the location and the more agents from the other population at the location.

We restrict attention to the case in which ρ^1_ℓ is homogeneous of degree zero, so that there are no congestion effects at locations. We refer to this as the case of **constant returns to scale**, in the sense that increasing the number of agents at a location yields a like increase in the expected number of matches at that location. The proportional matching rule used in the previous section is an example of a constant returns to scale rule. Other interesting possibilities are **increasing returns to scale**: $\rho^i_\ell(an_1, an_2) > \rho^i_\ell(n_1, n_2)\ \forall a > 1$; and **decreasing returns to scale**: $\rho^i_\ell(an_1, an_2) < \rho^i_\ell(n_1, n_2)\ \forall a > 1$. Increasing returns to scale capture the idea that more agents at a location make it more likely that agents will "meet" each other (such as in search models), while decreasing returns to scale capture congestion effects—more agents at a location make it less likely that agents will "meet" each other.

The preference model of the previous section is a special case of the location model. To obtain the preference model we let J_1 be the set of locations associated with population 1, J_2 the set of locations associated with population 2, $J_1 \cap J_2 = \emptyset$, and for $i \in J_1, j \in J_2$, $\gamma_{\alpha ij\ell} = 1/2$ if $\ell = i, j$, and 0 otherwise.

More generally, the matching function ζ implied by the location model is given

[19]It is necessary to exclude $(0,0)$ from the domain where continuity is required: Suppose $\rho^i_\ell(n_1, n_2)$ is given by the proportional matching rule of the previous section. Then, as n_i and $n_k \to 0$, $\rho^i_\ell(n_1, n_2) \to 0$ if $n_k/n_i \to 0$ while $\rho^i_\ell(n_1, n_2) = 1$ if $n_k > n_i$.

by:[20]

$$\zeta(\alpha_1, \alpha_2, \theta) = \sum_{\ell \in \lambda} \frac{\rho_\ell^1 \gamma_{\alpha_1 \ell} Q_{\alpha_1} \gamma_{\alpha_2 \ell} Q_{\alpha_2}}{\sum_{\alpha_2' \in A_2} \gamma_{\alpha_2' \ell} Q_{\alpha_2'}} = \sum_{\ell \in \lambda} \frac{\rho_\ell^2 \gamma_{\alpha_1 \ell} Q_{\alpha_1} \gamma_{\alpha_2 \ell} Q_{\alpha_2}}{\sum_{\alpha_1' \in A_1} \gamma_{\alpha_1' \ell} Q_{\alpha_1'}}.$$

5.2 Efficient Outcomes

We now derive conditions under which the outcome in the location model is not only homogeneous, but also satisfies an efficiency condition. The sufficient conditions for efficiency come in three parts. First, we need the matching technology to be unrestricted, meaning that all agents have access to the same set of activities. Second, we need it to be possible that an isolated group of agents can form. Third, we require that neither large nor small groups are penalized in terms of matching probabilities, which holds if the matching technology exhibits constant returns to scale.[21]

Isolated groups of agents will be able to form if the matching technology contains pure activities:

Definition 5 *An activity α is* pure *if there exists a location ℓ with the property that $\gamma_{\alpha\ell} = 1$.*

Theorem 2 *Let all agents (of both populations) have access to the same set of activities in the location model. Suppose that game G has a Nash equilibrium whose payoffs strictly dominate the payoffs of all other correlated equilibria. Suppose that the matching process exhibits constant returns. Let there be at least three locations, with a pure activity for each population associated with each location. Then in every state in an asymptotically stable set, every agent is choosing the efficient equilibrium action.*

The proof appears in the Appendix. To see why the result holds, consider first a state in which there is a "free" location, meaning that there is a location at which no agents appear. Suppose further that there is a pure activity corresponding to this location for each population, and that agents are currently earning less than the

[20] Note that number of matches by a population 1 agents choosing α_1 at ℓ is $\rho_\ell^1 \cdot \gamma_{\alpha_1 \ell} Q_{\alpha_1}$, the total number of matches at ℓ is $\rho_\ell^1 \cdot \sum_{\alpha_1' \in A_1} \gamma_{\alpha_1' \ell} Q_{\alpha_1'}$, and the number of matches at ℓ between a population 1 agent choosing α_1 and an opponent choosing α_2 is $\rho_\ell^1 \cdot \gamma_{\alpha_1 \ell} Q_{\alpha_1} \cdot \gamma_{\alpha_2 \ell} Q_{\alpha_2} / \left(\sum_{\alpha_2' \in A_2} \gamma_{\alpha_2' \ell} Q_{\alpha_2'} \right)$.

[21] We assume throughout that equilibria exist that give all agents higher payoffs than not being matched, so that efficiency requires that the game be played.

efficient payoff. Then the existing state is not part of an asymptotically stable set. Instead, consider a perturbation that attaches a small number of agents from both populations to the pure activity corresponding to the vacant location, with these agents playing the actions of the efficient equilibrium. This activity/action pair will earn a higher payoff than any other activity/action pair, attracting agents to the previously vacant location and the efficient equilibrium and thus destabilizing the existing state and precluding its membership in an asymptotically stable set.

The potential difficulty is that the current state may entail agents visiting every location, so that there is no free location with which to work. However, we show that every candidate for an asymptotically stable set of states includes states with unused locations. Intuitively, the system can drift among the states in the stable set, until a state is reached with an unused location. We refer to this process as "freeing a location". The assumption of constant returns to scale is used in this demonstration, as it allows agents to drift between locations without creating congestion affects at these locations. Once a state with a free location is reached, the perturbation described in the previous paragraph prompts dynamics that lead away from the state, precluding its stability.

In order to conclude that asymptotically stable sets are efficient, Theorem 2 assumes the existence of an equilibrium that *strictly* dominates all other equilibrium payoffs. We cannot reduce this requirement to weak domination. Suppose, for example, that the stage game is given by:

	L	R
T	$1,1$	$2,1$
B	$1,2$	$2,2$

Then evolution puts no pressure on strategy choices and the only candidate for an asymptotically stable set is the entire state space, not all of which is efficient.

The forces ruling out inefficient behavior are analogous to those appearing in cheap talk games, including Kim and Sobel [**?**], Matsui [**?**], and Sobel [**?**]. Consider a model in which a round of communication is allowed before a game is played. Consider an equilibrium that is not efficient, and suppose a state yielding this equilibrium belongs to a stable set. If the message space is rich enough to ensure that some messages are not sent, then the stable set contains a state featuring the same play of the game, but in which all agents would respond to the unsent message by playing their part of the Pareto dominating equilibrium. A perturbation that causes that message to be sent now creates dynamics that lead away from the allegedly stable set, which is a contradiction.

A similar mechanism is at work in our model. The place of the unsent message is taken by an unused location with a matching pair of pure activities. The condition that the message space be sufficiently rich to ensure that there is an unsent message is replaced by the condition that there exists at least three locations with pure activities, which we show is sufficient for drift to be able to free a location.[22]

The assumption of constant returns to scale in the matching technology plays two roles in the argument. It ensures that a location can be freed and that agents can be relocated to the freed location. If ρ_ℓ^i exhibits decreasing returns to scale, so that locations suffer from congestion, it may be difficult to free a location by adding agents to other, populated locations.[23] If ρ_ℓ^i exhibits increasing returns to scale, then it may be difficult to attract agents to playing the efficient outcome at a freed location.[24]

Our result says nothing about the *existence* of asymptotically stable sets. General sufficient conditions for existence of an asymptotically stable set in evolutionary models are severe. In particular, evolutionary models readily yield characterization results but only grudgingly offer general existence results. A trade-off then arises. Existence can be ensured under mild conditions by appealing to a solution concept that places little structure on the resulting "stable" sets, allowing these sets to be large and placing little structure on the states that appear in such sets. Alternatively, one can ask for well-behaved stable sets, such as asking that each state in such a set be a rest point of the dynamics (our condition (4.1)), but existence can then be ensured only under strong assumptions. For example, Kim

[22] The destabilizing perturbation in cheap talk models requires agents from a single population to adopt an unused message, even if both populations have the opportunity to send messages before play. In contrast, our destabilizing perturbation sends agents of *both* populations to the free location. If our dynamics allowed the population proportion playing a strategy to grow whenever that strategy is a best reply, even if it is currently played by a zero proportion of the population, then it would suffice for members of a single population to switch to a vacant location (at which point members of the other population would be drawn to the new location). Notice that because actions in a cheap talk model are contingent on messages, the sending of a new message is equivalent to switching *both* agents to a new location. Moreover, if a population-1 agent deviates to the new message in a cheap talk model, then it is possible that drift within the equilibrium component has already endowed population-2 agents with the behavior at this new message that is needed to destabilize the current component.

[23] While the payoff conditional on matching is unchanged as agents are added to a location, the payoff from the corresponding activity falls since the probability of matching has fallen.

[24] The payoff conditional on matching is high at this location, but the payoff from the corresponding activity is low because the probability of matching is low.

and Sobel [**?**] (in a model of cheap talk games) and Nöldeke and Samuelson[**?**] (in a model of extensive form games) find that demanding conditions, involving essentially the existence of an outcome whose payoffs strictly dominate the payoffs of every other outcome, are needed to establish the existence of well-behaved stable sets.[25] In the presence of endogenous interactions, where there are activities that affect payoffs only indirectly through their effect on the matching process, additional strong assumptions on the matching technology would be required, such as assumptions that the matching scheme is proportional and all activities are pure.

6 Conclusion

Our results provide two main insights into local interactions. First, they give rise to correlated equilibria, providing a new interpretation for the latter. Second, outcomes can depend crucially on the ability of agents to affect their interactions. We have seen that the ability to avoid undesired matches can make the difference between efficient outcomes and outcomes that doom some agents to inferior payoffs. This suggests that more attention might usefully be devoted to the process by which agents are matched to play games; a topic we hope to pursue further.

7 Appendix: Proof of Theorem 2

Let s^* be the strategy profile of the strict Nash equilibrium identified in the Theorem. Suppose Θ' is an asymptotically stable set and $\theta \in \Theta'$. Because all agents have access to the same activities, all agents of population i must receive the same payoff in state θ (otherwise it cannot be that $d\theta/dt = 0$); let π_i' denote this payoff. Let s^* be the efficient Nash equilibrium. We now suppose that $\pi_i' < \pi_i(s^*)$ and derive a contradiction.

By hypothesis, there are three locations associated with pure activities; label them as $1, 2$, and 3. Let $H_\ell^i \equiv \sum_{s_i \in S_i} H_{\ell s_i}$, the size of the population of player i agents at location ℓ.

[25]To see the difficulties that can arise, notice that Ritzberger and Weibull [**?**] present an example (Figure 3) of a game with a unique Nash equilibrium component that satisfies condition (4.3) from our definition of stability but not (4.2). In a normal form game, this type of outcome could be excluded by requiring the Nash equilibrium to be strict. In the presence of cheap talk or in an extensive form game, strict equilibria generally do not exist, so that a very strong condition on payoffs, such as an outcome producing payoffs strictly higher than any other, is needed.

Suppose that at one of the locations, say $1, H^i_1 = 0$, for $i = 1, 2$, so that no agents arrive at that location under θ'. Then consider the following perturbation of state θ. Reduce by a factor δ the proportions of agents at every activity/action combination for each population. (Since the ratio of the population of player 1 agents to the size of the population of player 2 agents at each location is unchanged by this, the payoffs to the activity/action choices in θ' have not changed.) Place these agents at the pure activity associated with location 1, and have them play s^*. This yields a payoff at location 1 that is strictly higher than the payoff under θ'. As a result, the dynamics will converge to a state in which all agents are choosing the pure activity associated with location 1 and playing s^*_i. This ensures that θ is not stable (i.e., fails condition (4.2)) and hence that Θ' is not asymptotically stable.[26]

It remains to show that Θ' contains a state in which $H^i_\ell = 0$ for $i = 1, 2$ and some $\ell \in \{1, 2, 3\}$. Suppose, first, that Θ' contains a state θ' in which $H^1_\ell = H^2_\ell$ for some $\ell \in \{1, 2, 3\}$. Without loss of generality, suppose this is location 1. Let f^1_i denote the distribution of actions faced by a player j ($j \neq i$) agent at location 1. Now consider the following perturbation. Reduce by δ the proportions of agents at every activity/action combination for each population. (As before, since H^1_ℓ/H^2_ℓ for all ℓ is unchanged by this, the payoffs to the activity/action choices in θ' have not changed.) Place these agents at the pure activity associated with location 1, and have the agents of player i play choose actions in accordance with f^1_i. Since $H^1_1 = H^2_1$, the payoffs at this pure activity have not changed. This perturbed vector is thus also an element of Θ'. Continuing in this way will eventually result in a state, contained in Θ', in which all agents choose the pure activity associated with location 1, ensuring that no agents choose the pure activities associated with locations 2 and 3.

Finally, suppose, $H^1_\ell \neq H^2_\ell$ for all $\ell \in \{1, 2, 3\}$. Without loss of generality, suppose H^1_ℓ/H^2_ℓ is maximized at location 1 and minimized at location 2. Note in particular that $\eta_1 \equiv H^1_1/H^2_1 > 1$ and $\eta_2 \equiv H^1_2/H^2_2 < 1$, so that $\tau \equiv \frac{(1-\eta_2)}{(\eta_1-\eta_2)} \in (0, 1)$. Let f^ℓ_i denote the distribution of actions faced by a player j ($j \neq i$) agent at location ℓ. Consider the following perturbation. Reduce by δ the proportions of agents at every activity/action combination for each population. (As before, since H^1_ℓ/H^2_ℓ for all ℓ is unchanged by this, the payoffs to the activity/action choices in θ' have not changed.) Assign a fraction $\tau\eta_1$ of the player 1 population and a fraction τ of

[26]Note that the rate of inflow to location 1 may be different for the two populations, and hence rationing at location 1 may occur. However, the non-rationed population will still have an inflow, eventually relieving the rationing of the other population at location 1.

the player 2 population to location 1, and a fraction $1 - \tau\eta_1 = (1 - \tau)\eta_2$ of the player 1 population and a fraction $(1 - \tau)$ of the player 2 population to location 2. Specify that the agents of player i at location ℓ play in accordance with f_i^ℓ. By construction, the ratio of the size of the population of player 1 agents to the size of the population of player 2 agents at locations 1 and 2 have not changed, and so the payoffs at these locations have not changed. Thus the perturbed vector is an element of Θ'. The argument is now completed as in the previous paragraph. □

References