SOCIAL GAMES: Matching and the Play of Finitely Repeated Games

Matthew O. Jackson and Alison Watts

Revised: January 29, 2008

Abstract: We examine a new class of games, which we call social games, such that players not only choose strategies but also choose with whom they play. A group of players who are dissatisfied with the play of their current partners can join together and play a new equilibrium. This imposes new refinements on equilibrium play in games, and we show how play depends on the relative populations of players in different roles, among other things.

We also introduce finitely repeated social games where players may choose to rematch in any period. Some equilibria of fixed-player finitely repeated games cannot be sustained as equilibria in a finitely repeated social game. Conversely, the set of repeated matching equilibria includes some plays that are not part of any subgame perfect equilibrium of the corresponding fixed-player repeated games. We explore existence of finitely repeated matching equilibria, the relationship to renegotiation-proof equilibrium, and show how new predictions are made in trust and centipede games.

Journal of Economic Literature Classification Numbers: A14, C71, C72, C78, J41.

Keywords: Social Games, Matching, Games, Repeated Games, Renegotiation

Department of Economics,
Stanford University, Stanford, CA 94305-6072, USA
(jacksonm@stanford.edu)
http://www.stanford.edu/~jacksonm

Department of Economics,
Southern Illinois University, Carbondale, IL, 62901, USA
(wattsa@siu.edu)

Financial support under NSF grant SES-0316493 is gratefully acknowledged, as is support from the Lee Center for Advanced Networking and the Center for Advanced Studies in the Behavioral Sciences. We thank Federico Echenique, Jan Eeckhout, Hubert Stahn, Phil Reny, Myrna Wooders, and the associate editor and anonymous referees for helpful comments and suggestions.
1. Introduction

In many social and economic interactions, players have choices not only of what actions to play, but also with whom they interact. For instance, if an employee does not like the behavior of his or her employer, he or she can quit and seek a job with another firm. Similarly, in some contexts dissatisfied employers can fire their employees and hire new ones. We see similar examples of rematching in a variety of settings ranging from divorce, excommunication, ostracism, to the breakup of partnerships or alliances, and the formation of new ones. This ability to rematch has strong implications for behavior within the relationships. Although this is a relatively obvious statement and has been studied in some specific contexts such as those mentioned above, there is no systematic or general method of modeling the play within a game when such play depends on players’ ability to rematch. In this paper we introduce such a methodology and show that it can have strong and intuitive implications for behavior. We examine a new class of games called social games such that players not only choose strategies but also choose with whom they play.

We examine two situations: one such that the choice of matching is made just once, and another where the interaction occurs over a finite number of periods and players may rematch in any period. In the one-shot version of a social game, a “matching equilibrium” consists of a matching of players into various groups who will each play the game together, as well as a description of what each player will play. This equilibrium must satisfy two requirements: first, the play of each group must be a Nash equilibrium; and second, no set of players could all improve their payoffs by leaving their current groups, forming a new group, and playing some other Nash equilibrium. In the finitely repeated version of the game, a “matching equilibrium” includes both a specification of what each player will play given each possible history of matching and play (by all players), as well as a specification of who is matched with whom given each possible history. The equilibrium definition is an inductive one. It requires that no group of players could jointly deviate and play a different matching equilibrium in the continuation and all improve their payoffs. We provide two different definitions of repeated matching equilibrium depending on how we treat the possibilities for rematching with other players.

Our results explore the existence of matching equilibria, as well as their structure. The existence of matching equilibrium depends on whether the game is repeated and how large the groups are that can rematch. We provide detailed results outlining these issues. In terms of what we learn about the structure of matching equilibria, here is a partial list of some of the results:

• the threat of rematching selects equilibria that are Pareto undominated by other Nash equilibria,
• player roles with smaller populations are relatively favored,
• payoffs across matched groups cannot differ by too much (a form of equal treatment),
• in repeated games, the threat of rematching supports play that is not supportable without rematching,
• in repeated games, equilibrium play can differ completely from all previous equilibrium concepts (without rematching).

Our results and model make three contributions: First, we provide a general framework and methodology for analyzing how matching and the play of games interact. Second, we show that a number of results regarding relative sizes of populations in markets, contracting, and bargaining, have a fairly general analog where the smallest population(s) are relatively favored in equilibrium play. Third, we show how repetition can change behavior in interesting ways. For example, it can induce trust as a unique equilibrium outcome in finitely repeated trust games. Moreover, it provides new predictions about how trusting behavior should depend on the relative numbers of players in different roles.

Related Literature

One obvious strand of related literature is the study of matching markets that followed the seminal paper of Gale and Shapley (1962) and is detailed in Roth and Sotomayor (1989). There players choose with whom they are matched, but not what happens subsequent to being matched. Here we combine the two decisions and see how relative population sizes of players can impact the play of the game.\(^1\)

Another related strand of literature concerns renegotiation-proof equilibria in finitely repeated games (e.g., Bernheim and Whinston (1987), Bernheim and Ray (1989), Farrell and Maskin (1989), and Benoit and Krishna (1993)). This corresponds to the other extreme of our model where the matching is degenerate but the game is not. In that special case, our definitions correspond exactly to renegotiation-proofness. In contrast, when there are multiple possible matchings, then the relationship between our equilibria and renegotiation-proof equilibria is more complicated. New equilibria emerge due to the threat of rematching, and in some cases our equilibria differ completely from the set of renegotiation-proof equilibria, and include plays that are not supported by any previous equilibrium concept.

Endogenous interactions and implications for behavior have also been studied in some specific contexts. These include market and bargaining settings (e.g., Kelso and Crawford (1982), Rubinstein and Wolinsky (1985), Crawford (1991), Hatfield and Milgrom (2003)\(^2\)), implementation (e.g., Jackson and Palfrey (1999)); mutual investment

---

\(^1\)In the social games considered in this paper, players care only about how their partners play and not who their partners are, and so these do not generalize the matching markets of Gale and Shapley (1962). However, in a companion paper, Jackson and Watts (2008), we provide definitions of social games and matching equilibria that generalize both the definitions here and the definitions from the standard matching world. There we show that well-known results on existence and the lattice structure of matchings from the bipartite marriage market and matching world have analogs in social games.

\(^2\)Hatfield and Milgrom (2003) is probably the most related to our work in that it allows for an abstract set of contracts to be reached between matched players. Contracts can be seen as a special case in our setting by modeling a game where each player’s strategy is a choice of contract and payoffs and incompatible contract choices create negative payoffs. However, Hatfield and Milgrom (2003) focus on matchings between doctors and hospitals (who can have multiple partners), and do not deal with the self-enforcement of contracts, or more than two player roles, or repetition. So there is essentially no overlap in results between their work and ours.
(Watson (1999)), network formation in coordination games (e.g., Jackson and Watts (2002), Droste, Gilles, and Johnson (2003), Corbae and Duffy (2003), and Goyal and Vega-Redondo (2004)), principal-agent games (e.g., Casas-Arce (2005)), games where agents choose their location (Mailath, Samuelson, and Shaked (2001)), and endogenous partnerships in prisoner’s dilemma games (Rob and Yang (2003) and Ghosh and Ray (1996)), as well as a large literature on evolutionary game theory where players are repeatedly randomly matched. Together, these papers make it clear that endogenizing interaction can affect play, albeit in specific settings. Our contribution is to develop a general framework for analyzing matching in game theoretic settings and providing results that give us some systematic understanding of how the ability to choose with whom one interacts affects the play of a game.

Another related literature is that dealing with cooperation and trust in repeated trust or centipede games (e.g., McKelvey and Palfrey (1992), Buskens (2000), Bolton and Ockenfels (2000), Anderhub, Engelmann and Guth (2002), and Kreps et al. (1982), among others). In these incomplete information models there are some players who are payoff-maximizing and others who are altruistic; trust is achieved in equilibrium when the payoff-maximizers attempt to mimic the altruistic types for the first part of the game. Trust can also be achieved in learning models where non-opportunistic players must learn to play the trusting strategy (e.g., Selten and Stoecker (1986)) or in evolutionary models where only high-payoff strategies survive (e.g., Rosenthal (2001)). In contrast, trust is achieved in a repeated matching equilibrium due to either fear of ostracism if one is not trusting or due to the ability to ostracize others for not being trustworthy. These two different mechanisms for trust relate to different relative sizes of populations, and the corresponding equilibrium plays differ as well. For instance, for some population configurations only one player role will exhibit cooperation, for other configurations both roles may exhibit cooperation, and for still other configurations neither role may exhibit cooperation.

Finally, before proceeding let us say a few words about the combined cooperative/noncooperative approach that we take (we discuss this more fully after providing formal definitions). Our approach combines noncooperative and cooperative game theoretic concepts, as we use Nash equilibrium to ensure that the play within any group is self-enforcing; while we use core-based definitions to ensure that the matching is stable and to derive implications from potential rematchings. We could have approached the problem entirely from one perspective or the other. However, we feel that we have chosen the right balance and coupling. The non-cooperative nature of Nash equilibrium really captures the self-enforcing nature of behavior that is necessary to make sure that each player will follow through with the strategies that are part of an overall equilibrium.

---

Our approach is very different from many of these in style and substance. For instance, the random matching in evolutionary games allows one to study how strategies perform against varieties of other strategies and subsequent selection patterns. Here we really isolate the idea that players have discretion in with whom they play, and try to derive the implications. There are circumstances where these approaches lead to the same conclusions and others where they lead to very different conclusions. For instance, in simple coordination games, the matching equilibria will pick out the payoff dominant equilibrium, whereas the play that emerges from various evolutionary approaches will depend on the specifics of the random matching and more details of the payoffs.
The "cooperative" nature of the core-stability captures the idea that dissatisfied groups of agents can communicate and jointly deviate and that this has implications for behavior. If instead one tried to use extensive form games explicitly to model a general matching process where arbitrary groups of agents can communicate, coordinate and deviate at many points in the process, one would end up with very complex and potentially intractable extensive forms, difficulties in defining appropriate equilibrium concepts, and a sensitivity to protocol. For these reasons the core has been the dominant approach in the literature on matching and marriage markets.

2. The One-Time Matching Model

We first provide definitions for settings where the choice to match is only made once.

The starting point is a standard game in normal form, with player roles denoted by \( i \in \{1, \ldots, n\} \). We use the term "role" to indicate that there may be a number of players who can fill a given role in our setting.

Associated with player role \( i \) is a strategy set \( S_i \) and profiles of pure strategies which lie in \( S = S_1 \times \ldots \times S_n \), with generic elements \( s_i \) and \( s \), respectively. Except where otherwise stated, we take \( S \) to be a finite set, and let the set of mixed strategies for player role \( i \) be denoted by \( \Delta(S_i) \), the set of probability distributions on \( S_i \).

Player role \( i \) has a von Neumann-Morgenstern utility function \( u_i \), where \( u_i(s) \) is the payoff to a player in role \( i \) when \( s \) is the n-vector of strategies played. Let \( u_i(\mu) \) denote the expected utility corresponding to an n-vector of mixed strategies \( \mu \).

Given the above base game, we now introduce the full society of players. A standard game would just have one player for each player role. Here we allow a society to have many players to fill each role. The players will then be matched into different groups, such that each group has one player in each role.

In particular, there is a finite population, denoted \( P_i \), of players who can fill role \( i \). For instance, \( P_1 \) would be a set of women and \( P_2 \) would be a set of men in a society if the game to be played is the "Battle of the Sexes" game. As another example, the populations could be firms and workers. We take these populations to be disjoint, so that each player in the society can play in exactly one role. Let \( n_i \) be the cardinality of population \( P_i \), and label the player roles so that \( n_i \geq n_k \), whenever \( i > k \). \( N = UP_1 \) is the set of all players in the society, and \( P = P_1 \times \ldots \times P_n \) is the set of vectors of players such that there is one player from each role \( i \), with generic element \( p \). We use \( i,j \), and \( k \) to denote indices of different player roles and \( a,b \), and \( c \) to denote generic players.

---

4 In this definition, a player cares about the play of the other players with whom he or she is matched but not the other players' identities. In Jackson and Watts (2008) we consider situations where players can care about both.
A profile \((n; P_1, \ldots, P_n; S, u)\) is called a **social game**.

Given that \(S\) is finite, the existence of a (possibly mixed strategy) Nash equilibrium for any group of players is assured.\(^5\)

Given that we now have an underlying normal form game, and populations of players to fill each role, we then need to specify how players are matched and play. This comes in two parts. First, we need to specify how the players are matched into groups, and then we need to specify the strategy that each player in the society plays in his or her respective game.

A **matching** is a mapping \(f : N \rightarrow P \cup N\), such that for each \(i\) and \(c \in P_i\)
(i) either \(f(c) = p \in P\) such that \(c = p_i\) or \(f(c) = c\), and
(ii) if \(f(c) = p\) and \(b = p_j\), then \(f(b) = p\).

The set of all matching functions for a given social game is denoted \(MF(n; P_1, \ldots, P_n; S, u)\), although we simply write \(MF\) when the underlying social game is fixed.

This is a standard definition of matching with the interpretation that \(f(c)\) is the vector of players with whom \(c\) is matched. Item (i) states that either player \(c\) is matched in a group \(p\) or player \(c\) is unmatched (matched to him or herself). Item (ii) states that if a player \(c\) is matched in a group that includes player \(b\), then \(b\) has to be matched in that same group.

We normalize the payoff of an unmatched player to 0.

In order to specify a play of the game, we need to first specify strategies for every player in the whole society \(N\). Let \(m\) denote an \(|N|\)-dimensional vector that specifies a mixed strategy for each player in the society, with \(m_c\) denoting player \(c\)'s strategy, such that \(m_c \in \Delta(S_i)\) if player \(c\) is in player role \(i\). Let \(M\) denote the set of such profiles.

Given a profile of mixed strategies for all players in the society, \(m \in M\), and a matching function \(f\), let \(U_c(m, f)\) be the expected utility that player \(c\) receives if the matching \(f\) is in place and \(m\) is the profile of mixed strategies played. Thus, \(U_c(m, f) = u_i(m_p)\), if \(f(c) = p\) with \(c = p_i\) and \(U_c(m, f) = 0\) if \(f(c) = c\), where \(m_p\) denotes the mixed strategy profile of the players in the group \(p\).

A **matching equilibrium** is a mixed strategy profile \(m \in M\), and a matching function \(f\) in \(MF\) such that
(a) if \(f(c) = p \in P\) for some player \(c\), then \(m_p\) is a Nash equilibrium and \(U_c(m, f) \geq 0\); and
(b) there does not exist \(p \in P\), and a Nash equilibrium \(\mu\) such that \(u_i(\mu) > U_{P_i}(m, f)\) for all \(i\).

\(^5\)Our definitions and results extend to games such that each group of matched players has a nonempty and compact set of equilibria, and continuous payoffs across (mixed) strategies. And, although we provide the definitions for the case of a game in normal form, there are obvious analogs for the case where the game played is in extensive form.
Condition (a) requires that each matched group plays a Nash equilibrium, so that equilibrium holds within a group, and also that each player at least receives the payoff they could obtain by leaving the match and being unmatched. This requires that no player have a unilateral deviation from the play of the overall game which would make him or her better off. Condition (b) requires that no set of players, one from each role, could form a new group and play some Nash equilibrium and all receive strictly higher payoffs than what they get in their assigned matches. This embodies the potential for groups of players to break away from the existing matching, form a group among themselves and play a strategy that would make them each better off. We require that the play of the deviating group must be a Nash equilibrium, in order for this to be a credible alternative to the current match.

Discussion of the Matching Equilibrium Definition

Note that just as with many solution concepts, there are various definitions of blocking that one might use. In a sense, we place strong restrictions on equilibrium, as a deviating group can block or upset a matching if they can find any Nash equilibrium that makes them better off. In this way, our definition is a core definition, where a group of players can block a given allocation if they can find some alternative allocation that is feasible for them and improving for them. In our case, feasibility corresponds to their anticipated play after the deviation being a Nash equilibrium. As will become clear in what follows, the definition that we employ is the (only) one that is a proper extension of renegotiation-proofness when we move to repeated games.

One can also view our definition of matching equilibrium as an extension of the standard definition from the matching literature, simply augmented to deal with the fact that each matched group could have multiple potential payoffs. The payoffs in a matching equilibrium can be seen as those corresponding to the core of an NTU game where only coalitions with one player from each player role generate non-zero payoffs and the set of possible payoffs for any group of players (one in each role) is the set of all Nash equilibrium payoffs derived from the underlying game. So, the definition of a matching equilibrium combines the core of a matching game with that of Nash equilibrium to predict potential payoffs that can be feasibly generated by groups of players; thus drawing from both cooperative and non-cooperative game theory and marrying prominent solution concepts from each. As an alternative, one could attempt to model things entirely non-cooperatively. For example, one could model some extensive form process by which groups form or reform. However this would be cumbersome and there does not appear to be any natural extensive form that would capture all the possible rematching opportunities and still be tractable. Indeed, the matching process that we are modeling here through use of core-stability would be a difficult process to model non-cooperatively, as in most applications it is a very open-ended process where there is no prespecified order of moves. There are strong advantages to the approach we employ here: Nash equilibrium is very well-suited to analyzing stability and self-enforcing play.

---

6 We require a strict form of blocking which, similarly to many other core-based concepts, aids in existence.
within groups once the matching process is over, while the core enables modeling group formation in a protocol-free manner and is well suited to capturing coalitional incentives.

3. Analysis with One-Time Matching

To preview some of the basic ideas, we begin with a simple example that illustrates the selection of equilibrium imposed by the threat of rematching.

Example 1: Battle of the Sexes with Uneven Populations

There is one woman P₁ = {1} and two men P₂ = {2, 3}. The woman is in the row player role, while the men are in the column role. The payoffs to the players are described by the following matrix.

\[
\begin{array}{cc}
A & B \\
A & 1,3 & 0,0 \\
B & 0,0 & 3,1 \\
\end{array}
\]

There are three Nash equilibria to the game: the pure strategy equilibria (A,A) and (B,B), and a mixed strategy equilibrium where the man plays A with probability 3/4 and the woman plays A with probability 1/4.

There are two matching equilibria: one with a matching f(1)=f(2)=(1,2), and another with a matching f'(1)=f'(3)=(1,3). In both equilibria the matched couple plays (B,B).

The other two Nash equilibrium strategies are not part of any matching equilibrium, as for instance, under f where 1 and 2 are matched, if the intended play is not (B,B) then players 1 and 3 can deviate to match and play (B,B) and both be better off and expect their agreement to be self-enforcing. This rules out play of (A,A) where 1 is matched with 2. Thus, under the "core" nature of the matching equilibrium solution, it is implicit that 1 and 3 can communicate and play something self-enforcing that will strictly improve the payoffs for both.

Note that this example shows that in order to guarantee existence of equilibrium it is necessary that a deviation can only block a proposed matching equilibrium if the deviating players are all strictly better off. With a weaker notion of blocking, where only some of the deviating players need to strictly benefit, equilibrium would fail to exist in the above game.

3.1 Equilibrium Existence

We now show that the set of matching equilibria is nonempty and compact. Compactness is used to prove existence of repeated matching equilibria.
Theorem 1: The set of matching equilibria of a social game is nonempty and compact.

Proof of Theorem 1: Let us first show that the set of matching equilibria is nonempty. Order player roles so that $n_i \geq n_k$, whenever $i > k$. Let $NE_1$ be the set of mixed strategy Nash equilibria that reach maximal payoff for the player role 1, subject to all other player roles getting at least 0. Let $NE_2$ be the subset of $NE_1$ that maximize player role 2’s utility, subject to being in $NE_1$. Inductively, let $NE_k$ be the subset of $NE_{k-1}$ that maximize player role $k$’s utility, subject to being in $NE_{k-1}$. If $NE_n$ is empty, then match all players to themselves. Otherwise, select any matching $f$ such that players in $P_1$ are all matched and pick a mixed strategy profile in $NE_n$ and set $m$ so that each player in role $i$ plays role $i$’s component of that mixed strategy equilibrium. It follows easily that this forms a matching equilibrium.

Next, let us argue that the set of matching equilibria is compact. Given the finite set of possible matchings and the compact nature of the strategy spaces, we need only show that the set of strategy profiles that are part of a matching equilibrium for any fixed matching $f$ is closed. Let $m^f \rightarrow m$, where $(m^f, f)$ is a matching equilibrium for every $r$. It is immediate that (a) is satisfied by $m$. We need only verify that (b) is satisfied by $m$. Suppose to the contrary that there exists $p \in P$, and a profile of strategies $m'_p$ for the players in $p$ such that $u_i(m'_p) > U_c(m, f)$ for all $i$, where $c = p$, and such that $m'_p$ is a Nash equilibrium. Then for large enough $r$, it follows that $u_i(m'_p) > U_c(m^f, f)$ for all $i$, which is a contradiction.

3.2 Characterization of Equilibrium

We now offer a characterization of the set of matching equilibria.

Before stating the characterization theorem, let us first explore some of the intuitive properties of equilibria. The following simple example shows that in a matching equilibrium the play across groups is restricted, including in situations where the populations of different player roles are exactly balanced.

Example 2: Bargaining, the Nash Demand Game, Double Auctions, and Ultimatum Games.

There are two player roles corresponding to populations called “Buyers” and “Sellers”. A matched buyer and seller play a version of a Nash-demand game where they each announce a price in $[0,1]$ (so here we allow for an infinite pure strategy set). If the price of the buyer exceeds the price of the seller, then there is trade of a single unit of a good that has value 1 to the buyer and value 0 to the seller, and at the average of the two prices. So $S_1 = S_2 = [0,1]$ and the payoff to the buyer (say role 1) is $1 - (s_1 + s_2)/2$ if $s_1 \geq s_2$, and 0 otherwise; and the payoff to the seller (role 2) is $(s_1 + s_2)/2$ if $s_1 \geq s_2$, and 0 otherwise. There are a continuum of Pareto efficient pure strategy Nash equilibria to this game (where both
players say the same price), as well as a Pareto dominated equilibrium where the seller says 1 and the buyer 0.

If there are fewer buyers than sellers, so \( n_1 < n_2 \), then all matching equilibria have all buyers matched and each matched pair playing the equilibrium where \( s_1 = s_2 = 0 \). If there are more buyers than sellers, so \( n_1 > n_2 \), then all matching equilibria have all sellers matched and each matched pair playing the equilibrium where \( s_1 = s_2 = 1 \). This is reminiscent of Aumann's "gloves game," or other games with complementarities, where the rarer items end up with higher prices or payoffs. In situations where the populations are evenly matched so that \( n_1 = n_2 \), there are infinitely many matching equilibria. However, even then there are still restrictions on the matching and play (when in pure strategies): all players must be matched in a pair and every match must play the same strategy, announcing \( s_1 = s_2 = x \), where \( x \) is the same across matches. To see this, note that, for instance, if two different pairs traded at prices \( x' \) and \( x \) where \( x' > x \), then the buyer from the first group and the seller from the second group would both be strictly better off by rematching and trading at any price strictly between \( x' \) and \( x \).

While fairly simple, this example shows that relative population sizes affect the equilibrium play. It also shows how an "equal-treatment" property (that is commonly implied by core solutions in settings with replications) manifests itself here. With evenly matched populations equilibrium payoffs cannot differ by too much. The general implications of core stability on the selection of equilibrium within matchings is made precise in Theorem 2, which we now present.

Let \( PO(k) \) represent the set of \( k \)-vectors of utility \( (u'_1, u'_2, \ldots, u'_k) \) such that

- player roles \( \{1, \ldots, k\} \) receive a nonnegative payoff, and
- if \( \mu \) is a Nash equilibrium such that all player roles get a positive payoff, then there exists \( i \leq k \) such that \( u'_i \geq u_i(\mu) \).

Thus, \( PO(k) \) is the set of possible utility vectors for a set of the first \( k \) player roles, that could not be improved upon for all of them by some Nash equilibrium that leads to nonnegative payoffs for all players.

Given mixed strategy profile and matching, \( (m,f) \), of some social game, for each \( i \) let

\[
v_i(m,f) = \min_{c \in \mathbb{P}_i, f(c) \neq c} U_c(m,f).
\]

Thus, \( v_i \) is the minimum utility obtained under \( (m,f) \) by any player in role \( i \) who is matched under \( f \).

**Theorem 2:** Suppose that a social game has at least one Nash equilibrium such that all player roles have a positive payoff\(^7\) and let \( k \) be the smallest \( i \) such that \( n_{i+1} > n_i \), letting \( k = n \) if there is no such \( i \). Then \( (f,m) \) is a matching equilibrium if and only if

---

\(^7\) The "if" part of the statement still holds if there only exists a Nash equilibrium where all player roles get a nonnegative payoff (as opposed to positive). Clearly, if in every Nash equilibrium some player role gets a negative payoff then all matching equilibria have all players unmatched.
Conclusions (i) and (ii) are fairly clear: all players in the least populous roles must be matched, as otherwise there would be at least one full unmatched group that could deviate, match, and play an equilibrium that makes all of them better off. Conclusion (iii) implies that any equilibrium played by a group has to be Pareto optimal for the players in the least populous roles. It also says that if we take the worst-off player in each of least populous roles, the resulting profile of payoffs cannot be below some Nash equilibrium that provides positive payoffs for all players. This limits how much the payoffs can differ across players within the same player role. The idea behind this is relatively straightforward. If (iii) were not satisfied, then a group of players who got the lowest payoff in each player role (with that payoff being 0 for roles above k) could deviate to match together and play a Nash equilibrium that would be strictly improving for each player involved.

The properties illustrated in Examples 1 and 2, in the case of bipartite matchings then generalize as follows.

**Corollary 1 (Favoring of Less Populous Roles):** Consider a social game where population 1 is the smallest ($n_1 < n_k$ for all $k > 1$), and where at least one of player role 1’s most preferred Nash equilibria gives all player roles a positive payoff. In any matching equilibrium, all players of role 1 are matched and all groups of players play one of player role 1’s most preferred Nash equilibria.

**Corollary 2 (Nearly Equal Treatment):** Consider a bipartite setting such that no player is indifferent between any two Nash equilibria. If (m,f) is a matching equilibrium, then for any two players c and b in the same player role i, there does not exist any Pareto efficient Nash equilibrium $\mu$ such that $U_c(m,f) < u_i(\mu) < U_b(m,f)$.

Corollary 2 does not claim that all players in the same role must receive the same payoff (which is not true), but that the payoffs cannot be too dispersed, in that there cannot lie any Pareto efficient Nash equilibrium payoffs between any two players’ payoffs. Thus, the players in the same role are nearly equally treated under a matching equilibrium.

In the one-time matching setting, we see that the fact that players have choices over their partners affects the strategies that are played. Equilibrium play favors roles that have fewer players and does not allow the payoffs of the least populous roles to differ too much across players.

### 4. Finitely Repeated Social Games
While the one-time matching provides some insight into how the ability of players to choose partners can influence play, when players have the opportunity to change partners over time, it influences repeated play in interesting ways, by both selecting from repeated game equilibria, and introducing new ones.

### 4.1 Repeated Matching Equilibrium

Consider a sequence of social games played over the finite set of periods \{1,2,3,...,T\}. Players receive the discounted sum of payoffs of per period plays, with a discount rate of \(\delta\) in \([0,1]\). Rematchings are possible in any period.

Let \(h=[s_1^1,f_1^1,s_2^2,f_2^2;...s_t^t,f_t^t]\) denote a generic history of the game through some time \(t\), which includes a list of the strategies played and the matches that were in place in each period. Let \(H(t)\) denote the union of all histories of the game through time \(t\). Let \(H=\bigcup_0^{T-1} H(t)\) be the set of all finite histories that could have been observed in some period through the beginning of period \(T\). We adopt the convention that \(H(0)\) is a singleton (empty) history denoted by \(\emptyset\).

A \(T\)-period matching function is a mapping \(F:H \rightarrow MF\), which indicates the current period matching following any history \(h\) in \(H\). Let \(F(h,c)\) denote \(c\)’s match after history \(h\). The set of all \(T\)-period matching functions is denoted \(MF(T)\), and \(MF(1)=MF\).

A (behavioral) strategy for a player \(c\) in role \(i\) is a map \(\sigma_c:H \rightarrow \Delta(S_i)\). So a strategy for a player indicates which mixed strategy she plays following any finite history of length no more than \(T-1\). The \(|N|\)-dimensional behavioral strategy profiles for the \(T\) period game are denoted \(S(T)\). Let \(U_c(\sigma,F)\) denote player \(c\)’s discounted expected utility if the profile of strategies \(\sigma\) in \(S(T)\) is played and the matchings are governed by \(F\) in \(MF(T)\).

For any \(t\) and history \(h\) in \(H(t)\), let \((\sigma_h,F_h)\) denote the continuation behavioral strategies and repeated matching. \(^9\)

Repeated matching equilibria of a \(T\)-period game are defined inductively. In a period \(t\), the repeated match equilibria have the same structure as the original definition of matching equilibrium, except that the requirement of Nash equilibrium is replaced by an equilibrium continuation that is a perfect equilibrium anticipating play of a matching equilibrium in the remaining periods in every subgame.

More formally, the set of repeated matching equilibria of a \(T\)-period social game, denoted \(RME(T)\), is defined as follows.

\(^8\) Although we do not explicitly allow players to condition their play on their current match, the matching is tied down as a function of the history up through the last period by a matching function. So when strategies are coupled with a matching function, players implicitly know with whom they will play. Thus, given that we pair matching functions with strategies for all equilibrium definitions, it is irrelevant whether we allow players’ strategies to depend on their current match.

\(^9\) That is, \((\sigma_0,F_0)\) lies in \(S(T-1)\times MF(T-1)\) and is such that \((\sigma_0(h’),F_0(h’))=(\sigma(h,h’),F(h,h’))\) for any continuation history \(h’\) in \(\bigcup_0^{T-1} H(t’)\).
Let \( \text{RME}(1) \) be the matching equilibria of the 1 period social game.

Inductively, let \( \text{RME}(t) \) be the set of \((\sigma, F)\) in \(S(t) \times MF(t)\) such that

\( (i) \) \((\sigma_h, F_h)\) is in \( \text{RME}(t-1) \) for all \( h \) in \( H(1) \), and

\( (ii) \) no player wants to deviate from \( \sigma(\emptyset) \) given the current matching \( F(\emptyset) \) and anticipating the continuation governed by \((\sigma, F)\).\(^{10}\)

\( (iii) \) there does not exist any \( c \in N \) with \( 0 > U_c(\sigma, F) \), and

\( (iv) \) there does not exist any \( p \in P \) and \((\sigma', F')\) satisfying (i)-(iii) such that

\[ U_c(\sigma', F') > U_c(\sigma, F) \]

for all \( i \) and \( c = p_i \), and such that \( F'(h, c) = p \) or \( c \) for all \( h \) and for all \( c \) in \( p \).

The conditions are fairly straightforward. Parts (i) and (ii) are standard sequential rationality requirements requiring that the continuation strategies be matching equilibria of the continuation game and that no player wants to deviate from the prescribed strategies. Part (iii) is a simple requirement that no player be better off leaving the match altogether and remaining single. Part (iv) captures the implications of rematching: it requires that there not exist a group who could match together and find strategies that would make each of them strictly better off and such that their future matchings could only involve themselves. That is, in order for a deviation to be viable it must be that the deviating group does not need to count on being able to rematch with players outside of their group in order to implement their deviation.

### 4.2 Existence of Repeated Matching Equilibria

In the definition of repeated matching equilibria, as in the definition of matching equilibrium, only individual players or single matched groups are allowed to deviate. This assumption is made to ensure existence, we show below that if larger groups are allowed to deviate, then existence can fail.

**Theorem 3:** The set of repeated matching equilibrium is nonempty and compact for every finite \( t \). Moreover, there exists a repeated matching equilibrium where the repeated matching function is constant on the equilibrium path.

We remark that simply repeating a one period matching equilibrium is not always a repeated matching equilibrium, and in some cases, there is no repeated matching equilibrium where both matching and play is constant.

### 4.3 Relation of Repeated Matching Equilibria to Renegotiation

We now explore the structure of repeated matching equilibria. We begin by examining the relationship with renegotiation-proof equilibrium. We do this since if there were just

\(^{10}\) So, \( U_c(\sigma'_c, \sigma_c, F) \leq U_c(\sigma_c, F) \) for all players \( c \) and behavioral strategies \( \sigma'_c \) that differ from \( \sigma_c \) only in the first period, which by the induction implies that a player would not want to deviate to any other behavioral strategy that differs following other histories.
one player in each role, then the two definitions coincide, and so one might wonder whether more generally the repeated matching equilibria are some sort of generalization of renegotiation-proof equilibria.

In some special cases, there is a relationship between repeated matching equilibria and renegotiation-proof equilibria.

**Proposition 1:** Let all populations be of equal size and suppose that all renegotiation-proof equilibria are such that each player’s expected utility is non-negative in all subgames. Then every renegotiation-proof equilibrium is part of a repeated matching equilibrium with a constant matching function $f$ where all players are matched and where each group plays the same renegotiation-proof equilibrium.

Next we show that if populations are not evenly matched, then the set of repeated matching equilibria may no longer contain any renegotiation-proof equilibrium. This is because the least populous player roles will get to choose their favorite Nash equilibrium in the last period. This ties down play in the last period, which can change the set of possible plays in the second to last period, which then changes what is supportable earlier, and so forth. The changes in what is supportable occur in non set inclusive ways once we move back several periods.

**Example 3: All Repeated Matching Equilibria are neither renegotiation-proof nor Pareto Optimal**

This game is based on one in Benoit and Krishna [1993], extended to allow for populations of players. Let player 1 be in player role 1 (row) and players 2 and 3 be in player role 2 (column). The game is repeated twice, $\delta=.9$, and payoffs are

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td>0,0</td>
<td>1,3</td>
<td>0,0</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>3,1</td>
<td>0,0</td>
<td>6,0</td>
</tr>
</tbody>
</table>

This game has two pure strategy Nash equilibria: A,B and B,A, as well as a Pareto dominated mixed strategy equilibrium. There are three types of renegotiation-proof equilibria \(^{11}\) (as defined by Benoit and Krishna (1993)):

1. play B,A in period 1 and A,B in period 2, or vice versa.
2. play B,C in period 1 and play A,B in period 2. If player 2 deviates in period 1 then play B,A in period 2.
3. play A,B in period 1 and A,B in period 2.

All repeated matching equilibria have player 1 matched in each period and have the matched players play B,A in both periods. To see this notice that in period 2, all matching equilibria must play B,A (by Theorem 2). This eliminates (1), (2) and (3) as possible equilibria. Player 1 would like to play equilibrium (2) and see play of B,C in

\(^{11}\) Note that Nash equilibria and renegotiation-proof equilibria are defined as usual on a player set of size $n$ while repeated matching equilibria are defined on the overall population (larger than $n$).
period 1. However, his partner will always have incentive to deviate in period 1 (gaining a payoff of 1, while losing at most .9 in the second period, as only B,A could be played in the second period). It is now easy to see that B,A must also be played in the first period. At these repeated matching equilibria, the expected payoffs are (5.7, 1.9) to the matched players if the same players are matched in both periods. This is Pareto dominated by (6.9, 2.7) which are the expected payoffs from renegotiation-proof equilibrium 2.

Thus, in this example none of the repeated matching equilibria are renegotiation-proof, and any of them that have constant matchings are Pareto dominated by a renegotiation-proof equilibrium. The example shows how having players choose partners over time can limit play within a period, which then can lead to ramifications in repetition, completely changing what is possible.

**4.4 Group-Stable Repeated Matching Equilibrium**

Next, to develop a further understanding of equilibrium structure, we strengthen the definition of repeated matching equilibria to that of group-stable repeated matching equilibria, which allows blocking by larger groups than just one from each role. In repeated settings, this can be important, as equilibrium continuations to support play by some group can require threats of rematching in future periods. Thus, to support a deviation, a group might need several complete sets of player roles.

The set of **group-stable repeated matching equilibria** of a T-period social game, denoted GSRME(T), is defined inductively as follows.

Let GSRME(1) be the matching equilibria of the 1 period game.

Inductively, let GSRME(t) be the set of \((\sigma,F)\) in \(S(t) \times MF(t)\) such that

(i) \((\sigma_h,F_h)\) in GSRME(t-1) for all \(h\) in \(H(1)\), and

(ii) no player wants to deviate from \(\sigma(\emptyset)\) given the current matching \(F(\emptyset)\) and anticipating the continuation governed by \((\sigma,F)\).

(iii) there does not exist any \(c \in N\) with \(0 > U_c(\sigma,F)\), and

(iv) there does not exist any \(S \subseteq N\) and \((\sigma',F')\) satisfying (i)-(iii) such that \(U_c(\sigma',F') > U_c(\sigma,F)\) for all \(c\) in \(S\), and such that \(F'(h,c) \in S\) for all \(h\) and for all \(c \in S\).

This is analogous to the definition of repeated matching equilibrium, except that the deviating group is now allowed to include more than just one player from each role. This allows a deviating group to implement more complicated deviations, which could potentially be sustained by intricate rematching among themselves over time.

The following provides an example where all group-stable repeated matching equilibria differ completely from all subgame-perfect equilibria.

**Example 4: All Group-Stable Repeated Matching Equilibria differ from Subgame Perfect Equilibria (without matching)**
There are two periods and a discount factor of 1.

There are three player roles and nine players \{1,...,9\}. Players 1,4,7 are in role 1, players 2,5,8 are in role 2, and players 3,6,9 are in role 3.

Player role 1 is the row player, player role 2 is the column player, and player role 3 choose among the matrices. Payoffs are as follows.

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>center</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>up</td>
<td>5,1,3</td>
<td>-, -, -</td>
<td>-, -, -</td>
</tr>
<tr>
<td>middle</td>
<td>-,-,9</td>
<td>-26,-</td>
<td>25,25,25</td>
</tr>
<tr>
<td>down</td>
<td>-,-,-</td>
<td>6,8,10</td>
<td>28,9,-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>center</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>up</td>
<td>21,-,-</td>
<td>-13,16</td>
<td>8,10,6</td>
</tr>
<tr>
<td>middle</td>
<td>-,-,-</td>
<td>3,5,1</td>
<td>-,-,-</td>
</tr>
<tr>
<td>down</td>
<td>20,20,20</td>
<td>-,-,-</td>
<td>-9,-,-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>center</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td>up</td>
<td>-18,-</td>
<td>15,15,15</td>
<td>-,-,-</td>
</tr>
<tr>
<td>middle</td>
<td>10,6,8</td>
<td>-,-,-</td>
<td>-,-,-</td>
</tr>
<tr>
<td>down</td>
<td>13,-,23</td>
<td>-,-,13</td>
<td>1,3,5</td>
</tr>
</tbody>
</table>

The entries with a ‘-‘ have a largely negative payoff.

There are three pure strategy equilibria: (up, left, I); (middle, center, II); and (down, right, III). There are also mixed strategy equilibria, which for negative enough payoffs in the ‘-‘ entries results in largely negative payoffs for at least one player.

In a one-period setting, any matching equilibrium must have all players matched, and each group playing one of the three pure strategy Nash equilibria.

Next we show that there is a two-period matching equilibrium.

The first period matching is \(\{1,2,3\}, \{4,5,6\}, \{7,8,9\}\). In the first period, the first group plays (middle, right, I), the second group plays (down, left, II), and the third group plays (up, center, III).

If there are no deviations in the first period, then the second period matching is \(\{1,5,9\}, \{4,8,3\}, \{7,2,6\}\), with corresponding plays of (up, left, I), (middle, center, II), and (down, right, III), respectively.
If there is a deviation by some player \( k \) in the first period, then in the second period \( k \) is matched into some group that plays the Nash equilibrium that gives \( k \) a payoff of 1 (the particular matching is irrelevant). If there are deviations by more than one player in period 1, then some arbitrary one-period matching equilibrium is played in period 2.\(^{12}\)

The overall payoffs to the nine players are \((30,28,26,23,21,25,16,20,18)\), in order. Let us check that this is a group-stable repeated matching equilibrium. It is clearly a matching equilibrium for any history in the second period. So let us check the first period. First we check that no player wants to deviate from their prescribed strategy. Player 1 could benefit by deviating from middle to down in the first period, for a gain of 3. However, then in the second period instead of receiving a payoff of 5, player 1 would get a payoff of 1, so overall the payoff change would be \(+3-4=-1\). Similar calculations for each player verify that no player could benefit from a unilateral deviation.

Next, let us check that no group of players could benefit by reorganizing themselves. The key to this is that in each group in the first period there is some player who has a "high" potential payoff to a deviation (a gain of 3), another player who has a medium gain from a deviation (a gain of 1) and the last player who has no gain. Given that second period payoffs are vectors of the sort \(5,3,1\) to respective players and each player gets at least 1, the available deterrents to deviations are payoff changes from 5 to 1 (a loss of 4) and from 3 to 1 (a loss of 2). Note that the first player in a group who plays \((25,25,25)\) must get a 5 payoff in the second period. The second player must get at least a 3 payoff in the second period. This already requires two different groups in the second period. It is then easy to see that in order to sustain more than one group getting payoffs of \((25,25,25)\) in the first period would require at least four groups in the second period. Careful checking along these lines shows that getting \((25,25,25)\); \((20,20,20)\), and \((15,15,15)\) is the best that one can achieve in terms of first period payoffs. While there are still some details to check, this is the heart of verifying that this a two-period group-stable repeated matching equilibrium.

Lastly we show that this group-stable repeated matching equilibrium differs completely from all subgame-perfect equilibria of the game without matching. First notice that \((25,25,25)\) cannot be achieved in the first period at any subgame-perfect equilibria as it is impossible to keep both players 1 and 2 from cheating. Similarly \((20,20,20)\) and \((15,15,15)\) can not be achieved in the first period either. In fact the best subgame perfect equilibrium for player 1 would be to achieve \((10,6,8)\) in period 1 and \((5,1,3)\) in period 2. Thus the maximum payoff a player can receive at a subgame-perfect equilibrium is 15 which is less than 16 which is the minimum payoff a player receives at the above

\(^{12}\) Notice that in our definition of repeated matching equilibria, both new matchings and new action vectors can be conditioned on the past play of those with whom one was grouped with and not grouped with in the past. In this example, even though player 1 was not matched with 5 and 9 in the first period, in the second period player 1 knew whether or not 5 and 9 had deviated previously. If instead players did not have information regarding stranger’s past play, then if a player chose to rematch with strangers he could not condition play based on how the strangers had behaved in the past, which is similar to the assumptions of Ghosh and Ray (1996). Such an assumption would most likely force players not to regroup, if their partners had not deviated in the past. For further discussion of repeated matching equilibria with constant matching functions and their relationship to renegotiation-proof equilibrium, see Proposition 1.
repeated matching equilibrium. Since this group-stable repeated matching equilibrium Pareto dominates all subgame perfect equilibrium we know that no subgame-perfect equilibrium can be a group-stable repeated matching equilibrium. ◊

While the possibility of rematching can lead to new equilibria, group-stable matching equilibria face existence problems, even with evenly matched populations. This is illustrated in the following example.

**Example 4 continued: Nonexistence of Group-Stable Repeated Matching Equilibrium**

Add three extra players \{10,11,12\} to Example 4 where player 10 is in role 1, 11 in role 2 and 12 in role 3. Any equilibrium where at least two groups are not getting payoffs above \((15,15,15)\) in the first period will be blocked by some set of three groups deviating to play the three-group equilibrium of the two period game. So this suggests that we try to have some three groups play the three group matching equilibrium, and then one group left to play alone. The best the remaining group could get would be \((10,6,8)\) in the first period and \((5,1,3)\) in the second period - or some permutation of that (sustained by threats of rematching in the second period if a player deviates in the first period). However, in that case, the second, third and fourth groups could deviate (as S in (b) of the definition) to play the three-group equilibrium and all be made better off. Any permutation of players will still have such a deviation. This leaves us only with possibilities of trying to sustain, say, two groups getting \((25,25,25)\) in the first period, with two other groups getting lower payoffs. However, this would require the latter two groups getting payoffs of \((1,1,5)\) in the second period. That will not allow those two groups to sustain anything but a Nash equilibrium play in the first period. Either of those groups could deviate alone to improve by earning a permutation of \((10,6,8)+(5,1,3)\). Similar reasoning rules out any other attempt to sustain at least two groups getting payoffs of at least \((15,15,15)\) in the first period. Thus there is no group-stable repeated matching equilibrium to this game. ◊

**5. An Application to Centipede/Trust Games**

The set of repeated matching equilibria can differ dramatically from various sorts of equilibria in settings with fixed players, as we have seen. In order to get a better feeling for how this can work, we examine how repeated matching equilibria can be used to make new predictions in a well-studied class of games.

Consider the following general variation on a simple finitely repeated centipede/trust game where \(d>b>a>c>0\) and let \(\delta=1\).

**Extensive form:**

```
Player role 1
| Player role 2 |
---|---|
\(a,a\)| \(c,d\)

\(b,b\)
```

**Normal form:**

```
| Down | Across |
---|---|
Down | \(a,a\) | \(a,a\) |
Across | \(c,d\) | \(b,b\) |
```
There is a unique pure strategy Nash equilibrium to this game, which is for both players to play down. There are many mixed strategy equilibria, but they all have player role 1 playing down, and player role 2 playing down with probability at least \((b-a)/(b-c)\), and thus all Nash equilibria lead to the \(a,a\) outcome, even with repetition.

Here, it turns out that the set of repeated matching equilibria depends critically on the relative sizes of the populations of players in each role. If player role 1 is less numerous, then there is an equilibrium play that has all matched pairs playing (across, across) in all but the last period of the game, when it changes to (down, down). In contrast, if player role 2 is less numerous, then there is an equilibrium play has all matched pairs playing (across, down) until the last period of the game when it changes to (down, down).

In both cases we see players in role 1 exhibiting ``trust''. However, it happens for very different reasons. When role 1 players are less numerous, then they play across because they know they can threaten the second player with ostracism if the second player fails to play across. Here they are trusting because they are sure that they can enforce the other player’s behavior. When the players in role 1 are more numerous, then they are playing across because they are threatened with being ostracized themselves if they fail to do so. Now they play as if they are ``trusting,” even though they know that they will be taken advantage of; but because the alternative of being ostracized is worse. In cases where the two populations are evenly matched, all equilibria lead to down, down in all periods, as there are no credible threats of ostracism. Thus, in situations where players have the ability of rematching, our model provides some new predictions about when and how ``trusting” behavior should be exhibited.13

We now state more formal results outlining the structure of equilibria. Let us begin with the case where role 2 players are less numerous.

**Proposition 2:** Let \(n_1>n_2\) and \(T\geq 2\). There exists a GSRME (and RME) where all players in role 2 are matched in every period and where all matched players play (across, down) in all periods except for the last period.

Proposition 2 shows that playing (across,down) until the last period is always part of an equilibrium. What other equilibria exist depends on the relative sizes of the populations and the number of periods. To see this, note the following. First, it is clear that in a one period setting, the only equilibria are to play (down,down). Next, let us consider \(T=2\). As the proposition states, we can now have (across, down) in the first period and (down,down) in the second period as part of equilibrium play. Is it possible to have (across, across) in the first period and then (down,down) in the second? No, as by deviating a player 2 would strictly benefit. However, it can be checked that is still

---

13 We assume that there is no cost to switching partners. If there is a positive cost to switching partners, then a player will be unwilling to sever his current tie in order to ostracize a cheating partner, and so (down,down) will be played in all periods. However, in many situations (such as marriage, or business partnerships) there exists both a positive cost to switching partners and a positive cost to staying with a partner who has cheated previously. As long as the cost of switching partners is not larger than the cost of staying with a cheating partner, then our results will continue to hold true.
possible to play (down, down) in both periods as part of an equilibrium. Now, let us consider $T=3$. It is now possible to sustain play of (across, across) then (across, down) then (down, down). This is sustained as follows. If a player 1 deviates, the player is ostracized and the player 2 rematches. If a player 2 deviates in the first period, then the subsequent equilibrium is (down, down) in both remaining periods. Is it possible to play (down, down) in all three periods as part of an equilibrium? This now depends on the population sizes. If $n_1$ is sufficiently large (at least $3n_2 + 2$), then there must be at least two player 1’s who are unmatched for all three periods. A player 2 could deviate with two player 1’s and sustain an equilibrium where (across, down) is played in the first period and then (down, down) in the next two periods and each of the two player 1’s gets to play in one of the last two periods. As $T$ increases, depending on relative population sizes, there can be more complicated equilibria introduced that vary across $T$’s.

Next, let us examine the case where role 1 players are less numerous. Here the equilibria have a different structure, and, for some payoffs, we can rule out always playing (down, down) even without large disparities in population size.

Let $[b]_+$ equal $b$ if $b$ is an integer otherwise let $[b]_+$ represent the greatest integer not exceeding $(b+1)$.

**Proposition 3:** Let $n_1 < n_2$.

If $T \geq \lfloor (d-b)/a \rfloor + 1$, then there exists a group-stable repeated matching equilibrium (and RME) where matched pairs of players play (across, across) for the first $T-\lfloor (d-b)/a \rfloor$ periods.

If $T \geq 2\lfloor (d-b)/a \rfloor + 2$ and $b > a(1+T/(T-2\lfloor (d-b)/a \rfloor)-1))$, then there is no group-stable repeated matching equilibrium where players play (down, down) in each period.

This proposition also shows the difference between repeated matching equilibrium and group-stable repeated matching equilibrium. While the above is a repeated matching equilibrium, there is also a repeated matching equilibrium with a constant matching function where all players play down in all periods (even if $b > a(1+T/(T-2\lfloor (d-b)/a \rfloor)-1))$, as this cannot be improved upon by a group of only two players.

This points to a further difference between the two definitions. Note that it is clear that no two GSRME can be strictly Pareto ranked, at any time. This follows directly from the fact that otherwise the grand coalition could block one equilibrium with the other in the first period. In contrast, the RME can be strictly Pareto ranked (even as the number of periods goes to infinity). If $b > a(1+T/(T-2\lfloor (d-b)/a \rfloor)-1))$ then at the above RME, all matched players receive an average per period payoff strictly greater than $a$ (even as the number of periods goes to infinity), while at the RME where all matched pairs play down in every period, the average per period payoff of matched players equals $a$.

---

14 This property echoes average payoff properties of renegotiation proof equilibria, as for instance, found in Benoit and Krishna (1993).
Proposition 4: If \( n_1 = n_2 \) then in every group-stable repeated matching equilibrium (and every RME) all players are matched every period and all matched pairs play (down, down) in every period.

Let us remark on our modeling approach here. Note that in a three player centipede game with one player in role 1 and two in role 2, it is possible to get the same outcome as a subgame-perfect equilibrium of a game where in the second stage player 1 chooses whether to play with player 2 or 3. Thus, one might presume that a repeated matching equilibrium can generally be modeled as a subgame-perfect equilibrium in an extensive form game where players choose their partners at different stages. However, this presumption is not correct as more generally there may be many players who are contemporaneously choosing with whom to be matched, and trying to model such a protocol in an extensive form without introducing many spurious equilibria would be enormously cumbersome, if not impossible. A core based concept cuts right through such difficulties.

6. Discussion

We have defined and analyzed a new class of games called social games where players not only choose strategies but also choose with whom they play. This imposes new refinements on equilibrium play, where play depends on the relative populations of players in different roles, among other things. In finitely repeated settings, where players may choose to rematch in any period, we also find some interesting new aspects imposed by the threat of rematching. On the one hand, the threat of rematching can sustain new equilibria, sometimes with higher payoffs than without matching. On the other hand, the threat of rematching can limit the equilibria played within a period to those most beneficial to players in least populous roles, and this limits the types of threats that are available to sustain equilibria in repeated settings.

While we have analyzed both the existence and the structure of the equilibria in various settings, there is much more to be learned about such equilibria and their characteristics in various settings. Further studies include future study of the setting we have examined, as well as extensions and variations on that setting. For instance, one could introduce additional heterogeneity in allowing strategy spaces to be player specific, or by allowing for externalities across groups. We now briefly mention some extensions of the analysis that deserve future investigation, but are beyond the scope of this paper.

6.1 Symmetric Games with a Single Population of Players

The first extension is to settings where there is just one population of players who can play in any role. A natural (but certainly not the only) setting in which to consider this is

---

15 Note that our model does have some externalities in that if groups form and some people are excluded then those excluded suffer. However, one group does not care about how other groups play while in some applications such as IO games one group may care about how the other plays. This would be an interesting avenue for future research.
where the game is symmetric. Here, any set of n players may be grouped together and all players are ex ante identical. The definition of matching equilibrium extends in the obvious way.\textsuperscript{16}

The following example shows that once we are in a world with a single population of players, existence of a matching equilibrium is no longer guaranteed, even in symmetric two-player single-period games.

\textbf{Example 5. Nonexistence in a single-population, symmetric, one-period social game}

There are three players and a game involves two players, and has payoffs

\[
\begin{array}{cc}
\text{A} & \text{B} \\
\text{A} & 0,0 & 1,2 \\
\text{B} & 2,1 & 0,0
\end{array}
\]

There does not exist any matching equilibrium.\textsuperscript{17} Any equilibrium would necessarily have two players matched and play either (A,B) or (B,A) as the mixed strategy equilibrium is strictly Pareto dominated by either pure strategy equilibrium. However, in any such matching the player getting the lower payoff can deviate together with the unmatched player and both be made better off (by playing the equilibrium that is less favorable to the formerly unmatched player and more favorable to the previously matched player).

This example shows that it is important for existence in single population social games that there exist a matching that includes all players. In situations where there does exist a matching which includes all players, then there does exist an equilibrium, as described in the following proposition.

\textbf{Proposition 5} Consider a single population social game with identical players. If there exists a matching that includes all players, then there exists a matching equilibrium. Moreover:

\begin{itemize}
  \item If every Nash equilibrium yields a negative payoff for at least one player, then all matching equilibria have all players unmatched.
  \item At most n groups have different minimum payoffs, and in fact at most n-1 players get a payoff that is less than the maximal minimum payoff among Nash equilibria.
\end{itemize}

The proof is straightforward and left to the reader.

\textsuperscript{16} Simply let P in the definition be the set of all vectors of n players.

\textsuperscript{17} This example is similar to Gale and Shapley’s (1962) roommate example of nonexistence of stable matchings, where there is always one person matched with a roommate that he or she finds less attractive than the unmatched person. However, here there is no a priori cycle in preferences that leads to instability, but instead it comes from the asymmetry in Nash equilibrium payoffs.
Note that the proposition shows that such a matching equilibrium will have “most” groups playing a symmetric Nash equilibrium if one exists that is not Pareto dominated by another Nash equilibrium.

In settings where all players come from one population and each can play in any role we find that beyond differences in existence, we also see differences in play from what we would see if players could not fill any role.

**Example 6. Two-Period Prisoners’ Dilemma**

Let there be three identical players who play the following Prisoner’s Dilemma game which is repeated twice. Let the discount rate $\delta=1$.

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3,3</td>
<td>-2,4</td>
</tr>
<tr>
<td>D</td>
<td>4,-2</td>
<td>2,2</td>
</tr>
</tbody>
</table>

There exists a group-stable repeated matching equilibrium where two players are matched and where the matched pair plays C,C in period 1 and D,D in period 2. (Since this game has a unique Nash equilibrium, the matched pair must play D,D in period 2.) If one of the matched players deviates in period 1 then the other matched player will sever this tie and will link with the unmatched player in period 2. So if a player deviates in period 1 he receives an expected payoff of 4 while if he does not deviate he receives an expected payoff of 5. Notice that the unmatched player cannot offer either matched player a credible better deal in period 1. Thus cooperation is sustained in the first period, which differs from any other equilibrium concept - including repeated matching equilibrium where players have set roles that they can fill.

**6.2 Other Equilibrium Definitions**

Our definition of matching equilibrium has examined Nash equilibrium within a single period version of the game. We could also examine other alternative definitions of equilibrium, within the purely static setting. There are of course, many equilibrium notions to choose from both in terms of analyzing how a given group would play within a game, as well as taking into account rematching. One can make arguments for looking at Strong Nash equilibrium, rather than Nash equilibrium within the play of a group, given that we are allowing for coalitional deviations. One could also use variations, where deviating groups should be immune to further subdeviations, as in Coalition Proof Nash Equilibrium. In addition, one might look at farsighted notions, where players anticipate further reactions by other players.

Alternatively, one could replace the notion of Nash equilibrium with that of correlated equilibrium. This switch would have the advantage of eliminating some matching equilibrium where all players end up unmatched. For example, consider a modified version of the battle of the sexes where there is one player in role 1 and one player in role 2 and where payoffs are
There are 3 Nash equilibrium of this game generating payoffs of \((3,-1), (-1,3),\) and \((-0.5,-0.5)\). Since all Nash involve a negative payoff for at least one player, at the unique matching equilibrium all players are unmatched. If instead we replaced the notion of Nash equilibrium with that of correlated equilibrium, then the two players could agree to match and then to toss a coin and play the first Nash if heads comes up and the second if tails comes up. This agreement would generate a positive expected payoff for each player, and so each player would a priori be better off than they would be if they were unmatched. However, after the coin is tossed at least one player would like to sever the tie and remain unmatched. To prevent such a severing one would need to use a simple contract to prevent breaking a match that would involve paying a fine of at least 1 for such a break. However, there are two possible obstacles with this method. First, there are many social and economic matching examples (such as a worker and firm with the worker’s strategies shirk or not) where players may not agree to leave the equilibrium choice up to a lottery. Second, there are some examples of social matches where writing and/or enforcing the necessary contract to prevent severing a match is unlikely to occur. For instance, in a battle of the sexes game where players are deciding what to do for the evening by flipping a coin it is unlikely that a punishment contract would be written and/or enforced. Alternatively, consider a game where a new graduate student is offered an academic job and accepts the job and then later discovers that this job will result in a negative utility for him. In many academic disciplines (such as sociology and geography), the graduate student can renege on the contract with little or no consequence.

Given the many options available in terms of defining an equilibrium, we have started with what we feel is a very natural definition, and leave variations for future research.

### 6.3 Money Transfers

We could also analyze situations where in addition to choosing actions and partners, players could offer to make transfers to one another. As in the case of correlated equilibrium discussed above, money transfers would be important for economic applications where contracting is possible and would require new definitions. For instance, in the one-period model we could redefine the game so that first players agree to a match, and then they agree to a specific Nash equilibrium and a set of transfer payments that would depend on whether or not the agreed upon Nash was played (thus players would need to agree upon transfer payments if the correct Nash was played and on payments if something else was played.) A binding contract would be needed to enforce the agreed upon transfer payments. At a transfer matching equilibrium no group would prefer to change partners, Nash equilibria, or transfer payments and no individual would prefer to sever all ties.
In the one-period model, groups would always play the Nash equilibrium with the largest aggregate payoff. The game played would then be quite similar to the bargaining game of example 2, where in a two player role setting transfer payments would be made so that the minority group would end up with all the surplus, and if populations were evenly matched, then all matched groups would agree to the same set of transfer payments. Results similar to Theorem 2 and the corresponding corollaries would continue to hold.

In the repeated setting, one could redefine transfer repeated matching equilibrium and transfer group stable repeated matching equilibrium in the obvious manner. Then transfer payments could be used to support non subgame perfect equilibrium play. For instance, consider the centipede trust game of section 4.1 with one player in role 1 and one player in role 2 and let the game be repeated twice. The following would be a transfer repeated matching equilibrium. Let players play across, across in period 1, followed by down, down in period 2. However, if player 2 cheats in period 1, then he must make a transfer payment to player 1 in period 2 equal to at least (d-b). Thus player 2 can be punished with transfer payments instead of threatened with ostracism. Additionally, one could use a similar transfer payment system to prevent cheating to eliminate the non-existence of a group-stable repeated matching equilibrium in Example 4. Here we would let groups of 3 play the strategies corresponding to payoffs of (25,25,25) in period 1, and any of the three Nash in period 2 with transfer payments in period 2 so that period 2 payoffs end up as (4,3,2). If any one cheats in period 1, then the transfer payments could be adjusted so that the cheater(s) receives a 0 payoff in period 2.

However, a contract system that would enforce such transfer payments would have to be quite powerful. Additionally, the contract system would have to allow transfer payments to be contingent upon what the players agreed upon up front. However, if such a contract system exists, then players would just agree to play whichever strategies would maximize joint payoffs in every period (even if such strategies were not part of a Nash equilibrium). Transfer payments held up by such a powerful contract system would then be used to enforce these maximum payoff strategies and the game played would be quite trivial as the same maximum payoff strategies would be played every period.

References


Appendix - Proofs and Examples Omitted from the Text.

Note that in Theorem 1, the set of matching equilibria is compact. However, the matching equilibrium correspondence (as payoffs are varied) is not upper hemicontinuous. This is in contrast with Nash equilibrium, and so the matching leads to changes in equilibrium structure, just as many equilibrium refinements (e.g., undominated Nash equilibrium, trembling hand perfect equilibrium, Pareto undominated Nash equilibrium, ...) do. The failure of upper hemicontinuity is illustrated in the following example.
Example A1: Failure of upper hemi-continuity

There are two players and two player roles. Player 1 (the row player) has only one strategy, while player 2 (the column player) has two strategies: \{left, right\}. The payoffs are as follows.

<table>
<thead>
<tr>
<th></th>
<th>left</th>
<th>right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1, 1</td>
<td>-1, 1+1/k</td>
</tr>
</tbody>
</table>

For any k>0, there is a unique matching equilibrium which is to have both players remain single, as the only Nash equilibrium is “right” which gives player 1 a negative payoff. In the limit where k is infinite, “left” is a Nash equilibrium, and there is a matching equilibrium where both players are matched and get payoffs of 1. This means that both players remaining single is no longer a matching equilibrium in the limit, as then (b) is violated.

Proof of Proposition 1: If the player roles are of the same size, then all players are matched. So consider the case where there are fewer players in role 1. Again, call the players in role 1 men and the players in role 2 women. First, note that in any matching equilibrium all men must be matched (since n1<n2, and otherwise an unmatched man and woman can improve their situation by matching) and all matches must play a positive payoff equilibrium. Second, suppose to the contrary of the proposition that there exists a matching equilibrium (call it ME1) where some woman, say W2, is unmatched while some other woman, say W4, is matched and that there exists another matching equilibrium (call it ME2) where W2 is matched and W4 is unmatched. In order for ME1 to be a matching equilibrium it must be that the man W2 is matched with at ME2, say M1, strictly prefers (this preference will be strict since we have assumed no indifference) his ME1 spouse/Nash to playing the ME2 Nash with W2 (otherwise at ME1, W2 and M1 will prefer to link and play their ME2 Nash). Similarly, in order for ME2 to be a matching equilibrium it must be that the woman M1 is matched with at ME1, say W5, strictly prefers her ME2 spouse/Nash to her ME1 spouse/Nash (otherwise at ME2, W5 and M1 will prefer to link and play their ME1 Nash). In order for ME1 to be a matching equilibrium it must be that the man W5 is matched with at ME2 strictly prefers his ME1 spouse/Nash to his ME2 spouse/Nash. If we keep repeating this process we will end up with all women who are matched at ME1 must strictly prefer their ME2 spouse/Nash. However this is not possible. To see this recall that all men must be matched at every matching equilibrium. Thus if there are n1 men then there must be n1 women who are matched at ME1 and who strictly prefer their ME2 spouse/Nash. However, since W2 is unmatched at ME1 but matched at ME2 and since we assumed no indifference, it must be that W2 also strictly prefers her ME2 spouse/Nash, thus there are (n1+1) women who strictly prefer the ME2 equilibrium. Since only n2 woman are matched at ME2 this is not possible. Thus it must be that the set of women who are unmatched is the same at both equilibria. ◊
Proof of Theorem 2: Let us first show the ’’if’’ part. By (ii) \( m_p \) is a Nash equilibrium for any matched group \( p \) under \( f \) and payoffs are nonnegative and so (a) of matching equilibrium is satisfied. (b) is satisfied since by (iii) any deviating group \( p \) with deviating strategy \( m'_p \) must have some player role \( i \leq k \) for whom \( u_i(f,m) \geq v_i(f,m) \geq u_i(m'_p) \) (by the definition of PO(k) and since all players in roles 1 to k are matched).

Next, let us show the converse. If not all players in roles 1 to k are matched, then there must be some complete group \( p \) of players who are not matched. Consider any equilibrium \( m'_p \) that gives positive payoffs to all players (and such a Nash equilibrium exists by the assumption of the theorem). We then contradict (b) in the definition of matching equilibrium, as all players in \( p \) can strictly benefit by forming a group playing \( m'_p \). Thus, all players in \( P_1 \) to \( P_k \) are matched and so (i) holds. Next, note that (ii) follows directly from (a) in the definition of matching equilibrium. Finally, let us show (iii). Suppose to the contrary that \( (v_1(f,m), v_2(f,m), \ldots, v_k(f,m)) \notin PO(k) \). Consider a group \( p \) consisting of a player from each player role 1 to k who is obtaining \( v_i(f,m) \). Since that \( (v_1(f,m), v_2(f,m), \ldots, v_k(f,m)) \notin PO(k) \), along with players in roles above k that were unmatched under \( f \) (if k<n). It follows from the definition of PO(k) that there exists \( m'_p \in NE \) that gives all players in \( p \) a positive payoff and makes all the players in the roles 1 to k in \( p \) strictly better off than under \( (f,m) \), and players in roles above k better off than being unmatched. This contradicts (b), and so the supposition was incorrect, implying (iii).

Proof of Theorem 3: We prove the theorem by induction. Recall that player roles are ordered so that \( n_i \geq n_k \) whenever \( i > k \).

First note that the set of possible histories is finite, and so strategies can be represented as a finite list of vectors, where each vector represents a mixed strategy to be played following a given history and thus belongs to a simplex. The set of repeated matching functions is finite.

Theorem 1 established existence and compactness (and constant matching) for \( t=1 \). Suppose that the claim has been established for all \( t<T \), we show that it is true for \( T \). Let us consider the case where there is some matching equilibrium for the one-period game that has players matched (so there is at least one Nash equilibrium that gives all players a nonnegative payoff), as the other case is obvious.

Let RPE(T) be \((\sigma,F)\) satisfying (i) and (ii). We first show that RPE(T) is nonempty and has an element with a constant matching function. Let \( f \) be a matching that has some equilibrium in RME(T-1) for which the matching is constant and equal to \( f \). Let \( F(\emptyset)=f \). \( H(1) \) is a finite set of possible histories that can occur in the first period. Associate with each \( h \) in \( H(1) \) the same continuation matching equilibrium in RME(T-1) that has the constant matching \( f \). This then defines \((\sigma,F)\) except for \( \sigma(\emptyset) \). Given that continuation matching is independent of play, pick any Nash equilibrium that gives all players nonnegative payoffs for \( \sigma(\emptyset) \). Specifying this as \( \sigma(\emptyset) \) defines a \((\sigma,F)\) in RPE(T), since (i) is satisfied by construction, and (ii) is satisfied since \( \sigma(\emptyset) \) is a Nash equilibrium that
gives all players nonnegative payoffs and the continuation is independent of play in the first period. Thus, RPE(T) is nonempty and has at least one element with a constant matching function.

Next, let us argue that RPE(T) is compact. Let \((\sigma^k,F^k) \rightarrow (\sigma,F)\), where \((\sigma^k,F^k)\) is in RPE(T) for each \(k\). By the compactness of RME(T-1), it follows that (i) is satisfied. To see (ii), note that by the finiteness of the number of repeated matching functions, we can restrict attention to the case where \(F^k=F\) for each \(k\). (ii) then follows, since any improving deviation from \(\sigma\) would also be an improving deviation from \(\sigma^k\) for large enough \(k\).

We now argue that RME(T) is nonempty and compact.

We first argue that RME(T) is nonempty. Here we repeat the arguments of theorem 1, but using the elements in RPE(T) that have constant matchings and give all players nonnegative payoffs (and we know that this set is nonempty as argued above). Order player roles so that \(n_i \geq n_k\), whenever \(i>k\). Let RPE_1 be the subset of elements of RPE(T) that have constant matchings and reach maximal payoff for player role 1, subject to all player roles getting at least 0 and the matching being constant (and we know that a maximum exists since this set is compact, as it is a compact subset of a compact set). Let RPE_2 be the subset of those that maximize player 2 types utilities, subject to being in RPE_1. Inductively, let RPE_k be the subset of those that maximize player role k’s utility, subject to being in NE_{k-1}. Pick an element \((\sigma,F)\) of RPE_n. This satisfies both (a) and (b) by construction.

Compactness of RME(T) now follows along similar lines as the proof of the corresponding claim in Theorem 1 (given the compactness of RPE(T)), as a violation of (a) or (b) at the limit of a sequence of equilibria would imply a violation far enough along the sequence. ◊

**Proof of Proposition 1:** We use the definition of renegotiation-proof equilibrium (abbreviated RNE(t)) found in Benoit and Krishna (1993). First, we show that matching every player and having every matched group play the same \(\sigma\) in RNE(1) must be a matching equilibrium of the 1 period game. By definition of renegotiation-proof equilibrium, \(\sigma\) must be a Nash equilibrium of the 1 period game and by assumption \(u_j(\sigma) \geq 0\) thus condition (a) of matching equilibrium is met. Next consider condition (b) of matching equilibrium. By definition of renegotiation-proof equilibrium, players who are currently grouped together do not want to play a different Nash. Since every group plays the same renegotiation-proof equilibrium, it must be that every player in role \(i\) receives the same payoff thus there is no group of agents who want to rematch and play a different Nash, and so condition (b) must hold true. Thus having every player matched and all groups play \(\sigma\) is a matching equilibrium of the 1 period game.

Next, we show that having every player matched by some constant matching function \(F\) and having all groups play the same \(\sigma\) in RNE(t) (for any \(\sigma\) in RNE(t)) must be a group-stable repeated matching equilibrium of the t period game. We do this by induction,
presuming it to be true up through t-1. Let us also presume that any \((\sigma, F)\) in RME (t-1) such that some group is matched after all histories must have that group playing a renegotiation-proof equilibrium.\(^{18}\) We will show that the same things are true for t. Let us first show that \((\sigma, F)\) is in RPE(t). By definition of renegotiation-proof equilibrium, all continuation payoffs of \((\sigma, F)\) are in RNE(t-1). So, by induction and the supposition that all continuation payoffs are nonnegative, all continuations of \((\sigma, F)\) are in RME (t-1) for the (t-1) period game. Thus condition (i) is met. Condition (ii) is met by the definition of renegotiation-proof equilibrium. Next, we show that \((\sigma, F)\) is in RME(t). By assumption \(u_i(\sigma, F) \geq 0\), and so condition (a) of repeated matching equilibrium is met. Next we show that condition (b) is met. Since every matched group plays the same \(\sigma\), we know that every player in role i must have the same expected payoff (if one player i deviates, then all have incentive to deviate) and so we can just show that no group who is currently matched wants to change strategies as there will be no extra gain from a new group forming and changing strategies. Given the induction step that any continuations of \((\sigma, F)\) must have all constant groups playing a renegotiation-proof equilibrium, and that no other renegotiation-proof equilibrium can Pareto dominate the current renegotiation-proof equilibrium (by the definition of renegotiation-proofness), it follows that (b) is satisfied. Now suppose that \((\sigma, F)\) is in RME (t) and that \(F\) is constant for some group and that group is not playing a renegotiation-proof equilibrium. By (ii) of RPE(t) it must be that this group is playing a subgame-perfect Nash equilibrium. And by definition of renegotiation-proofness, it must then be that there exists a renegotiation-proof equilibrium which does better for all the individuals in this group than the subgame-perfect Nash that they are playing does. However, the existence of such a renegotiation-proof equilibrium would violate (b) of the definition of RME. \(\Box\)

**Proof of Proposition 2:** The following is a group-stable repeated matching equilibrium:

We begin with a description of the play on the equilibrium path. In all periods, all players in role 2 are matched with the same players and every matched pair plays (across, down) in all periods prior to T, and in period T all matched pairs play (down, down).

In order to show that this is part of an GSRME, we proceed by induction on T. It is clearly true for T=1. So, let us suppose that it is true for T-1 and show that it is also true for T, where T is at least 2. Let us specify what happens if some player deviates in period 1 (as we know that the continuation is already a GSRME.) If exactly one player in role 1 deviates, then she is unmatched for the rest of the game, and the remaining players are matched into a GSRME of the T-1 period game. If there are any other combinations of deviations, then simply continue with the prescribed strategies in the remainder of the game.

To see that this is a GSRME (and RME) we first show that no player in role 1 will ever cheat. The maximum gain from deviating to a role 1 player is \(a-c\) but the loss from deviating is at least \(a\) (as the player is unmatched in the last period), and thus there is no

\(^{18}\) This does not imply the more general equivalence between GSRME and RNE, as there are other non-constant RME.
incentive for players in role 1 to deviate. Players in role 2 receive their maximal possible payoff and have no incentive to deviate.

To see that there is no group that can reorganize itself and achieve higher payoffs, simply note that the player 2’s are receiving their maximal possible payoff among all continuation equilibria and among current payoffs.

**Proof of Proposition 3:** Let us start with the first part of the proposition. The following is a group-stable repeated matching equilibrium:

We begin with a description of the play on the equilibrium path. In all periods, all players in role 1 are matched with the same players and every matched pair plays (across, across) for the first $T-[(d-b)/a]$ periods. For the last $[(d-b)/a]$ periods all matched pairs play (down, down).

In order to show that this is part of an GSRME, we proceed by induction on $T$. It is clearly true for $T=1$. So, let us suppose that it is true for $T-1$ and show that it is also true for $T$, where $T$ is at least 2. Let us specify what happens if some player deviates in period 1 (as we know that the continuation is already a GSRME.) If exactly one player in role 2 deviates, then she is unmatched for the rest of the game, and the remaining players are matched into a GSRME of the $T-1$ period game. If there are any other combinations of deviations, then simply continue with the prescribed strategies in the remainder of the game.

To see that this is a GSRME (and RME) we first show that no player in role 2 will ever cheat. The maximum gain from deviating to a role 2 player is $(d-b)$ but the loss from deviating is at least $a[(d-b)/a]$ (as the player is unmatched in the last $[(d-b)/a]$ periods), and thus there is no incentive for players in role 2 to deviate. Players in role 1 cannot gain from deviating since deviating in early periods (by playing down) creates a per period loss of $(b-a)$.

Next we check that there is no group that can reorganize itself and achieve higher payoffs. First we check that no such group exists in the last $[(d-b)/a]$ periods. Here there is no way for a pair of players to ever play (across, across) as role 2 always has incentive to cheat since the gain from cheating is $(d-b)$ but the loss from cheating (and being unmatched for the rest of the game) is at most $a((d-b)/a) - 1)$. Thus in the last $[(d-b)/a]$ periods the best any pair (or group) can do is to play (down, down). Second we check that no group can reorganize and do better in the first $T-[(d-b)/a]$ periods. We know that in the last $[(d-b)/a]$ periods (down, down) must be played to prevent role 2 from cheating. Thus the best that role 1 can hope to do is to achieve a payoff of b in the first $T-[(d-b)/a]$ periods. No group can form and make role 1 better off.

We now handle the second part of the proposition. We demonstrate that under these conditions, there exists a repeated matching equilibrium that is better for all players involved than playing down, down in each period. The following is such a group-stable repeated matching equilibrium.
In period 1, all players in role 1 are matched and every matched pair plays (across, across). In period 2, all players in role 1 are matched and all players in role 2 who were unmatched in period 1 are matched. (If $2n_1 < n_2$ then only $n_1$ of the unmatched period 1 role 2 players are matched.) Every matched pair plays (across, across). In period 3, repeat the matching and play of period 1. In period 4, repeat the matching and play of period 2. Continue to repeat the matching and play of periods 1 and 2 until period $T-2\left(\frac{(d-b)}{a}\right)_+ + 1$ is reached. In period $T-2\left(\frac{(d-b)}{a}\right)_+ + 1$, players are matched as in period 1 and every matched pair plays (down, down). In period $T-2\left(\frac{(d-b)}{a}\right)_+ + 2$, players are matched as in period 2 and every matched pair plays (down, down). Repeat periods $T-2\left(\frac{(d-b)}{a}\right)_+ + 1$ and $T-2\left(\frac{(d-b)}{a}\right)_+ + 2$ until period $T$ is reached.

In order to show that this is part of an GSRME, we proceed by induction on $T$. It is clearly true for $T=1$. So, let us suppose that it is true for $T-1$ and show that it is also true for $T$, where $T$ is at least 2. Let us specify what happens if some player deviates in period 1 (as we know that the continuation is already a GSRME.) If exactly one player in role 2 deviates, then she is unmatched for the rest of the game, and the remaining players are matched into the GSRME of the $T-1$ period game described in the proof of Proposition 3. If there are any other combinations of deviations, then simply continue with the prescribed strategies in the remainder of the game.

To see that this is a GSRME (and RME) first we show that no player in role 2 will ever deviate. The maximum gain from deviating for a role 2 player is $(d-b)$ but the loss from deviating is at least $a\left(\frac{(d-b)}{a}\right)_+$, so role 2 will never deviate. Additionally, player 1 is never able to gain from deviating.

Next we check that no group of players can gain from deviating. First we check that no such group exists in the last $2\left(\frac{(d-b)}{a}\right)_+$ periods. Clearly role 1 can be made better off by joining a group that plays the GSRME of Proposition 3 which plays (across, across) until period $T-\left(\frac{(d-b)}{a}\right)_+ + 1$. However, this GSRME is sustained by the existence of at least one unmatched role 2 player who acts as an outside option and clearly moving from the current strategy to the Proposition 3 strategy would make such a role 2 player worse off (or no better off). Thus there is no way to make role 1 better off except by forming a group which would hurt (or not help) a role 2 player and so such a reorganization will not occur. Similarly there is no way to make role 1 better off in earlier periods. Thus this strategy forms a GSRME (and RME).

Lastly we show that this GSRME Pareto dominates play where (down,down) is played in every period. At the current GSRME all players in role 1 receive an average per period payoff strictly greater than $a$ and if $b > a\left(1+\frac{T}{T-2\left(\frac{(d-b)}{a}\right)_+-1}\right)$ then all players in role 2 (who are matched in either period 1 or 2) receive an average per period payoff strictly greater than $a$. Thus if (down,down) is played by any group of players a new group can form (consisting of at least one role 1 player and two role 2 players) which makes everyone in the group strictly better off. Thus there is no GSRME where players play (down,down) in every period.
**Proof of Proposition 4:** By Proposition 1, we know that having everyone matched to the same partner in every period and having all play (down, down) is a RME. If we tried to improve everyone’s payoff and have everyone play (across, across) in the beginning of the game, then in the last period of playing (across, across) players in role 2 will cheat as they know that since $n_1=n_2$ ostracism is not a credible threat. ◇