1-2005

Social Games: Matching and the Play of Finitely Repeated Games

Alison Watts
Southern Illinois University Carbondale

Matthew O. Jackson
California Institute of Technology

SOCIAL GAMES:
Matching and the Play of Finitely Repeated Games

Matthew O. Jackson and Alison Watts

Draft: January 31, 2005

Abstract: We examine a new class of games, which we call social games, where 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, where play depends on the relative populations of players in different roles, among other things.

We also examine finite repetitions of games where players may choose to rematch in any period. Some equilibria of fixed-player repeated games cannot be sustained as equilibria in a repeated social game. Conversely, the set of repeated matching (or social) equilibria also includes some plays that are not part of any subgame perfect equilibrium of the corresponding fixed-player repeated games. We explore existence under different equilibrium definitions, as well as the relationship to renegotiation-proof equilibrium. It is possible for repeated matching equilibria to be completely distinct from renegotiation-proof equilibria, and even to be Pareto inefficient.

Humanities and Social Sciences 228-77,
Caltech, Pasadena, CA 91125, USA
(jacksonm@hss.caltech.edu)
http://www.hss.caltech.edu/~jacksonm/Jackson.html,

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

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 work for 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. While this is a relatively obvious statement, we have no systematic 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 has strong, systematic, and intuitive implications for behavior. We examine a new class of games called social games where players not only choose strategies but also choose with whom they play.

We examine two situations: one where 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, 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. We show that matching equilibria generally exist in bipartite settings - where the matchings of players are into pairs who play a two-person game. However, when the matchings of players are into groups of three or more players matching equilibria may not exist under the most general formulation, but do exist when players care only about the play of the game and not about the identities of the players with whom they are matched. We also show that the requirement that no group of players desire to leave their current group and match together has strong implications for the play in the game. It implies that only Nash equilibria which are not strictly Pareto dominated by other Nash equilibria can be played, and in fact results in a selection from that set, even when the population of players is completely evenly matched. We also show that play can depend on the relative populations of players available for different roles in the game, with the selection among equilibria favoring players who are less populous.

1

When the game is finitely repeated, and rematching is possible in any period, we show that existence depends on whether deviating players are allowed only to rematch once and for all, or whether they can suggest more complicated rematching plans that include other players outside of their group. The possibility of rematching results in a set of equilibria that is neither a superset nor a subset of the set of fixed-matching subgame perfect equilibria. In particular, through rematchings players can be rewarded and punished with payoff combinations that cannot be achieved in fixed-matching games. This changes the structure of play that can be sustained.

In terms of what we learn about how the ability to rematch affects play, 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, and
- in repeated games with mismatched population sizes, play can differ completely from renegotiation-proof equilibria.

Related Literature

One obvious strand of related literature is the matching literature that followed the seminal paper of Gale and Shapley (1962) and is detailed in Roth and Sotomayor (1989). That standard matching world is the special case of our model where the game played between players is degenerate and players' payoffs depend only on their partners' identities. We show that well-known results on existence and the lattice structure of matchings from the bipartite matching world have analogs in our setting when the game is bipartite, but also that there are new aspects to the matching structure that emerge due to subsequent interaction among players.

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. If it is possible to match all players at the same time and all players are identical within each role, then our equilibria are a superset of the set of renegotiation-proof equilibria. The new equilibria that emerge are supported by the threat of rematching. More generally, in cases with an imbalance in the possible groupings of players, our equilibria can differ completely from the set of renegotiation-proof equilibria, and include plays that are not supported by any previous equilibrium concept.

Finally, there are other[1] papers that have examined endogenous interactions and implications for behavior. These include market and bargaining settings as studied by Kelso and Crawford (1982), Rubinstein and Wolinsky (1985), Hatfield and Milgrom (2003), among others; implementation as studied by Jackson and Palfrey (1999), and 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)). There are also papers showing that endogenous interactions can lead to efficient play. Mailath, Samuelson, and Shaked (2001) examine a local interactions model where agents have the ability to seclude themselves from undesirable opponents in games that have one Nash equilibrium that strictly Pareto dominates all other correlated equilibria. Rob and Yang (2003) and Ghosh and Ray (1996) examine endogenous formation of long term relationships where partners play a Prisoner's Dilemma game with each other and where players are of different types. 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.

2. The One-Period Model

We begin by providing definitions for settings where the choice to match is taken once.

Given is a normal form game with **player roles** denoted by $i \in N = \{1,...,n\}$.

There is a **population**, P_i, of players who are of role i. For instance, P_1 could be all of the women and P_2 could be all of the men in the society if the game is the ``Battle of the Sexes'' game. Let

$$P = \{p \subset \cap P_i \text{ such that for each i there exists one } a \in P_i \text{ such that } a \in p\}$$

be collection of all sets of players consisting of one player in each role. We use i,j, and k to denote indices of different player roles. We use a,b, and c to denote generic players. We denote generic elements of P by p.

Let n_i be the cardinality of P_i, and order player roles so that $n_i \geq n_k$, whenever i>k.

Each player role i has a strategy set S_i, and a player $c \in P_i$ in that role receives a payoff $u_c(s,p)$ if s is the vector of strategies that is played and c is matched in group p, where s is in $S = S_1 \times ... \times S_n$. The payoff function u_a is a von Neumann-Morgenstern utility function. Mixed strategies for a player c in role i are denoted μ_c in $\Delta(S_i)$. We let $u_c(\mu,p)$ denote the expected utility for a player c who is matched in the set of players p and μ denotes the n-vector of mixed strategies played by the players in p.

---

[1] See also Watson (1999) who studies mutual investment with uncertainty over types.

Note that payoffs can depend on both the strategies and the set of players with whom a given player is matched. This allows players to differ, for instance, in skill or ability among other things.

A specification of (n; P_1,...,P_n; S, u) is called a *social game*.

Unless otherwise stated we will assume that S is finite, and hence the existence of a Nash equilibrium for any group of players is assured. It will be clear that our definitions and results extend directly to any game such that each group of matched players has a nonempty and compact set of equilibria, and continuous payoffs across (mixed) strategies. We also note that 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 or is a Bayesian game.

A **matching** is a mapping f from $\cup P_i$ into $(\cup P_i) \cup P$, such that
(i) either f (a)= {a} or f(a)=p$\in$P such that a$\in$p, and
(ii) if f(a)=p$\in$P and b$\in$p, then f(b)=p.

The interpretation is that f(a) is the set of players that a is matched with. (i) states that either player a is unmatched, or else is a matched in a group p. (ii) states that if player a is matched in a group that includes player b, then player b has to be matched in that same group.

We normalize the payoff of an unmatched player to 0. Given a mixed strategy m profile for all players 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 played.   So, $U_c(m,f)=u_c(m_p,p)$, if f(c)=p and $U_c(m,f)=0$ if f(c)={c}, where m_p denotes the mixed strategy profile of the players in p under m.

A **matching equilibrium** is a mixed strategy profile m, and a matching function f such that

(a) if f(c)=p$\in$P for some player c, then m_p is a Nash equilibrium for the players in p and $U_c(m,f) \geq 0$; and

(b) there does not exist p$\in$ P, and a profile of strategies μ for the players in p such that $u_c(\mu,p) > U_c(m,f)$ for all c$\in$ p and such that μ is a Nash equilibrium for the players in p.[2]

A matching equilibrium draws from both non-cooperative and cooperative game theory, marrying prominent solution concepts from each.[3] It requires Nash equilibrium play

---

[2] We require a strong form of blocking which, similarly to many other core-based concepts, aids in existence.

[3] As an alternative, one could attempt to model things entirely non-cooperatively (or cooperatively). For example, one could model some extensive form process by which groups form or reform. There are strong

within matchings, and a form of core stability across matchings.   The Nash play imposes stability of play within a matching, while the core approach imposes restrictions on play both across and within groups. The restrictions within groups derives from the threat of rematching across groups. 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 P_1= {1} and two men P_2={2,3}**. The woman is in the row player role, while the men are in the column role and are both identical. The payoffs to the players are described by the following matrix, and the woman's payoff is independent of which man she plays with.

	A	B
A	1,3	0,0
B	0,0	3,1

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 of 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.

Note that this example also illustrates 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. **Heterogeneous Players within Populations**

We begin with an analysis of the general case, where players within the same role can be heterogeneous. This means that players within a given role might have different utility functions, and also that players might care about the identity of the other players with whom they are matched. For instance, in a battle of the sexes game different men might differ both in their preferences over plays of the game and with whom they are matched.

---

advantages to the approach we employ here. Nash equilibrium is very well-suited to analyzing play within groups, while the core enables one to model group formation in a protocol-free manner and is especially suited to capturing coalitional incentives.

Moreover, any given man might have preferences over the play of the game that depend on which woman he is matched with. Thus, each group of players can have different sets of Nash equilibria and Nash equilibrium payoffs.[4]

It is easy to see that without any restrictions on payoffs and with more than two player roles, existence will generally not be guaranteed. This follows directly from what is known in the multipartite matching literature (e.g., see Roth and Sotomayor (1989)), and is illustrated in the following example.

Example 2: Nonexistence of equilibria in a multi-partite, heterogeneous player setting.

Consider 6 players in 3 player roles. Players 1 and 4 are in role 1, players 2 and 5 are in player role 2, and players 3 and 6 are in role 3. Let there be a single Nash equilibrium for each matched group of players. Let the payoffs from those Nash equilibria be as follows: (3,3,3) for groups {1,2,3} and {4,5,6}; (4,4,4) for group {4,2,3}; (1,1,1) for groups {1,5,6} and {4,5,3}; (2,5,2) for {1,2,6}; and (0,0,0) for all other groups.
The only potential matchings are then {{1,2,3},{4,5,6}}; {{4,2,3},{1,5,6}}; and {{4,5,3},{1,2,6}}, and combinations where some players are unmatched.[5]
Note that {4,2,3} blocks the first matching, {1,2,6} blocks the second matching, and {4,5,6} blocks the third matching, and that any matching where some players are single is blocked as well. Thus, there is no matching equilibrium.

While existence is a problem in general multipartite settings, it is not a problem in bipartite settings nor in settings where players care only about the play in the game and not about the identity of the agents with whom they are matched. Let us examine these two settings in turn.

3.1 Bipartite Settings

Let NE(p) denote the set of Nash equilibria for the group of players p.

Say that a strategy matching profile (m,f) is *plausible* if (a) in the definition of matching equilibrium is satisfied (m_p is an individually rational Nash equilibrium for any matched set of players p), and for any c and $p=f(c) \in P$, m_p is not Pareto dominated by any m' in NE(p).

Say that *players are never indifferent* if for any two plausible strategy-matching pairs (m,f) and (m',f'), $U_c(m,f) \neq U_c(m',f')$ whenever $f(c) \neq f'(c)$ or $f(c)=p= f'(c)$ and $m_p \neq m'_p$.

In situations where players are never indifferent, let $\geq_i$ be the partial order defined by saying that (m,f) $\geq_i$ (m',f') if all players in role i weakly prefer (m,f) to (m',f').

---

[4] This is clearly a generalization of the usual matching world. If each group had a single Nash equilibrium in their game, then we would be in the standard matching world. If, however, they have multiple Nash equilibria, then the problem becomes more intricate, as now the preferences are not uniquely defined.
[5] We can represent a matching directly by the induced partition over players.

Let us say that a matching equilibrium f is *Player role i-optimal* if $(m,f) \geq_i (m',f')$ for all matching equilibria f'.

In the bipartite setting, there is a nice structure to the set of matching equilibria. This is well-known for the standard marriage-market setting, and turns out to extend to the social-game setting. The following theorem is an extension of well-known results in the standard matching literature of Gale and Shapley (1962) and Conway (as reported by Knuth (1976)).

Theorem 1: If there are two player roles, then there exists a matching equilibrium. Moreover, if players are never indifferent, then there exists both a Player role 1-optimal matching equilibrium and a Player role 2-optimal matching equilibrium. Additionally,
(A) for any two distinct matching equilibria (m,f) and (m',f'): $(m,f) \geq_1 (m',f')$ if and only if $(m',f') \geq_2 (m,f)$, and
(B) the set of matching equilibria forms a distributive lattice (based on either $\geq_1$ or $\geq_2$).

The first part of Theorem 1 is proven through an extension of the Gale-Shapley deferred-acceptance algorithm. The intuition for the proof is as follows. Let us refer to player role 1 as men and player role 2 as women. Each man proposes to form a match with a woman, and also specifies a Nash equilibrium to be played by the couple. The men start by proposing their most preferred match-equilibrium (breaking ties in some fixed manner, and making no proposal if they prefer to remain single). Each woman views her proposals and selects the most preferred one, provided she would rather not remain single. Next, each rejected man makes a new proposal, and each woman considers any new proposals received and selects the best one provided it is preferred to her current situation. The algorithm continues until each man is either matched or has made all the proposals that he prefers to remaining single. Note also that as in the traditional matching world of Roth and Sotomayor (1989) this algorithm makes it a dominant strategy for each man to reveal his true preferences.

Proposition 1: Consider a social game with two player roles such that players are never indifferent and all pairs of players (from different populations) have at least one Nash equilibrium which generates positive payoffs for both players. Every matching equilibrium has the same set of unmatched players.

Theorem 1 and Proposition 1 are both generalizations of results from the bipartite matching setting. Next we explore aspects of the model that involve factors not present in the standard matching setting.

We say one player b *weakly dominates* another player c (in the same player role) if for every potential matched group of players p with c in p, and every Nash equilibrium μ for p that gives all players a nonnegative payoff, there exists a Nash equilibrium μ' for p'

where b replaces c that strictly Pareto dominates μ for the players other than b in p' and gives b a nonnegative payoff.

Players are *well-ordered* if for every pair of players b and c in the same player role, either b weakly dominates c, or c weakly dominates b.

When players are well-ordered, we have an unambiguous ordering over the players from all players' perspectives.   It would seem natural to expect that in this case, any matching equilibrium would turn out to be assortive (with highest ranked players matched with other highest ranked players, etc.), or at least that there should exist one such equilibrium. This is true in the standard marriage-market setting, but turns out not to be true in the social game setting. The following example shows a case where the only matching equilibrium involves mismatching high ranked players with low ranked players.

Example 3: Non-Assortive Matching.

Let there be two player roles and two players in each role. Players 1 and 3 are those in role 1 and players 2 and 4 are those in player role 2.   When players 1 and 2 are matched, they have two possible Nash equilibria, leading to payoffs of (4,2) and (2,4).[6] When players 1 and 4 are matched (and the same for 3 and 2 matched), they have two possible Nash equilibrium payoffs of (3,1) and (1,3). When players 3 and 4 are matched they can only generate a payoff of (0,0).

Note that here, players 1 and 2 unambiguously dominate their counterparts in the same roles, 3 and 4, by generating higher payoffs regardless of their matching.   Nevertheless, the only matching equilibrium has player 1 matched with player 4 and player 2 matched with player 3. To see this, first note that if we try to match players 1 and 2, then their payoff must be either (2,4) or (4,2). Given the symmetry, let us assume, without loss of generality, that it is (2,4). Then players 1 and 4 can block and get (3,1), which is better for both players.   Indeed, the only matching equilibrium has players 1 and 4 matched with payoff (3,1), and players 2 and 3 matched with payoffs (1,3) (with the higher payoff for player 2 who is in role 2).

This example shows that in order to get a clearly assortative matching a stronger definition of "domination" would be required. It is not enough to have each corresponding equilibrium have a higher payoff. It would be sufficient to have the *worst* equilibrium of a matching with one player dominate the *best* equilibrium of a matching with another.

We know from Theorem 1 that in the two-role case, there exists both a Player 1-role optimal and a Player 2-role optimal matching equilibrium. However, as the next example

---

[6] Note that generically, there will be an odd number of Nash equilibria. The example is easily modified to include a third possible Nash equilibrium payoff for each set of players. For instance, have these be payoffs to Battle of the Sexes games, with a mixed strategy equilibrium that leads to lower payoffs for both player roles than either of the pure strategy equilibria.

shows having an uneven number of players does not guarantee that the players in the minority receive their optimal matching equilibrium.

Example 4: Not all Matching Equilibria are Man-Optimal when Men are the Minority.

Let there be two men and three women, with preferences that allow indifference. Assume each man-woman pair has two Nash equilibria that are not strictly Pareto dominated by another Nash in their game. Let the Nash payoffs be as follows: If M_1W_1 are matched then the game played results in Nash equilibria with payoffs of (4,1) or (2,3). If M_2W_2 are matched then the game played results in Nash equilibria with payoffs of (4,1) or (2,3). If M_1W_2 or M_2W_1 are matched then the Nash payoffs are (3,2) or (1,4). If M_1W_3 or M_2W_3 are matched then the Nash payoffs are (3,2) and (1,4). Thus M_1 prefers W_1 to W_2 or W_3 in the sense that the man's best equilibrium gives M_1 a higher payoff if he is matched to W_1 and the woman's best equilibrium gives M_1 a higher payoff if he is matched to W_1. Similarly M_2 prefers W_2. Here there are four matching equilibria: M_1W_1 and M_2W_2 matched and both play the (4,1) Nash and W_3 unmatched; M_1W_2 and M_2W_1 matched and both play the (3,2) Nash and W_3 unmatched. (There are two other matching equilibria like the last one with W_1 or W_2 being unmatched, respectively.) So having more women than men guarantees that each man plays a man's favorite Nash, but it does not guarantee that each man receives his first choice of mate, even though this is feasible for both men.

4. Play-Only Preferences

While matching equilibria may not exist in general multipartite settings, they do exist in multipartite settings where players care only about the play of the game and not about with whom they are matched. We now examine this case in detail.

We say that players have *play-only preferences* if:
- $b \in P_i$ and $c \in P_i$ implies $u_b = u_c$, and
- $u_b(s,p)$ is independent of p.

Thus, players in the same role have the same preferences, and also players do not care about the players they are matched with - only how those players behave. Note that this does not require that players from different populations be similar, only that players who might play the same role be similar. Under play only preferences, we sometimes write u_i for the utility function of players in role i.

4.1 Existence of Equilibrium

We now show that the set of matching equilibria is nonempty and compact. Compactness is important in establishing the existence of repeated matching equilibria, which we explore below.

9

Theorem 2:   The set of matching equilibria of a social game with play-only preferences is nonempty and compact.

While the set of matching equilibria is compact, we note that the matching equilibrium correspondence (as payoffs are varied) is *not* upper hemi-continuous. This is in contrast with Nash equilibrium, and so the matching leads to changes in equilibrium structure, just as many equilibrium refinements do. The failure of upper hemi-continuity is illustrated in the following example.

Example 5: 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.

left	right
1 , 1	-1, 1+1/k

For any k, 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, ``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.

4.2 Characterization of Equilibrium

We now offer a complete characterization of the set of matching equilibria in social games with play-only preferences.

Before stating the characterization theorem, let us first explore some of the intuitive properties that we should see. 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 6: Bargaining, the Nash Demand Game, Double Auctions, and Ultimatum Games.

There are two populations called ``Buyers'' and `Sellers''.   A matched buyer and seller each announce a price in [0,1].   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.[7]   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

---

[7] It should be clear that any variation on this game will have similar features. For instance, allowing the price to be any function of the announced prices, or only having trade if the announced prices coincide, or having the buyer pay the buyer's price and the seller receive the seller's price (and burning the remainder). etc.

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 $P_1 < P_2$, then all matching equilibria have all buyers matched and each match playing the equilibrium where $s_1 = s_2 = 0$. If there are more buyers than sellers, so $P_1 > P_2$, then all matching equilibria have all sellers matched and each match playing the equilibrium where $s_1 = s_2 = 1$. In situations where the populations are evenly matched so that $P_1 = P_2$, then there are infinitely many matching equilibria. However, there are strong restrictions across equilibria. All players are matched and every match announces $s_1 = s_2 = p$, where p is the same across matches. To see this, note that, for instance, if two matched groups traded at prices p' and p where p'>p, 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 p' and p.

This example shows not only how relative population sizes interact with equilibrium selection. 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 3, 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 μ is a Nash equilibrium such that all player roles get a positive payoff, then there exists $i \leq k$ such that $u_i' \geq u_i(\mu)$.

Given a matching and mixed strategy profile (m,f) of some social game with play-only preferences, for each i let
$$v_i(m,f) = \min_{\{c \in P_i, \, f(c) \neq \{c\}\}} U_c(m,f).$$

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

Theorem 3: Consider a social game with play-only preferences. Suppose that there exists at least one Nash equilibrium such that all player roles have a positive payoff[8] 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*

(i) all players in P_1 to P_k are matched,

(ii) each matched group p plays a Nash equilibrium m_p such that all players get a nonnegative payoff, and

---

[8] The ``if'' part of the statement still holds if there 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.

(iii) $(v_1(m,f), v_2(m,f),\ldots, v_k(m,f)) \in PO(k)$.

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

Corollary 1 (Favoring of Less Populous Roles): Consider a social game with play-only preferences 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 two-player role setting, where no player is indifferent between any two Nash equilibria. If (m,f) is a matching equilibrium, then for any two players a and b in the same role i, there does not exist any Pareto efficient Nash equilibrium μ such that $U_a(m,f) < u_i(\mu) < U_b(m,f)$.

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. In this sense the players in the same role are nearly equally treated under a matching equilibrium.

5. Finitely Repeated Social Games

We now extend the concept of matching equilibrium to finitely repeated games. For this analysis, we specialize to the setting of play-only preferences, which is a maintained assumption for the remainder of the paper unless otherwise stated.

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 δ in [0,1]. Rematchings are possible in any period.[9]

Let h=[$s_1,f_1;s_2,f_2;...s_t,f_t$] 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. Let H(t) denote the union of all histories of the game through time t.   Let $H=\cap_0^{T-1} H(t)$ 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 which we denote by $\varnothing$.

---

[9] If rematchings are only possible ex ante, then the analysis is similar to that in the previous sections (where one can modify definitions to require that play be according to subgame perfect equilibrium rather than just Nash equilibrium).

A (behavioral) strategy for a player j in role i is a map $\sigma_j:H \rightarrow \Delta(S_i)$. The behavioral strategy profiles for the T period game are denoted S(T). So a strategy for a player indicates which mixed strategy they play following any finite history of length no more than T-1.

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,i) denote i's match after history h. The set of all T-period matching functions is denoted MF(T), and so MF(1)=MF.

Let $U_c(\sigma,F)$ denote player c's discounted expected utility if the profile of strategies σ are played and the matchings are governed by F.

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

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

Inductively, let RPE(t) be the set of (σ,F) in S(t) x MF(t) such that
    (i) $(\sigma(h),F(h))$ in RME(t-1) for all h in H(1), and
    (ii) no player wants to deviate from $\sigma(\varnothing)$ given the current matching $F(\varnothing)$ and anticipating the continuation governed by (σ,F).

Let RME(t) be the set of (σ,F) in RPE(t) such that
    (a) there does not exist any $c \in N$ with $0>U_c(\sigma,F)$, and

    (b) there does not exist any $S \subset N$ and (σ',F') in RPE(t) 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 definition has the same structure as the original definition of matching equilibrium, where the idea of Nash equilibrium is replaced by the notion of RPE(t) – an equilibrium that is a perfect equilibrium anticipating that the continuation will be a matching equilibrium of t-1 periods.
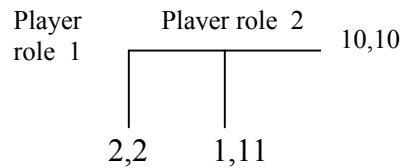
Note that here we allow a potentially large group S to deviate under (b), while in the single-period version we just allowed a single group p to deviate under (b). In a one-period setting, these two definitions are obviously equivalent. It is only with two or more periods that the definitions diverge. This will become clearer shortly, as we show that equilibria exist if we only allow a single group p to deviate, while they do not always exist (with two or more periods) if we allow larger groups to deviate.

Before discussing existence issues, let us examine an example where we see how the threat of rematching can influence play.

Example 7: Centipede and Trust Games: Repeated Matching Equilibria that are not Nash Equilibria (without matching)

13

Consider a simple version of a centipede game, which can also be viewed as a variation on a trust game. We illustrate the game both in normal and extensive forms.

Extensive form:

Player role 1 — Player role 2 — 10,10

2,2   1,11

Normal form:

	Down	Across
Down	2,2	2,2
Across	1,11	10,10

There are three players, one in role 1 and two in role 2, so that $P_1=\{1\}$ and $P_2=\{2,3\}$. 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 8/9, and thus all Nash equilibria lead to the 2,2 outcome.

Let us consider the matching equilibria in a twice repeated version of this game. First, let us note that without any possibility of rematching, the unique Nash (and hence subgame perfect, renegotiation-proof, etc.) equilibrium outcome is Down, Down in both periods. The following is a repeated matching equilibrium where play is Across, Across in the first period:

In period 1 players 1 and 2 are matched and both play across.

In period 2, if player 2 deviated in the first period, then players 1 and 3 are matched and both play down. Otherwise players 1 and 2 are matched and both play down.

Here, the threat of rematching induces player 2 to play across in the first period game, and makes it possible for the players to reach the Pareto efficient payoff of 10,10, which is not reached as part of any equilibrium (Nash, subgame perfect, renegotiation-proof, etc.) in a finitely repeated game without rematching.

While rematching can lead to new equilibria, it can also lead to existence problems, even with evenly matched populations. This is illustrated in the following example.

Example 8: Nonexistence of Repeated Matching Equilibrium

There are two periods and a discount factor of 1.

There are three player roles and twelve players $\{1,...,12\}$. Players 1,4,7,10 are in role 1, players 2,5,8,11 are in role 2, and players 3,6,9,12 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.

I

	left	center	right
up	5,1,3	-, -, -	-, -, -
middle	-, -, 9	-,26,-	25,25,25
down	-, -, -	6, 8, 10	28,9,-

II

	left	center	right
up	21,-,-	-, 13, 16	8, 10, 6
middle	-, -, -	3,5,1	-, -, -
down	20,20,20	-, -, -	9,-,-

III

	left	center	right
up	-,18,-	15, 15, 15	-, -, -
middle	10, 6, 8	-,-,-	-, -, -
down	13, -, 23	-, -, 13	1,3,5

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.

Let us now argue that there does not exist any repeated matching equilibrium of the two-period game.

To do this, we first show that if we consider just the set of players {1,...,9}, then 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 the

first period, then some arbitrary one-period matching equilibrium is played in the second period.

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 repeated matching equilibrium. It is clearly a matching equilibrium for any history in the second period. So let us check the first period. First let us 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 change in payoff 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, and 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 repeated matching equilibrium.

Now let us turn to the situation with four groups. 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 repeated matching equilibrium to this game. ◊

We remark that the above example also shows (to the careful reader) that even if we adopted a ``coalition-proof'' style equilibrium definition, where deviations have to be immune to deviations by further subgroups, existence would not be obtained.

The difficulty of existence of equilibrium is overcome by a weaker definition of repeated matching equilibrium, where only individual players or single matched groups are allowed to deviate. While this is not always as strong a concept, it is still of interest and often coincides with repeated matching equilibrium.

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 social game. Inductively, let GSRPE(t) be the set of (σ,F) in S(t) x 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 $\sigma(\varnothing)$ given the current matching $F(\varnothing)$ and anticipating the continuation governed by (σ,F).

Let GSRME(t) be the set of (σ,F) in GSRPE(t) such that
      (a) there does not exist any $c \in N$ with $0 > U_c(\sigma,F)$, and
      (b) there does not exist any $p \in P$ and (σ',F') in GSRPE(t) such that $U_c(\sigma', F') > U_c(\sigma,F)$ for all c in p, and such that F'(h,c)=p or c for all h and for all c in p.

We now show that group-stable repeated matching equilibria always exist.

Theorem 4: The set of group-stable repeated matching equilibrium is nonempty and compact for every t. Moreover, there exists such an equilibrium where the repeated matching function is constant on the equilibrium path.

Note that it is not the case that simply repeating a one period matching equilibrium is a group stable repeated matching equilibrium.

As we have remarked, the concept of (group-stable) repeated matching equilibrium incorporates some notions of renegotiation. That is, it is possible that the set of players p who consider changing equilibria in part (b) is actually already matched. It is very important to note, however, that although the definition of repeated matching equilibrium imposes some forms of renegotiation, it is neither a subset nor a superset of the set of renegotiation equilibria (except when t=1). This follows since changing the set of equilibria in a repeated game changes both the possibilities at a given date, and also the possible threat points offered for other dates. The changes in these combinations leads the sets to differ in a non-nested way when there are more than 2 periods.

In some cases, there is a relationship between group-stable repeated matching equilibria and renegotiation-proof equilibria. The following proposition illustrates this.

17

Proposition 2: 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 *group-stable* repeated matching equilibrium with a constant matching function f where all players are matched and where each group plays the same renegotiation-proof equilibrium.

The proposition holds for group-stable repeated matching equilibria, but not for repeated matching equilibria. This is clear from Example 8, as repeated matching equilibria need not exist even when there exists a renegotiation-proof equilibrium and populations are evenly balanced. It should also be noted that Proposition 2 requires all groups to play the same equilibrium: just like in the 1-period game, having some groups play one renegotiation-proof equilibrium and other groups play another renegotiation-proof equilibrium may not be a group-stable matching equilibrium.

While we already saw in Example 7, that when populations are not evenly matched there may exist (group-stable) repeated matching equilibria that are not Nash equilibria, much less renegotiation-proof equilibrium, we can also show that if populations are not evenly matched, then both the set of group-stable repeated matching equilibria and the set of repeated matching equilibria may no longer contain any renegotiation-proof equilibrium. This is because the minority player types will get to choose their favorite Nash equilibrium in the last period. Thus other threats will no longer be credible and the set of equilibria will change. This is illustrated in the following example.

Example 9: (Group-Stable) 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

	A	B	C
A	0,0	1,3	0,0
B	3,1	0,0	6,0

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 (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 (group-stable) 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 3). This eliminates (1), (2) and (3) as possible equilibria. Player 1 would like to play equilibrium (2) and see play of B,C in period 1. However, his or her 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 (group-stable) 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 (group-stable) repeated matching equilibria are renegotiation-proof, and any of them that have constant matchings are Pareto dominated by a renegotiation-proof equilibrium.

To see a situation where the set of Repeated matching equilibria differ completely from the set of subgame perfect equilibrium plays without rematching, let us revisit example 8.

Example 8 revisited:

Consider the centipede/trust game of example 8, but now repeated four times instead of just twice. The following is a repeated matching equilibrium:

In period 1, players 1 and 2 are matched and both play across.

In period 2, players 1 and 3 are matched and both play across if there was no deviation in period 1; and otherwise both play down.

In periods 3 and 4, all players play down. If player 2 deviated in period 1, then players 1 and 3 are matched in periods 3 and 4. If player 2 did not deviate in period 1, but player 3 deviated in period 2, then players 1 and 2 are matched in periods 3 and 4. Otherwise, players 1 and 2 are matched in period 3 and players 1 and 3 are matched in period 4.

The fact that this is a repeated matching equilibrium and that the overall payoffs are (24,12,12) means that no equilibrium where players play down in each period can be supported as a repeated matching equilibrium.

This example also shows the difference between repeated matching equilibrium and group-stable repeated matching equilibrium. While the above is a group-stable repeated matching equilibrium, there is a also group-stable repeated matching equilibrium where all players play down in all periods, as this cannot be improved upon by a group of only two players.

6. Discussion

19

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 several extensions of the analysis that obviously deserve future investigation.

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 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.[10]

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.

Example 10. Nonexistence in a single-population, symmetric, one-period social game

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

	A	B
A	0,0	1,2
B	2,1	0,0

There does not exist any matching equilibrium. 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

---

[10] Simply let P in the definition be the set of all vectors of n players.

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.

Proposition 3. 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:
- If every Nash equilibrium yields a negative payoff for at least one player, then all matching equilibria have all players unmatched.
- 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.

The proof is straightforward and left to the reader.

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 11. 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 $\delta=1$.

	C	D
C	3,3	-2,4
D	4,-2	2,2

There exists a 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

21

differs from any other equilibrium concept - including group-stable repeated matching equilibrium where players have set roles that they can fill.

6.2 Other Deviation Possibilities

We have examined two possibilities in terms of defining matching equilibria in finitely repeated settings.   There are still other possibilities. We mention one other definition here, but this definition and others are worthy of further investigation. In our definition of group-stable repeated matching equilibrium, the deviating group is just on its own, and cannot rely on rematching with any other players in any future continuation of the game. One could allow for a group to deviate in one period, but then still have rematchings with other groups possible in subsequent periods. We provide such a definition and an example of non-existence in the appendix.

6.3 Infinite Horizons

It is of interest to have definitions for repeated matching equilibria that cover the case of infinite horizon games.   As with renegotiation-proof equilibria, once one goes to an infinite horizon, it is no longer clear what a ``natural'' equilibrium definition is; especially as an inductive definition can no longer be used. This is an important problem that we leave for future research.

6.4 Other Equilibrium Definitions

Our definition of matching equilibrium has examined Nash equilibrium within a single period version of the game. We can also think of situations where the stage game is an extensive form game and we apply subgame perfection. As subgame perfect equilibrium and Nash equilibrium differ, it is clear that we will end up with different conclusions in some games.

For instance, reconsider the bargaining game of Example 5 in a case where buyers are less numerous than sellers (P1>P2), but where sellers move first in announcing their price. This is then essentially the same as an ultimatum game, and while there are Nash equilibria that lead trade at any price in [0,1], the only subgame perfect equilibrium outcome is trade at a price of 1.   This means that if we apply matching equilibrium where we apply subgame perfection as part of the equilibrium concept, the only outcome is that all matched groups trade at a price of 1, even though buyers are less numerous than sellers. This derives from the fact that no group of matched players can commit to play anything other than a subgame perfect equilibrium in the one shot extensive form. In order to escape such a conclusion, one would need to either allow some commitment on the part of players at the time where they are deciding on matchings, or else to allow for rematchings within the extensive form. Our analysis of repeated games considers such rematchings when the extensive form is a repeated game, but one could also examine

matching equilibria where rematchings are possible at a variety of times in more general extensive form games.

6.5 Cooperative Games

We have considered settings where the period game is a non-cooperative game. One can also provide analogous matching definitions where the period game is a cooperative game, either TU or NTU. Many of our results extend directly, especially in the static case, where one would simply replace Nash equilibrium play with a solution for the cooperative game. In the repeated context, the prospect of rematching alters the set of equilibria in some periods. This means that an analysis of a repeated cooperative game would in fact be easier than the analysis of a repeated non-cooperative game. For instance, provided there is existence whatever is seen as the solution concept for the cooperative game in the static setting, then there will be existence of repeated matching equilibria (and not just group-stable ones) in the repeated setting.

References

Benoit, J.P. and V. Krishna, 1993. "Renegotiation in Finitely Repeated Games." *Econometrica* 61, 303-323.

Bernheim, B.D. and D. Ray 1989. "Collective Dynamic Consistency in Repeated Games." *Games and Economic Behavior,* 1, 295-326.

Bernheim, B.D. and M.D. Whinston, 1987. "Coalition Proof Nash Equilibria, II: Applications," *Journal of Economic Theory* 42, 13-29.

Corbae, D. and J. Duffy (2003). ``Experiments with Network Economies,'' mimeo: University of Pittsburgh.

Droste, E., R. Gilles and C. Johnson, 2003. "Evolution of Conventions in Endogenous Social Networks," mimeo, Virginia Tech.

Farrell, J. and E. Maskin 1989. "Renegotiation in Repeated Games." *Games and Economic Behavior,* 1, 327-360.

Gale, D. and L. Shapley, 1962. "College Admissions and the Stability of Marriage," *American Mathematical Monthly* 69, 9-15.

Ghosh, P. and D. Ray, 1996. "Cooperation in Community Interaction without Information Flows," *Review of Economic Studies*, 63, 491-519.

Goyal, G. and F. Vega-Redondo, 2004. "Network Formation and Social Coordination," forthcoming in *Games and Economic Behavior*.

Hatfield, J.W. and P. Milgrom 2003. ``Auctions, Matching and the Law of Aggregate Demand,'' mimeo: Stanford University.

Jackson, M.O. and A. Watts, 2002. "On the Formation of Interaction Networks in Social Formation Games," *Games and Economic Behavior*, 41, 265-291.

Jackson, M.O. and T.R. Palfrey, 1998. ``Efficiency and Voluntary Implementation in Markets with Repeated Pairwise Bargaining,'' *Econometrica*, 66:6, 1353-1388.

Kelso, A.P. and V. Crawford (1982) ``Job Matching, Coalition Formation, and Gross Substitutes,'' Econometrica, 50:6, 1483-1504.

Knuth, D., 1976. Marriages Stables, Les Presses de l'Université de Montreal, Montreal.

Mailath, G.J., L. Samuelson and A. Shaked, 2001. "Endogenous Interactions" in The Evolution of Economic Diversity (A. Nicita and U. Pagano, Eds.), Routledge, New York.

Rob, R. and H. Yang, 2003. "Long-Term Relationships as Safeguards; Can dishonest types induce honest behavior?" Mimeo, University of Pennsylvania.

Roth, A. and M. Sotomayor, 1989. Two Sided Matching, Econometric Society Monographs No. 18: Cambridge University Press, Cambridge, UK.

Rubinstein, A. and A. Wolinsky (1985), ``Equilibrium in a Market with Sequential Bargaining,'' *Econometrica*, 53, pp. 1133-1150.

Watson, J., 1999. "Starting Small and Renegotiation," *Journal of Economic Theory*, 85, 52-90.

Appendix.

Proof of Theorem 1: Let us refer to player role 1 as men and player role 2 as women. To find a matching equilibrium we extend the Gale-Shapley deferred-acceptance algorithm, where a man proposes to a woman and also proposes a Nash equilibrium to be played by the couple. Let each man rank all the Nash equilibria from playing the game with every possible woman, where the man discards any Nash/woman pair which gives him a negative payoff. Artificially break ties, so that we have a strict ranking over acceptable mates and equilibria for each man, and similarly for each woman. The algorithm is as follows. First, each man simultaneously proposes to his best Nash/woman

pair (i.e., he proposes to this woman and proposes that they play this particular Nash equilibrium). Each woman then reviews her options and accepts the proposal of the man/Nash pair she likes best. If there is no proposal which gives her a nonnegative payoff, then all proposals are rejected. In the second round each currently unmatched man proposes to his second best Nash/woman pair. Again the women each accept their best acceptable proposal, where now a proposal from the first round is rejected if a woman receives a better proposal in the second round. This process continues iteratively, where each time a man is unmatched he proposes the best acceptable woman/Nash pair that he has not yet proposed, or else makes no proposal. The process ends when all unmatched men have exhausted their acceptable proposals. This process must end at a matching equilibrium: By construction, (a) of matching equilibrium is satisfied. The argument that (b) must also be satisfied is as follows. If there is a man who would prefer to be matched with someone else than his current mate and/or would prefer to play a different Nash equilibrium, then it must be that he already proposed this Nash to this woman and that at some prior step she turned him down, which means she had a better (or equivalent) offer. As the woman's ending match must be at least as good as the one she had at that time (by the structure of the algorithm), this woman would not be made better off by leaving her current Nash/man for this Nash/man pair. Thus, (b) is satisfied. If players are never indifferent, then this algorithm must end at the man-optimal matching equilibrium, since the algorithm ends with a matching equilibrium where each man is matched to his most preferred achievable Nash/woman. A woman-optimal matching equilibrium can be similarly constructed.

Let us now prove the remainder of the theorem. First we prove part (A) and show that if $(m,f) \geq_1 (m',f')$ then $(m',f') \geq_2 (m,f)$. Suppose to the contrary that $(m,f) \geq_1 (m',f')$ and that at least one woman, say W2, strictly prefers (m,f) to (m',f'). If W2's spouse at (m,f), say M1, also strictly prefers (m,f) to (m',f') then (m',f') is not a matching equilibrium since W2 and M1 prefer to sever their (m',f') ties and link with each other and play their (m,f) Nash. Thus it must be that M1 is indifferent between (m,f) and (m',f'). Since we assumed players are never indifferent this is only possible if M1 has the same spouse/Nash at both equilibria. But if this is true, then W2 would have the same spouse/Nash at both equilibria and thus would not strictly prefer (m,f). Thus the "if" statement of part (A) must be true. The "only if" statement follows from the above; simply replace the role 1 (2) players with the role 2 (1) players.

Next we prove part (B). Let (m,f) and (m',f') be two matching equilibria. Define $\sup_1\{(m,f),(m',f')\}$ to be the strategy matching profile where each man is matched with the spouse/Nash pair he most prefers (or points to) from either his (m,f) or (m',f') spouse/Nash pair. Define $\inf_1\{(m,f), (m',f')\}$ to be the strategy matching profile where each man is matched with the spouse/Nash he least prefers from either his (m,f) or (m',f') spouse/Nash pair. We show that $\sup_1\{(m,f),(m',f')\}$ and $\inf_1\{(m,f),(m',f')\}$ are both plausible matching profiles and that they are both in fact matching equilibria. First we show that $\sup_1\{(m,f),(m',f')\}$ is plausible. It is enough to show that there do not exist two men, say M1 and M3, who both point to the same spouse, say W2, when they point to their preferred spouse/Nash pairs. Suppose to the contrary that two such men exist and that M1 is matched to W2 at matching equilibrium (m,f) while M3 is matched to W2 at

25

matching equilibrium (m',f'). Since players are never indifferent, W2 must prefer either her spouse/Nash at (m,f) or at (m',f'). Say she prefers M1 or her spouse/Nash at (m,f). But then (m',f') cannot be a matching equilibrium since W2 prefers her spouse/Nash at (m,f) and M1 also prefers his spouse/Nash at (m,f); thus M1 and W2 prefer to sever their (m',f') links and link to each other and play their (m,f) Nash. Thus $\sup_1\{(m,f),(m',f')\}$ must be plausible.

Next we show that $\sup_1\{(m,f),(m',f')\}$ is also a matching equilibrium. Suppose to the contrary that $\sup_1\{(m,f),(m',f')\}$ is not a matching equilibrium, thus there exists a woman, say W2, who would like to sever her $\sup_1\{(m,f),(m',f')\}$ tie and link with a different spouse/Nash, say M1. At $\sup_1\{(m,f),(m',f')\}$ W2 must be linked with the same spouse/Nash she is linked with at either (m,f) or (m',f'), say it is (m,f). Since (m,f) is a matching equilibrium it must be that if W2 asks M1 to sever his (m,f) tie and link with her and play a certain Nash, M1 says no. Since M1 weakly prefers his spouse/Nash at $\sup_1\{(m,f),(m',f')\}$ to (m,f) he will also refuse to sever his $\sup_1\{(m,f),(m',f')\}$ link to link with W2. Thus even though W2 would like to sever her $\sup_1\{(m,f),(m',f')\}$ tie and link with another spouse/Nash she is unable to do so. Thus $\sup_1\{(m,f),(m',f')\}$ must be a matching equilibrium.

Next we show that $\inf_1\{(m,f),(m',f')\}$ is plausible and is a matching equilibrium. By part (a) we know that $\inf_1\{(m,f),(m',f')\}$ is the same as $\sup_2\{(m,f),(m',f')\}$. Thus from the above analysis $\sup_2\{(m,f),(m',f')\}$ must also be a matching equilibrium.

Lastly we show that if there exists (m'',f'') such that $(m'',f'')\geq_1(m,f)$ and $(m'',f'')\geq_1(m',f')$ then $(m'',f'')\geq_1 \sup_1\{(m,f),(m',f')\}$. This follows from the definition of $\sup_1\{(m,f),(m',f')\}$. Similarly if there exists (m'',f'') such that $(m'',f'')\leq_1(m,f)$ and $(m'',f'')\leq_1(m',f')$ then $(m'',f'')\leq_1 \inf_1\{(m,f),(m',f')\}$. Thus the set of matching equilibria must form a lattice (based on either $\geq_1$ or $\geq_2$).

To show that the lattice (based on either $\geq_1$ or $\geq_2$) is distributive involves straightforward manipulations, which we leave to the reader to verify. ◊

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 $n_1<n_2$, 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 n_1 men then there must be n_1 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 (n_1+1) women who strictly prefer the ME2 equilibrium. Since only n_1 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 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 players getting at least 0. Let NE_2 be the subset of those that maximize player 2 types utilities, subject to being in NE_1. Inductively, let NE_k be the subset of those that maximize player role k's utility, subject to being in NE_{k-1}.   If NE_n is empty, match all players to themselves. Otherwise, select any matching f such that players in P_1 are all matched and pick m such that all players play their role's component from some mixed strategy profile in NE_n (so that all players in the same role play the same mixed strategy). This forms a matching equilibrium.

Next, let us argue that the set of 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 an equilibrium for any given matching f is compact. Let $m^r \rightarrow m$, where (m^r,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 $u_c(m'_p) > U_c(m,f)$ for all $c \in p$ and such that m'_p is a Nash equilibrium. Then for large enough r, it follows that $u_c(m'_p) > U_c(m^r,f)$ for all $c \in p$, which is a contradiction.◊

Proof of Theorem 3:   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 th definition of matching equilibrium. Finally, let us show (iii). Suppose to the contrary that $(v_1(f,m), v_2(f,m),…, v_k(f,m)) \notin PO(k)$. Consider a group p consisting of a player from each player role 1 to k who is obtaining $v_i(f,m)$. Since that $(v_1(f,m), v_2(f,m),…, v_k(f,m)) \notin PO(k)$, 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 4: We prove the theorem by induction. Recall that player roles are ordered so that n$_i$ ≥ 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 2 established existence and compactness (and constant matching) for t=1. Supposing 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.

We first show that GSRPE(T) is nonempty and has an element with a constant matching function.   Let f be a matching that has some equilibrium in GSRME(T-1) for which the matching is constant and equal to f. Let F(0)=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 GSRME(T-1) that has the constant matching f. This then defines (σ,F) except for σ(0). Given that continuation payoffs are constant, pick any Nash equilibrium that gives all players nonnegative payoffs for σ(0). Specifying this as σ(0) defines a (σ,F) in GSRPE(T), since (i) is satisfied by construction, and (ii) is satisfied since σ(0) is a Nash equilibrium that gives all players nonnegative payoffs and the continuation is independent of play in the first period. Thus, GSRPE(T) is nonempty and has at least one element with a constant matching function.

Next, let us argue that GSRPE(T) is compact. Let $(\sigma^k, F^k) \to (\sigma, F)$, where (σ^k, F^k) is in GSRPE(T) for each k.   By the compactness of GSRME(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 σ would also be an improving deviation from σ^k for large enough k.

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

We first argue that GSRME(T) is nonempty. Here we repeat the arguments of theorem 1, but using the elements in GSRPE(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 $GSRPE_1$ be the subset of elements of GSRPE(T) that have constant matchings and reach maximal payoff for player role 1, subject to all player roles getting at least 0. Let $GSRPE_2$ be the subset of those that maximize player 2 types utilities, subject to being in $GSRPE_1$. Inductively, let $GSRPE_k$ be the subset of those that maximize player role k's utility, subject to being in NE_{k-1}. Pick an element (σ, F) of $GSRPE_n$. This satisfies both (a) and (b) by construction.

Compactness of GSRME(T) now follows along similar lines as the proof of the corresponding claim in Theorem 1 (given the compactness of GSRPE(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 2: 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 σ in RNE(1) must be a matching equilibrium of the 1 period game. By definition of renegotiation-proof equilibrium, σ 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 σ 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 σ in RNE(t) (for any σ 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 (σ, F) in GSRME (t-1) such that some group is matched after all histories must have that group playing a renegotiation-proof equilibrium.[11] We will show that the same things are true for t. Let us first show that (σ, F) is in GSRPE(t). By definition of renegotiation-proof equilibrium, all continuation payoffs of (σ, F) are in RNE(t-1). So, by induction and the supposition that all continuation payoffs are nonnegative, all continuations of (σ, F) are in GSRME (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 (σ, F) is in GSRME(t). By assumption $u_j(\sigma, F) \geq 0$, and so condition (a) of group-stable repeated matching

---

[11]This does not imply the more general equivalence between GSRME and RNE, as there are other non-constant GSRME.

29

equilibrium is met. Next we show that condition (b) is met. Since every matched group plays the same σ, 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 (σ,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 (σ,F) is in GSRME (t) and that F is constant for some group and that group is not playing a renegotiation-proof equilibrium. By (ii) of GSRPE(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 GSRME. ◊

A **strong group-stable repeated matching equilibrium** of a t-period game, denoted SGSRME(t), is defined inductively as follows.

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

Inductively, suppose that we have defined SGSRME through t-1.

Let SGSRPE(t) be the set of (σ,F) in S(t) x MF(t) such that
      (i) (σ(h),F(h)) in SGSRME(t-1) for all h in H(1), and
      (ii) no player wants to deviate from σ(0) given the current matching F(0) and
      anticipating the continuation governed by (σ,F).

Let SGSRME(t) be the set of (σ,F) in SGSRPE(t) such that
      (a) there does not exist any $p \in P$, (σ',F') in SGSRPE(t), and $c \in N$ with
      F(c,0)=p, F'(0,c)=c and $U_c(\sigma', F')>U_c(\sigma,F)$, and
      (b) there does not exist any $p \in P$ and (σ',F') in SGSRPE(t) such that
      $U_c(\sigma',F')>U_c(\sigma,F)$ for all c in p.

Notice that in the case of t=1, all repeated matching equilibrium definitions coincide with that of matching equilibrium.

In the case where there is just one group and no issue of individual rationality, all definitions coincide with the definition of finite period renegotiation-proof equilibrium (e.g., Benoit and Krishna (1993)).

Example 12: Nonexistence of Strong Group-Stable Repeated Matching Equilibrium

There are two player roles and four players {1,2,3,4}. Players 1 and 3 are in role 1, players 2 and 4 are in role 2.

Player role 1 is the row player and player role 2 is the column player. Payoffs are as follows

	left	center	right
up	1, 4	-10, -10	7, -10
middle	-10, -10	4,1	-10, -10
down	-10, 7	-10, -10	5,5

Here there are several matching equilibria in a one period game. These are where all players are matched and each group plays either of the pure strategy equilibria: up, left or middle center.   Note that any mixed strategy equilibrium gives a payoff of less than 1, and that down and right are weakly dominated strategies.

The following is in SGSRPE(2): in the first period:
players 1,2 are matched and play down, right and get payoffs of (5,5).
players 3,4 are matched and play up, left and get payoffs of (1,4).

In the second period --
If nobody from the first group deviated in the first period then
players 1 and 4 are matched and play middle center and get payoffs of (4,1)
players 3 and 2 are matched and play up left and get payoffs of (1,4)

If someone from the first group deviated in the first period then
players 1,4 are matched and play up left and get payoffs of (1,4)
players 3,2 are matched and play middle center and get payoffs of (4,1)

In terms of verifying that this is in SGSRPE(2) – it is clear that nobody will wish to deviate in the first period. If one of players 1,2 deviates in the first period, they will all get a payoff of 1 in the second period, while if they do not deviate, they get a payoff of (4,4) in the second period. So a deviating player will lose 3, but gain at most 2 in the first period.

Now, we can see that SGSRME(2) is empty – since the fact that the players can get a payoff of (5,5)+(4,4)=(9,9); cannot be given to two groups at once.
It is the second period regroupings that give players in roles 1 and 2 the incentives to play the prescribed strategies, and these regroupings are necessarily asymmetric.