Equilibrium Existence in Bipartite Social Games: A Generalization of Stable Matchings

Matthew Jackson
Stanford University

Alison Watts
Southern Illinois University

Abstract

We prove existence of equilibria in bipartite social games, where players choose both a strategy in a game and a partner with whom to play the game. Such social games generalize the well-known marriage problem where players choose partners but do not take actions subsequent to matching.

1. Introduction

The seminal paper of Gale and Shapley (1962), and the literature that has followed (see Roth and Sotomayor (1989)), examine stable matchings of individuals who care about who their partners are. While there are many applications where a partner's identity is enough information for a player to predict his or her payoffs, there are also many where other considerations are also important. For example, a worker cares not only about which firm he or she is employed by, but also about what contract he or she is offered and what other actions the firm may take. Similarly, a firm cares not only about which workers it hires, but also about what duties the workers perform and how they perform them.

We explore matching and equilibrium play in situations where both of these forces are important. There are a finite set of players who will be playing games. Players have discretion over which other players they are grouped with, as well as what action they play in the game. In the case where the game is degenerate, our model reduces to a standard matching model; while in cases where there is just one group of players, our model reduces to a standard game. The interesting case, where these two choices interact, is what we term a social game. We define such games and then show that existence of equilibria in these settings is guaranteed in bipartitate settings, but not in muli-partitite settings. We also provide some examples and results about the structure of equilibria in bipartite settings.

In a companion paper, Jackson and Watts (forthcoming), we study the special case of social games where players care only about the play of the game and whether they are matched or not, but do not care about which other players they are matched with. In such settings only the threat of rematching and relative population sizes matter in determining equilibrium play, and identities play no real role. In that special class of settings existence of matching equilibrium is guaranteed even in multi-partite settings, whereas here where identities matter, existence is only guaranteed in bipartite settings, and identities can play a central role in equilibrium structure, as we discuss below. The simplified structure in the companion paper permits an exploration of repeated social games, which is the bulk of the analysis in the companion paper, while with the general structure here, we concentrate on the one-shot setting.

2. Social Games: A Generalized Matching Model[1]

There is an underlying game in normal form, with *player roles* denoted by $i \in \{1,...,n\}$. The term "role" indicates 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 lie in $S = S_1 \times ... \times S_n$, with generic elements s_i and s, respectively. S is a finite set, and the set of

---

[1] The model and definitions are as in Jackson and Watts (forthcoming), but allow for play to depend on the identity of one's partners.

mixed strategies for player role i is denoted by $\Delta(S_i)$, the set of probability distributions on S_i.

There is a finite *population*, denoted P_i, of players who can fill role i. For instance, P_1 could be the set of firms and P_2 could be the set of workers if the game to be played is a shirking game. As another example, the populations could be men and women who play the "Battle of the Sexes" game. 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. $N = \cup P_i$ is the set of all players in the society, and $P = P_1 \times ... \times P_n$ 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.

Players' utilities can depend both on who they are matched with and what strategies are played. A player c in role i has a von Neumann-Morgenstern utility function over strategies and matches, such that $u_c(s,p)$ is c's payoff if s is the vector of strategies that is played and c is matched in group p, and $u_c(\mu,p)$ denotes the expected utility for a player c who is matched in group p and μ is the n-vector of mixed strategies played by p.

This formulation allows a player's payoffs to vary with the group with which he or she is matched. This means that the set of Nash equilibria could be quite different depending on which group of players are matched together. For example, it could be that players are playing two-by-two games, but that a player views the game as a prisoners' dilemma when matched with some players, and as a coordination game when matched with others.

A profile $(n; P_1,...,P_n; S, u)$ is called a *social 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 $f(b) = p$.

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 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.

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

1

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 played. Thus, $U_c(m,f)=u_c(m_p,p)$, 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 p.

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

So (a) requires that each matched group plays a Nash equilibrium, and also that each player receives a payoff greater than or equal to the payoff a player would receive if he left his current group and became unmatched. And (b) requires that no set of players, one from each role, could leave their current matches, form a new group, and play a Nash equilibrium that would give everyone in the new group a strictly higher payoff than what they receive in their current matches. This corresponds to a core definition, where the available feasible choices for any group of players (one from each role) is the set of Nash equilibrium plays in their game. In the special case where there is just one player per role, matching equilibrium corresponds with renegotiation-proof equilibrium.

If players within the same role are heterogeneous (or have different utility functions and care about the identity of the other players with whom they are matched), then 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)). The following example shows how this implication carries over directly from that literature.

Example 1: Nonexistence of a matching equilibrium 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.[2]
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. ◊

In Jackson and Watts (forthcoming), we show that existence of a matching equilibrium in multipartite settings is guaranteed in cases where players do not care about the identity of the players with whom they are matched, but only the play of the game in their match.

---

[2] We can represent a matching directly by the induced partition over players.

While existence is a problem in heterogeneous multipartite settings, it is not a problem in bipartite settings.

Matching Equilibrium in Bipartite Settings
The theorem below is an extension of well-known results in the standard matching literature of Gale and Shapley (1962) and Conway (as reported by Knuth (1976)).

Let NE(p) denote the set of Nash equilibria for the group of players $p \in P$.

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 each 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 $U_c(m,f) \geq U_c(m',f')$ for each player c in role i.

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'.

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

3

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.[3] 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. One might conjecture that in this case every matching equilibrium would be assortative (with highest ranked players matched with other highest ranked players, etc.), or at least that there should exist at least one assortative 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 2: Non-Assortative 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).[4] When players 1 and 4 are matched (and when 3 and 2 matched), there are 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).

Players 1 and 2 ``dominate'' their counterparts in the same roles, 3 and 4, by generating higher potential payoffs regardless of their matching. The dominance here is in terms of the frontier of payoffs that they bring to a partnership, although which point is selected can be an issue. Indeed, matching equilibrium is affected by this. The *unique* 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.

---

[3] See Gale and Shapley (1962), Conway (as reported by Knuth (1976)), and Roth and Sotomayor (1989).
[4] 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.

Thus, 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). ◊

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

Example 3: 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) or (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.

3. Concluding Remarks

We have introduced a new class of games, called social games, where players choose both their partners and strategies for the game. Such a specification enables us to study how population sizes and preferences over matchings interact with play within a match. The results here show that there can be subtle and unexpected implications about play in such settings. As there are many applications where this interaction is central, it presents a rich agenda for future research.

References

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

Jackson, M.O. and A. Watts, (forthcoming) "Social Games: Matching and the Play of Finitely Repeated Games," *Games and Economic Behavior*.

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

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

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 equilibrium/women pairs that he
might face from being matched 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 tentatively accepts the proposal of
the Nash/man pair that she likes best out of those proposed to her. 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 tentatively accept their best available 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 if he has already made all of his acceptable proposals. 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) is also 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
that 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.

We first 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 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 can block (m',f') by matching with each other and playing their (m,f) Nash, which would contradict equilibrium. 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, since it is plausible, there exists a woman, say W2, who would like to get rid of her $\sup_1\{(m,f),(m',f')\}$ spouse and match with a different spouse, say M1 and play some Nash equilibrium with M1. At $\sup_1\{(m,f),(m',f')\}$ W2 must be matched to the spouse/Nash she is matched to in at one of (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. From the proof of part (A) we know that $\inf_1\{(m,f),(m',f')\}$ is the same as $\sup_2\{(m,f),(m',f')\}$. 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 $\geq_2$).

7

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

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 women 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. $\Diamond$