COMPLEMENTARITIES, GROUP FORMATION AND PREFERENCES FOR SIMILARITY

MARCIN PĘSKI

Abstract. We present a model of two sociological phenomena: the tendencies to form groups and to favor others who are similar. Individuals divide society into friends and enemies. Individuals' payoffs depend on their own choices and on the choices of others. We assume different types of complementarities: Other things equal, each individual prefers to be friendly towards those who are friendly toward her (*second-degree complementarity*) and toward those who are friendly toward those ... who are friendly toward her (*higher-degree complementarity*). With second-degree complementarities, but no higher-degree externalities, individuals want to reciprocate friendship. Any additional amount of higher-degree complementarities pushes individuals to form groups. Next, we assume everybody may make mistakes that make him confuse individuals who are similar to each other. To minimize the cost of the mistakes, individuals want to keep their friends as different from their enemies as possible. Combined with group formation, individuals would like to be friendly toward others who are similar to them. Although individuals act *as if* they have preferences for similar others, in reality, their behavior is a best response to the equilibrium behavior of others.

1. Introduction

The tendencies to form groups and to favor others who are alike stand among two of the most pervasive properties of social interactions. The first property coordinates individual choices of friends. In a population divided into groups, any two friends of any given individual tend to be friendly toward each other. The second property,

*University of Texas at Austin, Department of Economics. Email: mpeski@gmail.com. I am grateful to Balazs Szentes, Leeat Yariv, Daaron Acemoglu, Matthew Jackson, David Easley, the participants of the meeting of the 2008 Game Theory Society meetings in Evanston as well as seminars at Cornell University and SFU for discussion and helpful remarks. All remaining errors are my own.

known also as *homophily* (McPherson, Smith-Lovin, and Cook (2001)), correlates the choice of friends with their individual characteristics. The two properties can be observed in various social situations and with respect to a wide range of relationships, and typically, these properties are observed together. They have important consequences for various social and economic outcomes, such as racial discrimination (Becker (1971), Schelling (1971)), social order and civil unrest (Horowitz (1985)), collective decision making (Alesina and Ferrara (2005)), information transmission (Jackson and Calvó-Armengol (2004)), and many others.

This paper presents a framework to discuss the causes, mechanisms, and connections between the two phenomena. We show that payoff complementarities push members of society to form groups. Similarity coordinates the composition of the groups. In an equilibrium, individuals act *as if* they have preferences toward befriending similar others, whereas in fact, their behavior is a best response toward others' seemingly homophiliac behavior.

The first part of the paper builds a model of group formation. There is a population of exchangeable individuals. Each individual divides the population into friends and enemies. (In the paper, we allow for more nuanced attitudes ranging from 'close friends' to 'worst enemies'.) There are two important assumptions. First, we focus on situations in which one cannot be friendly or hostile toward everybody. We assume that the number of potential friends or enemies is fixed, and the choice is about whom to be friendly to. Second, we assume that players' payoffs depend on the attitudes between them and other individuals and on the attitudes between other individuals. We discuss various types of complementarities between individuals' choices. Other things equal, a *second degree complementarity* makes an individual prefer being friendly toward individuals who are friendly toward her. Additionally, we assume *third* and *higher degree complementarities*: Each individual would like to be friendly toward individuals who are friendly toward individuals who ... who are friendly toward her. As a flipside, individuals would like to be hostile toward friends of their enemies and friendly toward enemies of their enemies. Although many social situations exhibit complementarities of second and higher degrees, to the best of our knowledge ours is the first paper that explicitly analyzes different types of complementarities. Consider the following examples:

- *Promotion within an organization.* Members of an organization (state bureaucracy, political party, army, firm) compete for promotion.[1] The number of available promotions is fixed. Each individual may support the advancement of other members, either by providing direct help or by blocking the advancement of a competitor. Because of the fixed number of promotions, a decision to support everybody is essentially equivalent to a decision to randomize the support. The strength of the support increases with the rank in the organization. This leads to complementarities. Because a promoted individual has a higher chance of helping others, i would rather support j if j will support i once promoted (second degree complementarity) or if j supports people who will support i (third degree complementarity). If the number of promotions available does not depend on the overall strength of the support, then the decision to support each individual is equivalent.[2]
- *Job search on a network.* Consider a model of job search through friends (Granovetter (1973), Jackson and Calvó-Armengol (2004)) There is a population of workers. Each worker fluctuates in and out of employment. If employed, workers receive information about available job opportunities and pass this information to their unemployed friends. If worker i passes information to worker j, i increases the probability that j becomes employed. The friendship choices exhibit complementarities. Worker i's payoff from passing information to individual j will increase if individual j is willing to pass information to i once i becomes unemployed, or j passes information to workers who are willing to pass information to i.
- *Peer-reviewed journal.* A peer-reviewed journal accepts an article if it is recommended by the referees. The weight of the referee's opinion increases with

[1]Competition for higher-level jobs in the colonial Ugandan Civil Service was one of the sources of the conflict between Africans and Indians that ended with the expulsion of the latter from Uganda by Idi Amin (see Motani (1977) and Twaddle (1975)).

[2]A related model of voting for voters can be found in Barbera, Maschler, and Shalev (2001).

> the number of papers he publishes. The individuals care only for the number of their own publications (and not, for example, for the quality of published papers). Because favorable reports lead to more publications and a higher reputation of referee j, the payoff of individual i from favorable reports on j's submissions increases if individual j tends to accept i's submissions (second degree complementarity) or if j accepts submissions of referees who ... who tend to accept i's submissions (higher degrees complementarity).

The friendship model is a large coordination game with a potential function (see Monderer and Shapley (1996)). We focus on equilibria that maximize the potential. These equilibria are known to have good equilibrium selection properties: they are robust to incomplete information Ui (2001), and they are uniquely stochastically stable under logit dynamics (Blume (1993), Young (1998), and Hofbauer and Sandholm (2002)).

We illustrate the results of the paper with an example of two profiles of friendship choices. In each of the two profiles, society is divided into two arbitrarily chosen, equal, and disjoint sets, A and B. In the first, *the group formation profile*, each individual is friendly toward the members of her own set and hostile toward the members of the opposite group. We say that A and B are groups of friends. In the second, *the contrarian profile*, individuals in set A are friendly toward members of set B, while hostile toward members of their own set. Similarly, members of set B are friendly toward As and hostile toward Bs.

In each of the two profiles, friendships are *reciprocated*, i.e., individuals are friendly toward those who are friendly toward them. In the first profile, any two friends of each player are also friends of each other; in the second profile, any two friends of each player are each other's enemies. We show that if there are only second degree complementarities, but no higher degree externalities, the two profiles lead to the same payoffs, and each maximizes potential.

With third or higher degree complementarities, only the group formation profile maximizes potential. This illustrates a sharp difference between second and higher degree complementarities for social behavior.

With more nuanced attitudes of friendship, we show that, given sufficient complementarities, the only profiles that maximize potential form a hierarchy of nested groups: Each player belongs to a small group of very close friends, which is contained in a larger group of slightly worse friends, which is contained in ...., and so on.

The second part of the paper modifies the original game by introducing a possibility of mistakes. Individuals may make mistakes and confuse other individuals. If individual i confuses individuals j and k, then i treats j with an attitude prescribed for k, and she treats k with an attitude appropriate for j. Second, we introduce heterogeneity into the population. Some individuals are more similar to each other than to others. The probability of a mistake increases with the similarity of two confused individuals.

Given second and higher degree complementarities, we show that, in each profile that maximizes potential in a game with mistakes, individuals are friendly toward those who are similar to them and hostile toward those individuals who are different. There is a simple intuition behind this result. Although confusing two friends or two enemies does not change anything, it is costly to confuse a friend with an enemy. To minimize this cost, individuals would like to keep their friends different from their enemies and similar to each other. If the population is divided into groups, then individual i's friends are also friends with each other. Hence, if i's friends are similar to each other, and their friends are similar to each other, than i must be also similar to her friends.

To see it in an example, suppose that each individual has one of two colors, Blue or Green. Mistakes are possible: two Blue individuals or two Green individuals can be confused, but nobody ever confuses a Blue individual with a Green one. Consider specific examples of the above profiles, in which groups A and B consist of, respectively, Blue and Green individuals. Each of the two profiles minimizes the costs of mistakes: Friends of any individual are similar to each other and different from her enemies. In the contrarian profile, individuals are friendly toward individuals who are different from them. On the other hand, the behavior in the group formation profile looks *as if* individuals prefer to interact only with similar others. With sufficient complementarities, only the latter profile maximizes the potential of the modified game.

The model of group formation contributes to a few strands of sociological literature. Social network theory distinguishes two types of interpersonal ties. The strong ties exhibit transitivity: if there are ties between individuals i and j and between i and k, then there is a tie between individuals j and k. A large amount of empirical literature document the fact that people choose friends of friends as their own friends more often than would happen if the friendships were chosen at random.[3] It is widely understood that strong ties enhance the existence of social norms.[4] The simplest argument attributes this to payoff complementarities: the emergence of social norms requires coordination, and its enforcement requires collective punishments which are possible only on dense networks. This paper shows that the payoff complementarities lead to the formation of friendship ties while abstracting from the exact source of complementarities. (On the other hand, weak ties seem to appear in the presence of payoff substitutes. In the job search example of Granovetter (1973), individual j can learn about information of individual k indirectly through contacts with i. This reduces j's utility of forming a direct tie with k.)

Structural balance theory discusses the connection between the transitivity of strong ties and group formation (see Wasserman and Faust (1994) for a review)[5]. Suppose that for all triples of individuals i, j, and k such that there are ties between pairs i and j as well as i and k, it must be that j and k are connected. Then, the set of individuals can be divided into disjoint sets, such that each individual is connected to all members of her own set, but there are no ties between sets. In this paper, we

[3]For early work on this topic, see Holland and Leinhardt (1970), Holland and Leinhardt (1976), Feld and Elmore (1982); Wasserman and Faust (1994) present an overview of the statistical methods involved in analyzing social networks. Backstrom, Huttenlocher, Kleinberg, and Lan (2006) find that the propensity of individuals to join online communities depends positively on the number of friends he or she has within the community and on whether these friends are mutual friends.

[4]For example, Granovetter (2004), Glaeser, Laibson, Scheinkman, and Soutter (2000), and Mobius and Szeidl (2007).

[5]I am grateful to David Easley for suggesting this connection.

allow for more nuanced differences in friendship attitudes, and extend the basic result to hierarchies of nested groups.

The literature has two approaches to homophily. In many situations (for example, in marriage), people have exogenous preferences for interacting with similar individuals (McPherson, Smith-Lovin, and Cook (2001)). Homophily can be a result of group selection: social groups that didn't develop preferences for in-group relationships simply dissolved and disappeared from society; it might be an outcome of transference of concepts and attitudes one cherishes toward one's kins to a wider social group (Horowitz (1985)); similar individuals tend to share similar backgrounds, which equips them with the same tastes, values, and attitudes, and which facilitates communication between them (Baccara and Yariv (2008)).

On the other hand, homophily becomes especially striking when there is no apparent connection between similarity and payoffs.[6] A series of papers argue that ethnic groups extend the threat of punishments from unilateral to punishments by the entire group, which enhances cooperation in long-term interactions (Greif (1993), Freitas (2007), Eeckhout (2006), see also Alesina and Ferrara (2005) for a review.) In the language of this paper, the group punishment can be attributed to higher degree complementarities in the punishment strategies: if player i trades with j and k, then i gains from j and k trading together, because such trade makes j fear punishment by k if j cheats on i.

[6]For example, sport fan riots in Byzantinian cities in the fifth and sixth century (Procopius (2007)), or genocide in Rwanda (Mamdani (2002)). In a series of classic experiments by Henri Tajfel, people acted as if they had preferences over a very arbitrary definition of similarity. Teenage boys from the same school and similar backgrounds were asked to estimate the number of dots flashed on the screen. Following their answers, the subjects were assigned one of two labels, "underestimators" and "overestimators." Next, the subjects were told to allocate small amounts of money between other individuals. To the surprise of the researchers, most of the subjects showed very strong bias toward individuals with the same label. The subjects either maximized the amount of money allocated to individuals with the same label (with a potential loss to the sum of payoffs in the whole population), or, in some cases, they maximized the total difference between the payoffs to both labels (with a potential loss to the sum of payoffs to their own label). See Tajfel (1970), Tajfel and Turner (1979), Tajfel, Billig, Bundy, and Flament (1971), and Haslam (2004).

The model of friendship choice is related to the network formation literature (for example, Jackson and Wolinsky (1996), Bala and Goyal (2000), Jackson and Watts (2002)). A choice of a friend corresponds to a unilateral formation of a directed link. As a contribution to this literature, this paper finds simple and natural conditions on payoffs that lead to the formation of novel network structures as hierarchies of groups or hierarchies of friends. Also, the paper is related to the matching literature. With no second degree complementarities, but no higher degree externalities, players' payoffs are independent of friendships (or lack of those) between other players. Typically, the matching literature focuses only on this case (many-to-many matching theory of Echenique and Oviedo (2006) seems to be the closest to our model).[7] Not surprisingly, our result about reciprocity in friendship is closely related to assortative matching in that literature.

Section 2 presents the model and the solution concept. Section 3 illustrates the role that different types of complementarities play in the group formation phenomenon. Section 4 analyzes the modified game, in which mistakes result from confusing the behavior toward similar individuals. The proofs can be found in the Appendix.

2. Model

There are N individuals with a typical individual denoted i, j. Each individual i chooses attitude $\lambda_{ij} \in \{1, ..., A\}$ that he feels toward any individual j, where A is a natural number. We interpret λ_{ij} as a measure of the strength of friendship with A being "the closest friendship" and 1 "the worst enmity." We say that individuals i and j are *a-friends* if $\lambda_{ij}, \lambda_{ji} \geq a$ for any $a \leq A$.

We fix natural numbers n_a such that $\sum_{a \leq A} n_a = N$, and we assume that for any individual i, there are exactly n_a individuals j such that $\lambda_{ij} = a$. Hence, each individual has at most $N_a = \sum_{a \leq a' \leq A} n_{a'}$ a-friends, and no individual can be friendly or hostile toward the entire population. . The assumption is natural in some situations. Consider the example of job search from the introduction. There, each individual chooses whom to inform about the employment opportunity first. Because there is exactly

[7] Exceptions are Sasaki and Toda (1996) and Hafalir (2008).

one person that will be informed first, there can be only one best friend. Similarly, because there is exactly one person to inform last, there is only one "worst enemy."

We assume that each individual i chooses attitude λ_{ii} toward himself. It is possible that player i chooses to be his own enemy, $\lambda_{ii} = 1$. Despite the inconvenient interpretation, we find that such an assumption simplifies the subsequent presentation. An alternative model in which players do not choose the attitude toward themselves becomes asymptotically equivalent to the current model when the size of the population becomes large, $N \to \infty$.

Let $\lambda_i = (\lambda_{ij})_j$ denote the strategy of player i. Let $\lambda = (\lambda_i)_{i \in N}$ denote the strategy profile, and let Λ be the set of all strategy profiles.

For any $k \geq 2$, a *k-interaction* is any sequence of k individuals $i_1, ..., i_k$. The individuals receive payoffs in interactions with other individuals. The payoff depends on the attitudes of the individuals in the interaction. Let $u^k : R^k \to R$ be a k times differentiable function. Player i_1-s payoff from k-interaction $i_1, ..., i_k$ in profile λ is equal to $u^k (\lambda_{i_1 i_2}, ..., \lambda_{i_k i_1})$.

For each $k \geq 2$, define the scaled sum of the payoffs of individual i in all k-interactions in which individual i participates:

$$U_i^k (\lambda) = \frac{1}{N^{k-1}} \sum_{i_1,...,i_k : i \in \{i_1,...,i_k\}} u^k (\lambda_{i_1 i_2}, ..., \lambda_{i_k i_1}) \tag{2.1}$$

Fix $K \geq 2$. Player i's payoff is defined as the sum of payoffs in all k-interactions for $k \leq K$:

$$U_i (\lambda) = \sum_{2 \leq k \leq K} U_i^k (\lambda).$$

For each $k \leq K$, define a discrete version of the partial derivative of function u^k with respect to all arguments[8]:

$$\Delta^k_{a_1 ... a_k} := \sum_{\delta_1 = 0,1} ... \sum_{\delta_k = 0,1} (-1)^{\delta_1 + ... + \delta_k} u^k (a_1 - \delta_1, ..., a_k - \delta_k). \tag{2.2}$$

[8]Notice that

$$\Delta^k_{a_1^* ,..., a_k^*} = \int_{a_1^* - 1}^{a_1^*} ... \int_{a_k^* - 1}^{a_k^*} \frac{\partial^k u^k}{\partial a_1 ... \partial a_k} (a_1', ..., a_k') \, da_1' ... da_k'$$

We say that payoffs are *(strictly) kth-degree complementary*, if $k \leq K$ and $\Delta^k_{a_1...a_k} \geq (>)\, 0$ for all $a_1, ..., a_k$. This is a standard definition (see Monderer and Shapley (1996)). For example, the second degree complementarity implies that in any interaction between players i and j, the payoff from i being friendly toward increases with the attitude of j toward i. The third degree complementarity implies that in any interaction between players i, j, and k, the difference between the benefit of i from being more friendly toward j if j is more friendly toward k and if j is less friendly toward k increases if k is more friendly toward i. Also, we say that payoffs *do not exhibit kth-degree externalities* if $\Delta^k_{a_1...a_k} = 0$ for all $a_1, ..., a_k$.

Typically, the above game has multiple equilibria. We focus on a subclass of equilibria with attractive equilibrium selection properties. For each strategy profile λ, define

$$V^k(\lambda) = \frac{1}{N^{k-1}} \sum_{i_1,...,i_k} u^k(\lambda_{i_1 i_2}, ..., \lambda_{i_k i}), \tag{2.3}$$

$$V(\lambda) := \sum_{2 \leq k \leq K} V^k(\lambda), \tag{2.4}$$

Here, $V^k(\lambda)$ is the sum of payoffs in all k-interactions, and V is the sum of payoffs in k-interactions for all $2 \leq k \leq K$. One checks that V is a *potential function* of Monderer and Shapley (1996): For each individual i, any profile λ, any two individual i's strategies λ_i, λ'_i,

$$u_i(\lambda_i, \lambda_{-i}) - u_i(\lambda'_i, \lambda_{-i}) = V(\lambda_i, \lambda_{-i}) - V(\lambda'_i, \lambda_{-i}).$$

Let $\Lambda_{\max}$ be the set of profiles that maximize the potential function. By Monderer and Shapley (1996), if $\lambda \in \Lambda_{\max}$, then λ is an equilibrium. Such equilibria are robust to incomplete information Ui (2001). Also, they have an evolutionary motivation: Blume (1993) shows that the potential maximizing equilibria are the only stochastically stable outcomes under payoff-dependent dynamics (see also Young (1998) and Hofbauer and Sandholm (2002)).

The potential maximization is closely related to the maximization of average payoffs. Fix $k \leq K$. Then, the average sum of payoffs in all k-interactions is proportional to the average k-interaction payoffs across all players plus a term that disappears with the size of the population,

$$\frac{1}{N}V^{k}\left(\lambda\right)=\frac{1}{k}\frac{1}{N}\sum\nolimits_{i}U_{i}^{k}\left(\lambda\right)+O\left(\frac{k}{N}\right). \tag{2.5}$$

This is because the payoffs in most k interactions $i_1, ..., i_k$ affect each player $i_1, ..., i_k$ equally. The last term on the right-hand side appears because some k-interactions include fewer than k different players and the total number of all such interactions is of order kN^{k-2}. It turns out that all results below that maximize potential $V(\lambda)$, uniformly maximize each of the terms $V^k(\lambda)$. Together with equation (2.5), this implies that each profile that maximizes potential maximizes the average payoffs up to term $O\left(\frac{k}{N}\right)$.

3. Complementarities and group formation

A set of individuals $I \subseteq N$ is a *group of a-friends* in profile λ, if for all individuals players $i, j \in I$ and $k \notin I$, $\lambda_{ij} \geq a$ and $\lambda_{ik} < a$. In a group of friends, individual choices are perfectly correlated: All members of a group of friends are a-friends of all other members. Also, if i is a member of a group, and if she has attitude at least a toward j, then j belongs to the same group. Hence, each group of friends consists of exactly N_a individuals. In this section, we analyze the relationship between complementarities and the emergence of groups of friends.

3.1. Correlation in players' choices.

The following numerical characteristic of profile λ is useful. For every k, every $a_1, ..., a_k$, define the number of k-tuples of individuals connected with attitudes, respectively, at least $a_1, ...,$ and a_k :

$$S_{a_1...a_k}^{k}\left(\lambda\right)=\left|\left\{\left(i_1,...,i_k\right):\lambda_{i_1i_2}\geq a_1,...,\lambda_{i_ki_1}\geq a_k\right\}\right|.$$

We argue that $S_{a_1...a_k}^{k}(\lambda)$ is a measure of the correlation between the individual choices. To see it, consider first the extreme case of the complete lack of correlation. Assume that each individual i chooses her strategy independently from the uniform distribution μ on the set of all strategies. The probability that the attitude of player i toward player j is not smaller than a is equal to $\frac{N_a}{N}$. The expected value of the number of k-tuples with each pair connected with attitude not smaller than a is equal to

$$E_\mu S^k_{a\ldots a}(\lambda) = N^k \left(\frac{N_a}{N}\right)^k = N_a^k,$$

where the expectation is taken over the uniform distribution $\mu \in \Delta\Lambda$.

On the other extreme, suppose that there exists a profile λ_a such that each player belongs to a group of a-friends. (Later, we discuss the existence of such profiles.) Thus, players' choices are completely correlated, and only k-tuples of individuals in the same group are connected with attitudes at least a. It is easy to see that

$$S^k_{a\ldots a}(\lambda_a) = N(N_a)^{k-1} = \frac{N}{N_a} N_a^k > E_\mu S^k_{a\ldots a}(\lambda).$$

Compute an upper bound on $S_{a_1\ldots a_k}(\lambda)$. Suppose that $a_k = \min(a_1, ..., a_k)$. There are at most N ways of choosing individual i_1; because individual i_1 has at most N_{a_1} a_1-friends, there are at most N_{a_1} ways of choosing individual i_2; ... ;because individual i_{k-1} has at most $N_{a_{k-1}}$ a_{k-1}-friends, there are at most $N_{a_{k-1}}$ ways of choosing individual i_k. Therefore, $S^k_{a_1\ldots a_k}(\lambda) \leq N N_{a_1} ... N_{a_{k-1}}$. More generally, for all $a_1, ..., a_k$,

$$S^k_{a_1\ldots a_k}(\lambda) \leq N \frac{N_{a_1} ... N_{a_k}}{N_{\min(a_1,...,a_k)}}. \tag{3.1}$$

Simple algebra leads to the following result. Suppose that payoffs are strictly kth-degree complementary for all k. Then, the potential of a profile increases if the players' choices become more correlated.

Lemma 1. *For any $k \geq 2$, any profile λ,*

$$V^k(\lambda) = \text{const} + \sum_{a_1,...,a_k \geq 2} \Delta^k_{a_1\ldots a_k} S^k_{a_1\ldots a_k}(\lambda), \tag{3.2}$$

where const *does not depend on profile λ.*

3.2. **Reciprocity in friendship.** Profile λ *reciprocates friendship* if $\lambda_{ij} = \lambda_{ji}$ for each pair of individuals i and j. Profiles that reciprocate friendship exist if, for example, the size of each category of friends n_a is even.[9] Let Λ_f be the set of all profiles that reciprocate friendship.

Suppose that there are profiles that reciprocate friendship. The next result shows that with second degree complementarities, but no other externalities, a profile maximizes potential if and only if the profile reciprocates friendship.

Proposition 1. *Suppose that $\Lambda_f \neq \varnothing$, payoffs are strictly second degree complementary, but they do not exhibit kth dergee externalities for any $k \geq 3$. Then, $\Lambda_{\max} = \Lambda_f$.*

Proof. It is easy to check that for every profile λ, if $S^2_{a_1 a_2}(\lambda) = N N_{\max(a_1,a_2)}$ for each a_1, a_2, then λ reciprocates friendship. Hence, $\arg\max V^2(\lambda) = \Lambda_f$.

For every $k \geq 3$, if payoffs do not exhibit kth dergee externalities, then $\Delta^k_{a_1 \ldots a_k} = 0$ for all $a_1, \ldots, a_k$. By Lemma 2.4, $V^k(\lambda)$ does not depend on profile λ. Hence, $\Lambda_{\max} = \arg\max V^2(\lambda) = \Lambda_f$. □

Proposition 1 is related to the pairwise stability of the assortative matching under supermodularity (Becker (1973)). Typically, the matching literature assumes that the payoffs in an interaction do not depend on interactions between other members of the population.[10] With no third or higher degree complementarities, our model has the same property.

3.3. **Hierarchies of groups.** Consider profile λ such that for each a, each individual i belongs to a group of a-friends. Because each group of a-friends has N_a members, there are exactly $\frac{N}{N_a}$ groups of a-friends. Clearly, any such profile reciprocates friendship. Since the set of a'-friends is included in the set of a-friends for $a < a'$, each group of a-friends contains $\frac{N_a}{N_{a+1}}$ groups of $(a+1)$-friends, and the groups form a hierarchy

[9]Suppose that n_a is even for each $A \leq A$. Arrange players $i = 1, \ldots, N$ on a circle and let $d(i,j)$ be the distance between players i and j (the length of the shortest arc between i and j). For each $a \leq A$, let $\lambda_{ij} = a$ for all players i and j such that $\frac{N_{a-1}}{2} < d(i,j) \leq \frac{N_a}{2}$. Because the distance is symmetric, profile λ returns friendship.

[10]For example, Echenique and Oviedo (2006) consider a model of many-to-many matching in which the agents' preferences for matchings do not depend on matchings of other individuals.

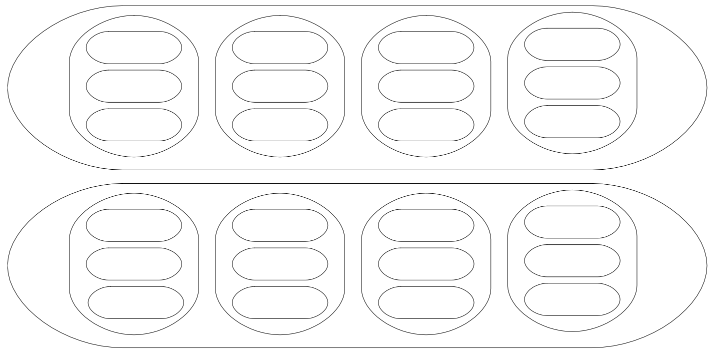

FIGURE 1. Hierarchy of groups.

of nested groups of better and better friends. This is illustrated in Figure 1. We refer to any such profile as a *hierarchy of groups.*

A hierarchy of groups exists if the sizes of categories of friends are appropriately divisible. More precisely, it exists if and only if for each a, n_a is a multiple of N_{a+1}. In such case, each hierarchy of groups divides society into $\frac{N}{N_a}$ groups of a-friends for each a. Let Λ_{hg} be the set of all hierarchies of groups.

Proposition 2. *Suppose that* $\Lambda_{hg} \neq \varnothing$*,* $K \geq 3$*, and payoffs are strictly* k*th-degree complementary for any* $2 \leq k \leq K$*. Then,* $\Lambda_{\max} = \Lambda_{hg}$*.*

There is a sharp contrast between second and higher degree complementarities. The existence of a group of friends is not necessary to maximize the potential with second degree complementarities (Proposition 1) . It becomes necessary when the payoffs are second and third degree complementary (Proposition 2). An addition of even higher degree complementarities of a higher degree does not change the result.

We need the following partial result.

Lemma 2. *Suppose that* $\Lambda_{hg} \neq \varnothing$*. For any* k*, if payoffs are* k*th degree complementary, then* $\Lambda_{hg} \subseteq \arg\max V^k(\lambda) \cap \Lambda_f$*. For any odd* k*, if payoffs are strictly* k*th degree complementary, then* $\Lambda_{hg} = \arg\max V^k(\lambda) \cap \Lambda_f$*.*

The first part of Lemma 2 is a consequence of Lemma 2.4 and the fact that inequality (3.1) turns into equality for any hierarchy of groups. (Observe that the right-hand side of (3.1) is equal to the number of all tuples of individuals inside groups of a-friends, where $a = \min(a_1, ..., a_k)$.) For the second part of Lemma 2, we show in the Appendix that for any profile that reciprocates friendship $\lambda \in \Lambda_f$, if k is odd and for each a, $S^k_{a...a}(\lambda) = N(N_a)^{k-1}$, then λ is a hierarchy of groups. To see the intuition behind this result, consider a simple example. There are two divisible categories, $A = 2$ and $N = 2N_2$. Category 2 corresponds to friends, and category 1 consists of enemies. Consider the payoffs in k-interactions for odd $k = 3$. If λ is a hierarchy of groups, then each two friends of an individual are also friends of each other, and the number of triples of mutual friends is equal to $S^3_{222}(\lambda) = N(N_2)^2$. If λ' is not a hierarchy of groups, then there is individual i such that not all of her friends are friends of each other, and the number of triples of individuals (i, i_2, i_3) such that all individuals in the triple are friends is strictly smaller than $(N_2)^2$. For any other individual i', the number of triples of individuals (i', i_2, i_3) such that all individuals in the triple are friends is not larger than the number of all pairs of friends of individual i', $(N_2)^2$. Hence, $S^3_{222}(\lambda') < N(N_2)^2$.

Proof of Proposition 2. As in the proof of Proposition 1, we show that $\arg\max V^2(\lambda) = \Lambda_f$. The result follows from Lemmas 1 and 2. □

3.4. **Strong complementarities and hierarchies of friends.** The existence of hierarchies of groups rely on restrictive divisibility assumptions. In this section, we show a different version of the group formation result with stronger payoff assumptions, but without divisibility restrictions.

Assume that each category of friends consists of exactly one individual, i.e., $n_a = 1$ for each a and $A = N$. Suppose that $N = 2^{\log_2^N}$, where $\log_2 N$, is a natural number. Player i's strategy induces a strict preference ordering of all individuals in the population.

Take any bijection $\sigma : \{1, ..., N\} \to \{0, 1\}^{\log_2 N}$. We interpret σ as an assignment of unique binary IDs to each individual. Construct profile λ_δ in which each individual i has relative lexicographic preferences over all members of society, relative to IDs σ.

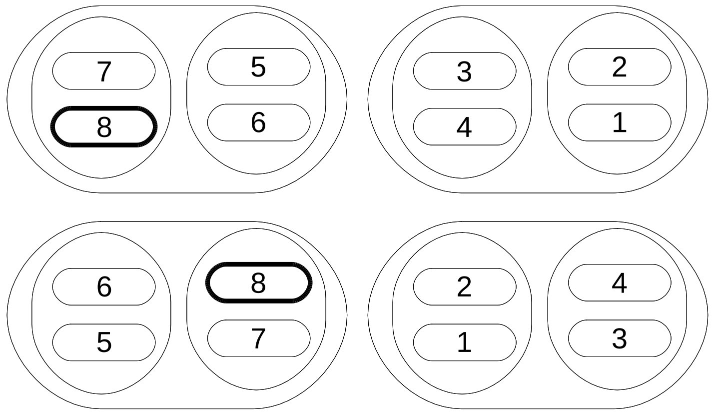

FIGURE 2. Two strategies in a hierarchy of friends (upper and lower part of the picture). Each strategy describes the attitudes of the individual encircled by the bold line toward all individuals in the population.

Precisely, for any j and k, i is more friendly toward j than toward k if

$$\sum_{n\leq\log_2 N} 2^{-n}\left|\sigma_n(i)-\sigma(j)\right| < \sum_{n\leq\log_2 N} 2^{-n}\left|\sigma_n(i)-\sigma(k)\right|. \tag{3.3}$$

For example, i prefers any individual j such that $\sigma_1(j)=\sigma_1(i)$ to any individual k such that $\sigma_1(i)\neq\sigma_1(k)$. Profile λ_δ is called a *hierarchy of friends.* Let Λ_{hf} denote the set of all hierarchies of friends.

In a hierarchy of friends, each individual is a member of a group of 2^l-friends for each $l\leq\log_2 N$. To see it, notice that there are exactly $\frac{N}{2}$ individuals j such that $\sigma_1(j)=\sigma_1(i)$. If each j chooses his friends using formula (3.3), the first coordinate of σ divides society into two equal groups of $\frac{N}{2}$-friends. Next, for any two individuals j and k such that $\sigma_1(j)=\sigma_1(k)$, i prefers j if $\sigma_2(j)=\sigma_2(i)$ and $\sigma_2(i)\neq\sigma_2(k)$. There are exactly $\frac{N}{4}$ individuals j such that $\sigma_m(j)=\sigma_m(i)$ for $m=1,2$, and the first two coordinates of σ divide society into 4 equal-size groups of $\frac{N}{4}$-friends. Similarly, the first l coordinates divide society into 2^l equally sized groups of $2^{\log_2 N-l}$-friends. See Figure 2.

Additionally, individuals have strict preference ordering over members of the other group. This is opposite to a hierarchy of groups, in which two members of a different group are treated in the same way.

The next result establishes the conditions under which the set of potential maximizing profiles is equal to the set of hierarchies of friends.

Proposition 3. *Suppose that* $\Lambda_{hf} \neq \varnothing$, $K \geq 3$, *and for each* k, *there exists* $\theta_k \geq k$ *and* $\psi_k > 0$ *such that* $u^k(a_1, ..., a_k) = \psi_k \theta_k^{a_1+...+a_k}$ *Then* $\Lambda_{\max} = \Lambda_{hf}$.

4. Preferences for similarity

4.1. **Similarity.** In the previous section, we show that sufficient complementarities lead to the formation of groups. However, none of the results sheds any light on the composition of groups. In fact, if $\Lambda_{hg} \neq \varnothing$, then any set of N_a individuals is a group of a-friends in some hierarchy of groups.

In order to discuss a composition of groups, we introduce some heterogeneity into the population of so far exchangeable individuals. We assume that some players are more similar to each other than to the rest. Let $d(i, j) \in [0, 1]$ denote the similarity between individuals i and j.

We interpret similarity $d(i, j)$ as a measure of how difficult it is to tell individuals i and j apart. For example, suppose that each individual is described by a certain number of attributes such as height, color of skin, facial hair, religion, wealth, etc.. Suppose that $d(i, j)$ is equal to the number of attributes shared by individuals i and j. If individuals" attitudes toward each other depend on their attributes, then two individuals with similar attributes are easier to confuse than individuals with distinct attributes.

Fix profile λ. For each a, define the measure of similarity between individuals connected with attitude at least a :

$$D_a(\lambda) := \sum_{j,j':\lambda_{jj'_i} \geq a} d(j, j'). \tag{4.1}$$

Intuitively, if $D_a(\lambda)$ is larger, then close friends are more similar to each other. Notice that $D_1(\lambda)$ is equal to the sum of similarities across all players, and it does not depend on the profile λ.

4.2. **Game with mistakes.** We consider a modification of the game from section 2. In the modification, individuals may mistake individual i for j, and the probability of mistake is proportional to $d(i,j)$.

Fix $c \in \left(0, \frac{1}{N}\right)$. For each i and $j \neq j'$, individual i makes a mistake and treats individual j with attitude$\lambda_{ij'}$ with probability $p^c_{jj'} = cd(j,j')$; with the remaining probability $p^c_{jj} = 1 - \sum_{j' \neq j} p^c_{jj'}$, she treats individual j with the correct attitude λ_{ij}. Parameter $c > 0$ scales the importance of mistakes. The payoffs in the game with mistakes are equal to the expected payoffs in the original game with expectations taken with respect to the distribution of mistakes:

$$U_i^{c,k}(\lambda) = \frac{1}{N^{k-1}} \sum_{i_1,i_2,\ldots,i_k : i \in \{i_1,\ldots,i_k\}} \sum_{j_1,\ldots,j_n} u^k\left(\lambda_{i_1 j_2}, \lambda_{i_2 j_3} \ldots, \lambda_{i_k j_1}\right) p^c_{i_2 j_2} p^c_{i_3 j_3} \ldots p^c_{i_1 j_1},$$

$$U_i^c(\lambda) = \sum_{2 \leq k \leq \infty} U_i^{c,k}(\lambda),$$

The game with mistakes has a potential function $V^c(\lambda)$ defined as:

$$V^{c,k}(\lambda) = \frac{1}{N^{k-1}} \sum_{i_1,i_2,\ldots,i_k} \sum_{j_1,\ldots,j_n} u^k\left(\lambda_{i_1 j_2}, \lambda_{i_2 j_3} \ldots, \lambda_{i_k j_1}\right) p^c_{i_2 j_2} p^c_{i_3 j_3} \ldots p^c_{i_1 j_1},$$

$$V^c(\lambda) := \sum_{2 \leq k \leq K} V^{c,k}(\lambda).$$

Let $\Lambda^c_{\max}$ denote the set of profiles that maximize potential.

It is instructive to relate the potential of the modified game to the potential of the original game minus the loss from mistakes. Let

$$C^k(\lambda) = \frac{1}{N^{k-1}} \sum_{i_1,i_2,\ldots,i_k} \sum_{\hat{k}=1}^{k} \sum_{j_l} \left(\begin{array}{c} u^k\left(\lambda_{i_1 j_2}, \ldots, \lambda_{i_{\hat{k}-1} i_{\hat{k}}}, \ldots, \lambda_{i_k j_1}\right) \\ -u^k\left(\lambda_{i_1 j_2}, \ldots, \lambda_{i_{\hat{k}-1} j_{\hat{k}}}, \ldots, \lambda_{i_k j_1}\right) \end{array} \right) d(i_l, j_l),$$

be the weighted sum of losses from one mistake in each interaction, where the weights are proportional to the probability of committing a mistake. For small $c > 0$, the sum of payoffs in k-interactions in the game with mistakes is equal to

$$V^{c,k}(\lambda) = (1 - ck) V^k(\lambda) - cC^k(\lambda) - O\left(c^2\right). \tag{4.2}$$

The first term is equal to the sum of payoffs in k-interactions in the original game, the second term is equal to the sum of losses from one mistake per interaction, and the last term consists of other losses from mistakes. The next result shows that, for sufficiently small c, each profile that maximizes the potential of the modified game minimizes the cost of mistakes across the profiles that maximize the potential of the original game.

Lemma 3. *There exists $c^* > 0$ such that for each $c \in (0, c^*)$,*

$$\Lambda^c_{\max} \subseteq \arg\min_{\lambda \in \Lambda_{\max}} C^k(\lambda).$$

Proof. This is an immediate consequence of (4.2). □

4.3. **Preferences for similarity.** The main result of this section shows that with sufficient complementarities, the cost of mistakes decreases when friends become more similar:

Proposition 4. *Suppose that $K \geq 3$, and either*

- *$\Lambda_{hg} \neq \varnothing$, and payoffs are strictly k-complementary for each $k \leq K$, or*
- *$\Lambda_{hf} \neq \varnothing$, and for each k, there exists $\theta_k \geq k$ and $\psi_k > 0$ such that $u^k(a_1, ..., a_k) = \psi_k \theta_k^{a_1 + ... + a_k}$.*

Then, there exist (profile-independent) constants $c^k_a > 0$ such that for each $\lambda \in \Lambda_{\max}$, for each $k \leq K$,

$$C^k(\lambda) = -\sum\nolimits_{a \leq A} c^k_a D_a(\lambda).$$

Together with Lemma 3, the Proposition says that, under appropriate assumptions, if a profile maximizes the potential of the modified game, then it maximizes the potential of the original game *and* a weighted average of the similarities between friends.

For small c, the direct cost of mistakes is small relative to the payoffs from interactions. However, in a potential maximizing equilibrium, there are substantial payoff differences between befriending people who are similar or different. From the point of view of an analyst who observes the choices and payoffs, the outcome may look *as if* individuals have strong preferences for interacting with similar others. However, the preference for similarity is an equilibrium effect: If everybody makes friends with similar others, the payoff complementarities push each individual to choose similar

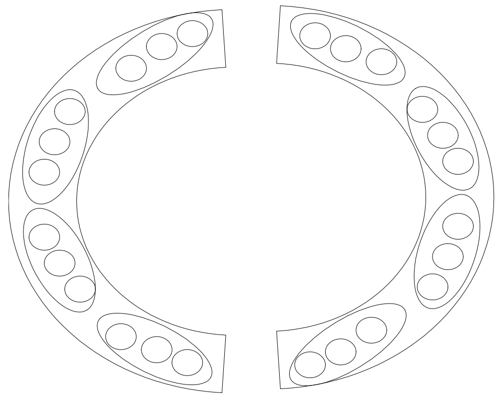

FIGURE 3. Hierarchy of groups on a circle

friends. Thus, one should be careful not to overestimate the strength of preferences toward similarity.

We illustrate the Proposition with an example of a similarity-maximizing hierarchy of groups. Suppose that N individuals are located on a circle. The similarity between two individuals i and j decreases with the length of the shorter arc between i and j :

$$d_c(j,k) = \min\left(|j-k|, |N+j-k \bmod N|\right).$$

An *interval* $I \subseteq \{1,...,N\}$ is a set of consecutive individuals

Corollary 1. *Suppose that $\Lambda_{hg} \neq \varnothing$, $K \geq 3$, and payoffs are strictly k-complementary for each $k \leq K$. Then, for sufficiently small c, $\lambda \in \Lambda^c_{\max}$ if and only if λ is a hierarchy of groups such that for each \a $\leq A$, each group of a-friends is an interval.*

See Figure 3.

Proof. Take any set of individuals $I' \subseteq I$, $|I'| \leq \frac{1}{2}|I|$. Then, $\sum_{j,j':j\neq j'} d_c(j,j') \leq 2\sum_{i=1}^{|I'|} iv(|I'|-i)$ with equality if and only if I' is a set of consecutive individuals. Recall that any hierarchy of groups divides society into $\frac{N}{N_a}$ disjoint groups of a-friends. Thus, a hierarchy of groups λ maximizes the potential of the modified game only if each group of a-friends consists of consecutive individuals. □

4.4. **Proof of Proposition 4.** Assume that $\Lambda_{hg} \neq \varnothing$, and payoffs are strictly k-complementary for each $k \leq K$. The proof in the other case is analogous, and we point to the only difference in the end of this section.

Take any two individuals i and i'. Define the average costs of mistaking individual i for individual i' as

$$C^k(\lambda; i, i') := \frac{1}{N^{k-1}} \sum_{\hat{k}=1}^{k} \sum_{i_1, i_2, \ldots, i_k : i_l = i} \left(\begin{array}{c} u^k\left(\lambda_{i_1 j_2}, \ldots, \lambda_{i_{\hat{k}-1} i}, \ldots, \lambda_{i_k i_1}\right) \\ -u^k\left(\lambda_{i_1 j_2}, \ldots, \lambda_{i_{\hat{k}-1} i'}, \ldots, \lambda_{i_k i_1}\right) \end{array} \right).$$

We have the following result:

Lemma 4. *Suppose that $\lambda \in \Lambda_{hg}$. For any pair of individuals i, i',*

$$C^k(\lambda; i, i') = \sum_{a_1, \ldots, a_l : a_l > \lambda_{ii'} \text{ for each } l} \frac{k N_{a_1} \ldots N_{a_k}}{N^{k-1} N_{\min(a_1, \ldots, a_k)}} \Delta^k_{a_1 \ldots a_k},$$

In particular, the cost of mistakes in λ depends only on the attitude between players i and i', and it decreases with $\lambda_{ii'}$.

The fact that the cost of mistake $C^k(\lambda; i, i')$ depends only on the attitude between players i and i' is a consequence of a simple symmetry argument. Let $C^k(a) := C^k(\lambda; i, i')$ for any profile λ and any two players i and i' such that $\lambda_{ii'} = a$.

The fact that the cost of a mistake decreases with the attitude is a consequence of the group formation phenomenon. Take any two players i and i' that are close friends in the hierarchy of groups λ. Because of group formation, for each player j, it is either that i and i' are friends of j or i and i' are enemies of j. In other words, the attitude between j and i is close to the attitude between j and i'. On the other hand, j's cost of mistaking two players increases with the difference in intended attitudes. If the attitudes toward i and i' are close, the cost of mistaking i for i' is small. Hence, the cost of mistaking two mutual friends is small. A similar argument shows that the cost of mistaking two mutual enemies is large.

Define

$$c^k_a := C^k(\lambda; a-1) - C^k(\lambda; a) > 0.$$

Then, we can rewrite the cost of mistakes $C^k(\lambda)$ as

$$\begin{aligned} C^k(\lambda) &= \sum_a \sum_{i,i':\lambda_{ii'}=a} d(i,j) \frac{1}{N^{k-1}} \sum_{\hat{k}=1}^{k} \sum_{i_1,i_2,\dots,i_k:i_l=i} \begin{pmatrix} u^k\left(\lambda_{i_1 j_2},\dots,\lambda_{i_{\hat{k}-1} i},\dots,\lambda_{i_k j_1}\right) \\ -u^k\left(\lambda_{i_1 j_2},\dots,\lambda_{i_{\hat{k}-1} i'},\dots,\lambda_{i_k j_1}\right) \end{pmatrix} \\ &= \sum_a \sum_{i,j:\lambda_{ij}=a} d(i,j)\, C^k(\lambda; i,j) \\ &= \sum_a \left(D_a(\lambda) - D_{a+1}(\lambda)\right) C^k(a) \\ &= \text{const} - \sum_{a\geq 2} c_a^k D_a(\lambda), \end{aligned}$$

where const $= D_1(\lambda)\, C^k(1)$ does not depend on profile λ. This ends the proof of the Lemma.

Assume that $\Lambda_{hf} \neq \varnothing$, and for each k, there exists $\theta_k \geq k$ and $\psi_k > 0$ such that $u^k(a_1,\dots,a_k) = \psi_k \theta_k^{a_1+\dots+a_k}$. The next result is analogous to Lemma 4.

Lemma 5. *Suppose that* $\lambda \in \Lambda_{hg}$. *For any pairs of individuals* i_0, i_0', *any hierarchies of friends* λ *and* λ', *if* $\lambda_{i_0 i_1} \geq (>)\, \lambda'_{i_0' i_1'}$, *then* $C^k(\lambda; i_0, i_1) \leq (<)\, C^k(\lambda'; i_0', i_1')$.

The rest of the proof proceeds in an analogous way.

5. Discussion

This paper explores strategic explanations for group formation and preferences for similarity. We discuss a game in which players choose their friends. We distinguish between different types of complementarities in players' choices. second degree complementarity pushes individuals toward reciprocating friendships. We show that third and higher degree complementarities push individuals to form groups. Next, we modify the original game by introducing the possibility of mistakes. The cost of mistakes is minimized by profiles that keep friends similar to each other and enemies similar to each other. With enough complementarities and group formation, similarity coordinates group's choices so that the behavior of society can be interpreted *as if* individuals have preferences toward similar others.

The main observation of the paper is that the apparent similarity-oriented behavior does not necessarily mean that people have intrinsic preferences for similarity. On the contrary, such behavior might be an equilibrium consequence of strong preferences toward group formation and similarity-based coordination. This paper discusses a very specific way in which similarity affects coordination. There are many others. The strategy "always befriend people who are similar to you" has three advantages. First, it is "simple": One compares another individual to oneself to decide whether that individual is worthy of friendship. There is no need to remember names or learn many details about social interactions in a new environment. Second, such a strategy facilitates coordination. If similarity is transitive, then choosing similar friends leads to group formation and to maximization of payoffs from third and higher degree complementarity. Finally, such a strategy is easily transferrable between various environments with different payoffs, populations of individuals, and, possibly, different meanings of similarity.

References

Alesina, A., and E. L. Ferrara (2005): "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43, 721–61.

Baccara, M. G., and L. Yariv (2008): "Similarity and Polarization in Groups," .

Backstrom, L., D. Huttenlocher, J. Kleinberg, and X. Lan (2006): "Group Formation in Large Social Networks: Membership, Growth, and Evolution," *Proc. 12th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining.*

Bala, V., and S. Goyal (2000): "A Noncooperative Model of Network Formation," *Econometrica*, 68, 1181–1229.

Barbera, S., M. Maschler, and J. Shalev (2001): "Voting for Voters: A Model of Electoral Evolution," *Games and Economic Behavior*, 37, 40–78.

Becker, G. (1973): "A Theory of Marriage: Part," *Journal of Political Economy*, 81, 813–846.

Becker, G. S. (1971): *The Economics of Discrimination.* University of Chicago Press, Chicago.

Blume, L. E. (1993): "The Statistical-Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5(3), 387–424.

Echenique, F., and J. Oviedo (2006): "A Theory of Stability in Many-To-Many Matching Markets," *Theoretical Economics*, 1, 233–273.

Eeckhout, J. (2006): "Minorities and Endogenous Segregation," *Review of Economic Studies*, 254, 31–53.

Feld, S. L., and R. Elmore (1982): "Patterns of Sociometric Choices: Transitivity Reconsidered," *Social Psychology Quarterly*, 45, 77–85.

Freitas, K. (2007): "The Indian Caste System as a Means of Contact Enforcement," UT Austin.

Glaeser, E. L., D. I. Laibson, J. A. Scheinkman, and C. L. Soutter (2000): "Measuring Trust," *Quarterly Journal of Economics*, 115, 811–846.

Granovetter, M. (1973): "The Strength of Weak Ties," *American Journal of Sociology*, 78, 1360–1380.

Granovetter, M. D. (2004): "The Impact of Social Structures on Economic Development," *Journal of Economic Perspectives*, 19, 33–50.

Greif, A. (1993): "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders Coalition," *American Economic Review*, pp. 525–48.

Hafalir, I. E. (2008): "Stability of Marriage with Externalities," *International Journal of Game Theory.*

Haslam, S. (2004): *Psychology in Organizations.* Sage Publications, Thousand Oaks, CA.

Hofbauer, J., and W. H. Sandholm (2002): "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, 70, 2265–2294.

Holland, P. W., and S. Leinhardt (1970): "A Method for Detecting Structure in Sociometric Data," *American Journal of Sociology*, 76, 492–513.

——— (1976): "Local Structure in Social Networks," *Sociological Methodology*, 7, 1–45.

Horowitz, D. L. (1985): *Ethnic Groups in Conflict*. University of California Press, Berkeley, California.

Jackson, M. O., and A. Calvó-Armengol (2004): "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 94, 426–454.

Jackson, M. O., and A. Watts (2002): "On the formation of interaction networks in social coordination games," *Games and Economic Behavior*, 41(2), 265–291.

Jackson, M. O., and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.

Mamdani, M. (2002): *When Victims Become Killers: Colonialism, Nativism, and the Genocide in Rwanda*. Princeton University Press, Princeton.

McPherson, M., L. Smith-Lovin, and J. Cook (2001): "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 27, 415–44.

Mobius, M., and A. Szeidl (2007): "Trust and Social Collateral," http://www.econ.berkeley.edu/ szeidl/papers/socialcollateral.pdf.

Monderer, D., and L. S. Shapley (1996): "Potential Games," *Games and Economic Behavior*, 14, 124–143.

Motani, N. A. (1977): *On His Majesty's Service in Uganda: The Origins of Uganda's African Civil Service: 1912-40*. Syracuse University, Syracuse.

Procopius (2007): *History of Wars: Books 1-2 (Persian War)*. Cosimo Classics, New York.

Sasaki, H., and M. Toda (1996): "Two-Sided Matching Problems with Externalities," *Journal of Economic Theory*, 70, 93–108.

Schelling, T. S. (1971): "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1, 143–186.

Tajfel, H. (1970): "Experiments in Intergroup Discrimination," *Scientific American*, 223, 96–102.

Tajfel, H., M. Billig, R. P. Bundy, and C. Flament (1971): "Social Categorization and Intergroup Behaviour," *European Journal of Social Psychology*, 2, 149–178.

Tajfel, H., and J. Turner (1979): "An Integrative Theory of Intergroup Conflict," in *The Social Psychology of Intergroup Relations*, ed. by W. G. Austin, and

S. Worchel. Brooks-Cole, Monterey, CA.

Twaddle, M. (ed.) (1975): *Expulsion of a Minority : Essays on Ugandan Asians.* Athlone Press for the Institute of Commonwealth Studies, London.

Ui, T. (2001): "Robust Equilibria of Potential Games," *Econometrica*, 69, 1373–80.

Wasserman, S., and K. Faust (1994): *Social Network Analysis: Methods and Applications.* Cambridge University Press, New York.

Young, H. P. (1998): *Individual Strategy and Social Structure.* Princeton University Press, Princeton.

Appendix A. Proofs

A.1. **Proof of Lemma 1.** We start with two observations. First, for each $a_1, ..., a_k$,

$$\left|\{(i_1, ..., i_k) : \lambda_{i_1 i_2} = a_1, ..., \lambda_{i_k i_1} = a_k\}\right| = \sum_{\delta_1=0,1} ... \sum_{\delta_k=0,1} (-1)^{\delta_1+....+\delta_k} S^k_{a_1+\delta_1,...,a_k+\delta_k}(\lambda)$$

Moreover, notice that

$$\begin{aligned} V^k(\lambda) &= \frac{1}{N^{k-1}} \sum_{a_1,...,a_k} u^k(a_1, ..., a_k) \left|\{(i_1, ..., i_k) : \lambda_{i_1 i_2} = a_1, ..., \lambda_{i_k i_1} = a_k\}\right| \\ &= \frac{1}{N^{k-1}} \sum_{a_1,...,a_k \geq 1} u^k(a_1, ..., a_k) \sum_{\delta_1=0,1} ... \sum_{\delta_k=0,1} (-1)^{\delta_1+....+\delta_k} S^k_{a_1+\delta_1,...,a_k+\delta_k}(\lambda) \\ &= \sum_{a_1,...,a_k \geq 2} \Delta^k_{a_1...a_k} S^k_{a_1...a_k}(\lambda) \\ &+ \sum_{a_1,...,a_k : \min(a_1,...,a_k)=1} S^k_{a_1...a_k}(\lambda) A_{a_1...a_k}, \end{aligned}$$

for some constants $A_{a_1...a_k}$. However, notice that $S^k_{a_1...a_k}(\lambda) = N_{a_1}...N_{a_k}$ for each $a_1, ..., a_k$ such that $\min(a_1, ..., a_k) = 1$ and that it does not depend on profile λ.

A.2. **Proof of Lemma 2.** Assume that k is odd and λ is a profile that reciprocates friendship and such that for each a, $S^k_{a...a}(\lambda) = N(N_a)^{k-1}$. We show here that λ is a hierarchy of groups.

For each individual i_1, denote the set of all sequences of k individuals that begin with individual i_1 and such that each individual in the sequence has attitude a toward the next individual:

$$S_{i_1} = \left\{(i_1, ..., i_k) : \lambda_{i_1 i_2}, ..., \lambda_{i_{k-1} i_k}, \lambda_{i_k i_1} \geq a\right\}.$$

For each $l \leq k$, denote also

$$\begin{aligned} S_{i_1}^{l} &= \{i_l : (i_1, ..., i_k) \in S_i\}, \\ S_{i_1}^{2,...,l} &= \{(i_2, ..., i_l) : (i_1, ..., i_k) \in S_i\}, \\ S_{i_1}^{l,...,k} &= \{(i_l, ..., i_k) : (i_1, ..., i_k) \in S_i\}. \end{aligned}$$

Because λ reciprocates friendship,

$$S_{i_1}^{2} = \{i' : \lambda_{ii'} \geq a\} = S_{i_1}^{k}. \tag{A.1}$$

I will show that for each $2 < l < k$,

$$S_{i_1}^{l-1} = S_{i_1}^{l+1}. \tag{A.2}$$

Indeed, for any $1 < l < k$,

$$S_{a...a}(\lambda) = \sum_{i_1} \sum_{(i_2,...,i_l) \in S_{i_1}^{2,...,l}} \sum_{i_{l+1} \in S_{i_1}^{l+1} : \lambda_{i_l i_{l+1}} \geq a} \left|\left\{(i_{l+2}, ..., i_k) : (i_{l+1}, i_{l+2}, ..., i_k) \in S_{i_1}^{l+1,...,k}\right\}\right|,$$

Because $S_i^{2,...,l} \leq (N_a)^{l-1}$, for each $i_{l+1} \in S_{i_1}^{l+1}$

$$k\left|\left\{(i_{l+2}, ..., i_k) : (i_{l+1}, i_{l+2}, ..., i_k) \in S_{i_1}^{l+1,...,k}\right\}\right| \leq (N_a)^{k-l-1}.$$

Hence, if $S_{a...a}(\lambda) = NN_a^{k-1}$, then for each individual i_1, for each $1 < l < k$, $\left|S_{i_1}^{2,...,l}\right| = N_a^{l-1}$, and for each individual $i_l \in S_{i_1}^{l}$, there are exactly N_a individuals $i_{l+1} \in S_{i_1}^{l+1}$ such that $\lambda_{i_l i_{l+1}} \geq a$. This further means that for each $i_l \in S_{i_1}^{l}$, all a-friends of individual i_l belong to set $S_{i_1}^{l+1}$. A symmetric argument shows that for each $1 < l < k$, for each $i_{l+1} \in S_{i_1}^{l}$, all a-friends of individual i_{l+1} belong to set $S_{i_1}^{l}$. Because λ reciprocates friendship, this yields (A.2).

If k is odd, (A.1) implies that

$$S_{i_1}^{2} = S_{i_1}^{k} = S_{i_1}^{k-2} = ... = S_{i_1}^{3}.$$

Thus, $S_{i_1}^{2} = S_{i_1}^{3}$, and the set of a-friends of individual i_1 is equal to the set of a-friends of a-friends of individual i_1. Since this is true for all i_1 and all a, λ is a hierarchy of groups.

A.3. **Proof of Lemma 4.** Let λ be a hierarchy of groups.

For all tuples $a^0, a^1 \in R^k$, $a^0_l \geq a^1_l$, denote a measure of kth-degree complementarity,

$$\Delta^{a^0}_{a^1} = \sum_{\delta_1=0,1} ... \sum_{\delta_k=0,1} (-1)^{\delta_1+....+\delta_k} u^k\left(a_1^{\delta_1}, ..., a_k^{\delta_k}\right) = \sum_{a_1,...,a_l: a^1_l < a_l \leq a^0_l \text{ for each } l} \Delta^k_{a_1...a_k},$$

Let $I(i)$ and $I(i')$ be the groups of $(a+1)$-friends of, respectively, individuals i and i'. Fix a bijection $\pi : I(i') \to I(i)$ such that $\lambda_{jk} = \lambda_{\pi(j)\pi k}$ for all individuals $j, k \in I(i)$ and $\pi(i) = i'$. Such a bijection exists because λ is a hierarchy of groups. For each j, denote

$$j^0 = \begin{cases} j, & j \notin I(i') \\ \pi(j), & j \in I(i') \end{cases}, j^1 = \begin{cases} j, & j \notin I(i) \\ \pi^{-1}(j), & j \in I(i) \end{cases}.$$

For any $i_1 = i, i_2, ..., i_k$, denote $\iota = (i_2, ..., i_k)$, and

$$\begin{aligned} a_1^{0,\iota} &= \lambda_{i_1^0 i_2^0}, a_2^{0,\iota} = \lambda_{i_2^0 i_3^0}, ..., a_k^{0,\iota} = \lambda_{i_k^0 i_1^0}, \\ a^0(\iota) &= \left(a_1^{0,\iota}, ..., a_k^{0,\iota}\right), \\ a_1^{1,\iota} &= \lambda_{i_1^0 i_2^1}, a_2^{0,\iota} = \lambda_{i_2^1 i_3^0}, ..., a_k^{0,\iota} = \lambda_{i_k^{k-1 \bmod 2} i_1^{k \bmod 2}}, \\ a^1(\iota) &= \left(a_1^{1,\iota}, ..., a_k^{1,\iota}\right), \end{aligned}$$

Because λ is a hierarchy of groups, for each l,

$$\lambda_{i_l^0 i_{l+1}^0} = \lambda_{i_l^0 i_{l+1}^0} = a_l^{0,\iota} \geq \min\left(a_l^{0,\iota}, a\right) = a_l^{1,\iota} = \lambda_{i_l^1 i_{l+1}^0} = \lambda_{i_l^0 i_{l+1}^1},$$

and, if $i_l \notin I(i) \cup I(i')$ for some $l \leq k$, then $a_l^{0,\iota} = a_l^{1,\iota}$.

For any $\delta_1, ..., \delta_k \in \{0,1\}$, any $i_1 = i, i_2, ..., i_k$

$$\begin{aligned} &u^k\left(\lambda_{i_1^0 i_2^{\delta_1}}, \lambda_{i_2^{\delta_1} i_3^{\delta_1+\delta_2 \bmod 2}} ..., \lambda_{i_{k-1}^{\delta_1+...+\delta_{k-2} \bmod 2} i_k^{\delta_1+...+\delta_{k-1} \bmod 2}}, \lambda_{i_k^{\delta_1+...+\delta_{k-1} \bmod 2} i_1^{\delta_1+...+\delta_k \bmod 2}}\right) \\ &= u^k\left(a_1^{\delta_1,\iota}, ..., a_k^{\delta_k,\iota}\right), \end{aligned}$$

and

$$\begin{aligned} &\sum_{\delta_1,\delta_2,...,\delta_k=0,1} (-1)^{\delta_1+...+\delta_k} u^k\left(\lambda_{i_1^0 i_2^{\delta_1}}, \lambda_{i_2^{\delta_1} i_3^{\delta_1+\delta_2 \bmod 2}} ..., \lambda_{i_k^{\delta_1+...+\delta_{k-1} \bmod 2} i_1^{\delta_1+...+\delta_k \bmod 2}}\right) \\ &= \sum_{\delta_1=0,1} ... \sum_{\delta_k=0,1} (-1)^{\delta_1+....+\delta_k} u^k\left(a_1^{\delta_1}, ..., a_k^{\delta_k}\right) = \Delta^{a^0(\iota)}_{a^1(\iota)}. \end{aligned}$$

Because $a_l^{0,\iota} = a_l^{1,\iota}$ if $i_l \notin I(i) \cup I(i')$ for some $l \leq k$, $\Delta_{a^1(\iota)}^{a^0(\iota)} = 0$ whenever there is $l \leq k$ such that $i_l \notin I(i) \cup I(i')$. On the other hand, if for each $l \leq k$, $i_l \in I(i) \cup I(i')$, then $a_l^{0,\iota} > a = a_l^{1,\iota}$, and $\Delta_{a^1(\iota)}^{a^0(\iota)} = \Delta_{aa...a}^{a^0(\iota)}$.

Therefore,

$$
\begin{aligned}
&\sum_{i_2,...,i_k} \left(u^k\left(\lambda_{ii_2},...,\lambda_{i_{k-1}i_k},\lambda_{i_k i}\right) - u^k\left(\lambda_{ii_2},...,\lambda_{i_{k-1}i_k},\lambda_{i_k i'}\right)\right)\\
&= \sum_{i_1,i_2,...,i_k:i_1=i} \sum_{\delta_k=0,1} (-1)^{\delta_k} u^k\left(\lambda_{i_1 i_2},...,\lambda_{i_{k-1}i_k},\lambda_{i_k i_1^{\delta_k}}\right)\\
&= \sum_{i_1,i_2,...,i_k:i_1=i} \frac{1}{2^{\left|\left\{l:i_l^0\neq i_l^1\right\}\right|}} \sum_{\delta_1,\delta_2,...,\delta_k=0,1} (-1)^{\delta_k} u^k\left(\lambda_{i_1^0 i_2^{\delta_1}},...,\lambda_{i_{k-1}^{\delta_{k-2}} i_k^{\delta_{k-1}}},\lambda_{i_k^{\delta_{k-1}} i_1^{\delta_k}}\right)\\
&= \sum_{\iota} \frac{1}{2^{\left|\left\{l:i_l^0\neq i_l^1\right\}\right|}} \sum_{\delta_1,\delta_2,...,\delta_k=0,1} (-1)^{\delta_1+...+\delta_k} u^k\left(\lambda_{i_1^0 i_2^{\delta_1}},\lambda_{i_2^{\delta_1} i_3^{\delta_1+\delta_2 \bmod 2}}...,\lambda_{i_k^{\delta_1+...+\delta_{k-1} \bmod 2} i_1^{\delta_1+...+\delta_k \bmod 2}}\right)\\
&= \sum_{\iota} \frac{1}{2^{\left|\left\{l:i_l^0\neq i_l^1\right\}\right|}} \sum_{\delta_1,\delta_2,...,\delta_k=0,1} (-1)^{\delta_1+...+\delta_k} u^k\left(a_1^{\delta_1,\iota},...,a_k^{\delta_k,\iota}\right)\\
&= \sum_{\iota=(i_2,...,i_k):i_l\in I(i)\cup I(i')} \frac{1}{2^{\left|\left\{l:i_l^0\neq i_l^1\right\}\right|}} \Delta_{aa...a}^{a^0(\iota)} = \sum_{\iota=(i_2,...,i_k):i_l\in I(i)\cup I(i')} \frac{1}{2^{k-1}} \Delta_{aa...a}^{a^0(\iota)} = \sum_{\iota=(i_2,...,i_k):i_l\in I(i)} \Delta_{aa...a}^{a^0(\iota)},
\end{aligned}
$$

where the second equality follows from the fact that there are exactly $2^{\left|\left\{l:i_l^0\neq i_l^1\right\}\right|}$ choices of different $(\bar{\imath}_2,...,\bar{\imath}_k)$ so that $(\bar{\imath}_l)^0 = i_l^0$ and $(\bar{\imath}_l)^1 = i_l^1$ for each l, and the last two equalities follow from the fact that there are exactly 2^{k-1} choices of different $(\bar{\imath}_2,...,\bar{\imath}_k)$ so that $\bar{\imath}_l \in I(i) \cup I(i')$ and $(\bar{\imath}_l)^0 = i_l^0$ and $(\bar{\imath}_l)^1 = i_l^1$ for each l.

The above is further equal to

$$
\begin{aligned}
&= \sum_{a^0:a_l^0>a \text{ for each } l\leq k} \Delta_{aa...a}^{a^0} \left|\left\{\iota=(i_2,...,i_k): i_l \in I(i), a^{0,\iota}=a^0\right\}\right|\\
&= \sum_{a^0:a_l^0>a \text{ for each } l\leq k} \sum_{a_1,...,a_l:a<a_l\leq a_l^0 \text{ for each } l} \Delta_{a_1...a_k}^k \left|\left\{\iota=(i_2,...,i_k): i_l \in I(i), a^{0,\iota}=a^0\right\}\right|\\
&= \sum_{a_1,...,a_l:a<a_l \text{ for each } l} \Delta_{a_1...a_k}^k \sum_{a^0:a_l^0\geq a_l \text{ for each } l\leq k} \left|\left\{\iota=(i_2,...,i_k): i_l \in I(i), a^{0,\iota}=a^0\right\}\right|\\
&= \sum_{a_1,...,a_l:a<a_l \text{ for each } l} \Delta_{a_1...a_k}^k \frac{N_{a_1}...N_{a_k}}{N_{\min(a_1,...,a_k)}},
\end{aligned}
$$

where the last equality follows from the fact that λ is a hierarchy of groups (compare with (3.1)).

Therefore, for $\hat{k}=1$,

$$\sum_{i_1,i_2,...,i_k:i_l=i}\left(u^k\left(\lambda_{i_1j_2},...,\lambda_{i_{\hat{k}-1}i},...,\lambda_{i_ki_1}\right)-u^k\left(\lambda_{i_1j_2},...,\lambda_{i_{\hat{k}-1}i'},...,\lambda_{i_ki_1}\right)\right)$$
$$=\sum_{a_1,...,a_l:a<a_l\text{ for each }l}\Delta^k_{a_1...a_k}\frac{N_{a_1}...N_{a_k}}{N_{\min(a_1,...,a_k)}}.$$

Because the above argument applies for each $\hat{k}\leq k$, this concludes the proof of the Lemma.

A.4. **Proof of Proposition 3.** Let $n=\log_2 N$. Define operation $\oplus$ on binary sequences $b,c\in\{0,1\}^n$:

$$(b\oplus c)_s=b_s+c_s \bmod 2\text{ for each }i\leq n.$$

Operation $\oplus$ works as the XOR operator on each of the coordinates in the sequence. The following properties of operation $\oplus$ are useful: For any $b,c\in\{0,1\}^n$,

$$b\oplus c=c\oplus b, b\oplus b=\mathbf{0},\mathbf{0}\oplus b=b\oplus\mathbf{0}.$$

For each $a\leq A$, let $\beta(a)\in\{0,1\}^n$ be the binary representation of $a-1=\sum_{m\leq n}2^{m-1}\beta_m(a)$. For any $a,a'\leq A$, define

$$a\oplus a'=\beta(a)\oplus\beta(a').$$

We need a technical result: For any sequence $a_1,...,a_k$, for any $l\leq k$, let $(a_l,...,a_{l-1})$ be a shorthand for the ordered tuple $(a_l,a_{l+1},...,a_k,a_1,a_2,...,a_{l-1})$.

Lemma 6. *For any $a_1,...,a_k$ such that $a_k\neq a_1\oplus...\oplus a_{k-1}$, any $\theta\geq k$, any m*

$$\theta^{a_1+....+a_k}<\frac{1}{k}\sum_{l=1}^{k}\theta^{a_{l+1}+...+a_{l-2}+(a_l\oplus...\oplus a_{l-1})},$$

Proof. Because $\theta\geq k$, it is enough to show that for each $a_1,...,a_k$, there is l such that $a_{l+1}\oplus...\oplus a_{l-1}>a_l$. Indeed, find the smallest $n'\in\{1,...,n\}$ such that $\sum_l\beta_{n'}(a_l)$ is odd and find l so that $\beta_{n'}(a_l) = 1$. Then,

$$\beta_m(a_{l+1} \oplus ... \oplus a_{l-1}) = \beta_m(a_l) \text{ for each } m < n', \text{ and}$$
$$\beta_{n'}(a_{l+1} \oplus ... \oplus a_{l-1}) = 0 < 1 = \beta_{n'}(a_l).$$

□

In order to shorten the notation, define

$$\upsilon(a_1, ..., a_k) := \psi_k \theta_k^{a_1+...+a_{k-1}+a_1\oplus...\oplus a_{k-1}}.$$

Notice that for any $l \leq k-1$,

$$a_l = a_1 \oplus ... \oplus a_{l-1} \oplus a_{l+1} \oplus ... \oplus a_{k-1} \oplus (a_1 \oplus ... \oplus a_{k-1}),$$

which implies that, if $a_k = a_1 \oplus ... \oplus a_{k-1}$

$$\sum_{l=1}^{k} u^k(a_l, ..., a_{k-1+l}) = \sum_{l=1}^{k} u^k(a_l, ..., a_{k-2+l}, a_l \oplus ... \oplus a_{k-2+l}) = \sum_{l=1}^{k} \psi(a_l....a_{k+k-1}).$$

Let λ^* be a hierarchy of friends, and let λ be a profile that is not a hierarchy of friends. Then,

$$\begin{aligned} V^k(\lambda^*) &= \frac{1}{N^{k-1}} N \sum_{a_1,...,a_{k-1}} u^k(a_1, ..., a_{k-1}, a_1 \oplus ... \oplus a_{k-1}) \\ &= \frac{1}{N^{k-1}} N \sum_{a_1,...,a_{k-1}} \frac{1}{k} \sum_{l=1}^{k} u^k(a_l, ..., a_{k-2+l}, a_l \oplus ... \oplus a_{k-2+l}) \\ &= \frac{1}{N^{k-1}} N \sum_{a_1,...,a_{k-1}} \frac{1}{k} \sum_{l=1}^{k} \psi(a_l, ..., a_{k-2+l}) \\ &= \frac{1}{N^{k-1}} \sum_{i_1,...,i_k} \frac{1}{k} \sum_{l=1}^{k} \psi\left(\lambda_{i_l i_{l+1}}, ..., \lambda_{i_{k-2+l} i_{k-1+l}}\right), \end{aligned}$$

where the last equality follows from the fact that each player has attitude a toward exactly one player for each $a \leq A$. Because of Lemma 6, and because λ is not a hierarchy of groups, the above is larger than

$$> \frac{1}{N^{k-1}} \sum_{i_1,...,i_k} \frac{1}{k} \sum_{l=1}^{k} u^k \left(\lambda_{i_l i_{l+1}}, ..., \lambda_{i_{k-1+l} i_l}\right)$$

$$= \frac{1}{N^{k-1}} \sum_{i_1,...,i_k} u^k \left(\lambda_{i_1 i_2}, ..., \lambda_{i_k i_1}\right) = V^k(\lambda).$$

A.5. **Proof of Lemma 5.** We use the same notation as in the proof of Proposition 3.

Fix $k \geq 2$. Define

$$\mathcal{A} = \left\{(a_1, ..., a_k) \in A^k : a_k = a_1 \oplus ... \oplus a_{l-1} \oplus a_{l+1} \oplus ... \oplus a_{k-1}\right\}.$$

Then, for any hierarchy of friends λ and i_0, i_1 such that $\lambda_{i_0 i_1} = m$,

$$C^k(\lambda; i_0, i_1) = \sum_{\bar{a} \in \mathcal{A}} \left[\theta_k^{a_1+...+a_{k-1}+(a_k)} - \theta_k^{a_1+...+a_{k-1}+(a_k \oplus m)}\right].$$

Fix $m < A$. There exists $n_0 \leq n$ such that $\beta_{n_0}(m) = 0 < 1 = \beta_{n_0}(m+1)$ and $\beta_{n'}(m) = \beta_{n'}(m+1)$ for each $n' \neq n_0$. For each $z \in \{0, 1\}$, define

$$\mathcal{A}(z) = \left\{(a_1, ..., a_k) \in A^k : \beta_{n_0}(a_k) =\right\}.$$

Then, there exists a mapping $\gamma : \mathcal{A}(1) \to \mathcal{A}(0)$ such that, for each $\bar{a} \in \mathcal{A}(1)$, $|l < k : \beta_{n_0}(a_l) \neq \beta_{n_0}(\gamma(a_l))| = 1$. Because of the definition of operation $\oplus$, it must be that if $\beta_{n_0}(a_l) \neq \beta_{n_0}(\gamma(a_l))$, then $\beta_{n_0}(a_l) = 1 \neq \beta_{n_0}(\gamma(a_l)) = 0$. Notice that for each $\bar{a} \in \mathcal{A}(0)$, $|\gamma^{-1}(\bar{a})| \leq k - 1$,

For each $\bar{a} \in \mathcal{A}(0)$, for each $\bar{a}' \in \{\bar{a}\} \cup \gamma^{-1}(\bar{a})$,

$$a_1 + ... + a_{k-1} + (a_k \oplus m) > a_1' + ... + a_{k-1}' + (a_k' \oplus (m+1))$$

Because $\theta_k \geq k$, it must be that

$$\sum_{\bar{a}' \in \{\bar{a}\} \cup \gamma^{-1}(\bar{a})} \theta_k^{a_1+...+a_{k-1}+(a_k \oplus m)} > \sum_{\bar{a}' \in \{\bar{a}\} \cup \gamma^{-1}(\bar{a})} \theta_k^{a_1+...+a_{k-1}+(a_k \oplus (m+1))}.$$

Because $\{\bar{a}\} \cup \gamma^{-1}(\bar{a}) \cap \{\bar{a}'\} \cup \gamma^{-1}(\bar{a}')$ for $\bar{a} \neq \bar{a}'$, $\bar{a}, \bar{a}' \in \mathcal{A}(0)$, and $\mathcal{A}(0) \cup \mathcal{A}(1) = \mathcal{A}$, it must be that

$$\sum_{\bar{a} \in \mathcal{A}} \theta_k^{a_1+...+a_{k-1}+(a_k \oplus m)} > \sum_{\bar{a} \in \mathcal{A}} \theta_k^{a_1+...+a_{k-1}+(a_k \oplus (m+1))}.$$

This ends the proof of the Lemma.

University of Texas at Austin, Department of Economics, *Email*: mpeski@gmail.com