Friends and Enemies: A Model of Signed Network Formation

Timo Hiller*

October 16, 2012

Abstract

I propose a game of *signed* network formation, where agents make friends to coerce payoffs from enemies with fewer friends. The model accounts for the interplay between friendship and enmity. Nash equilibrium configurations are such that, either everyone is friends with everyone, or agents can be partitioned into sets of *different* size, where agents within the same set are friends and agents in different sets are enemies. These results mirror findings of a large body of work on signed networks in sociology, social psychology, international relations and applied physics.

*Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK (e-mail: timo.hiller@bristol.ac.uk). This paper is based on the second chapter of my Ph.D. dissertation at the European University Institute. A previous version was titled "Alliance Formation and Coercion in Networks." I am grateful to my advisors Fernando Vega-Redondo and Massimo Morelli for their invaluable support and guidance. I also thank Dilip Abreu, Sergio Currarini, Matthew O. Jackson, David K. Levine, In-Uck Park, Paolo Pin, Brian Roberson, Tomas Sjöström, Francesco Squintani, participants of the Networks Working Group at the European University Institute, seminar participants at the Institute for Advanced Studies (Princeton), FGV-Rio, CORE (Université catholique de Louvain), University of Vienna, Ludwig Maximilian University of Munich, Universitat de Barcelona, University of Bristol, PUC-Rio and participants of SAET 2011, CTN 2011 and the UECE Lisbon Meetings 2010 for helpful comments. All remaining errors are mine.

1 Introduction

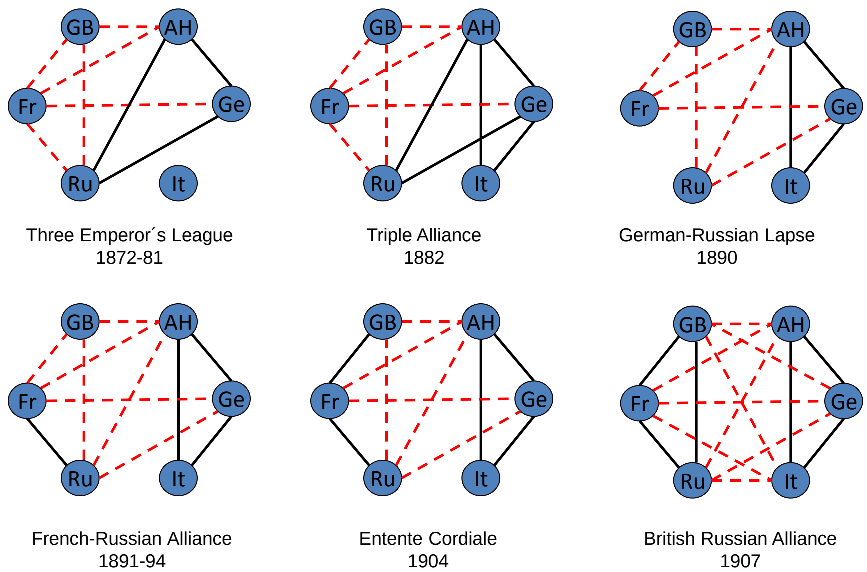
In much of the current economics literature on networks, links have a positive meaning and are commonly interpreted as friendship, collaboration or transmission of information. In many contexts, however, links may also be associated with negative sentiments, such as antagonism, coercion or even outright conflict. This paper sheds light on the interplay between these two forces by way of a game-theoretic model of *signed* network formation.

The study of signed networks, consisting of positive and negative links, has a long tradition in sociology and social psychology, dating back to Heider's seminal contribution on "cognitive dissonance" in 1946. The essential idea is that positively connected individuals tend to match their attitudes relative to third agents. That is, triads are expected to either consist of three positive, or one positive and two negative links. Cartwright and Harary (1956) proved that these local properties, which they coined as structural balance, yield sharp predictions globally.[1] In particular, the only network configurations that are structurally balanced are such that either all agents are friends, or there exist two distinct sets, also called cliques, where agents in the same set are friends and agents in different sets sustain antagonistic relationships.[2] Davis (1967) showed that, when allowing for triads of three negative links, "weakly balanced" graphs may consist of multiple cliques.

[1]For a very good introduction to the literature on structural balance see Easley and Kleinberg (2010).

[2]I provide a sketch of the proof. A graph with only positive links is balanced. For a graph with positive and negative links, pick an arbitrary agent i and divide the remaining set of agents into $i's$ friends and $i's$ enemies. All of $i's$ friends must be friends, as otherwise one obtains an unbalanced triad with two positive and one negative link. All of $i's$ enemies must be friends, as otherwise one obtains an unbalanced triad with three negative links. Links between $i's$ friends and $i's$ enemies must be negative, as otherwise one obtains an unbalanced triad with two negative and one positive link.

Structural Balance and The Path to WWI



The predictions of structural balance became an effective tool for the analysis of behavior of nations, especially in times of crisis and mounting threats of war. One of the earliest applications is Harary (1961), who examines the rapid shifts of relationships among nations in the Middle Eastern crisis of 1956 and observes a strong tendency towards balance. Moore (1979) also employs structural balance when explaining the "United States's somewhat surprising support of Pakistan..." in the conflict over Bangladesh's separation from Pakistan in 1972. Another interesting example is provided by Antal, Krapivsky and Redner (2006). They link the formation of alliances in the 19th century - ultimately leading up to WWI - to structural balance. The accompanying graph is depicted above, where alliances are denoted by straight lines and antagonistic relationships by dashed lines. Note that the system of relationships gradually moves towards a complete, structurally balanced network.[3]

[3] A point to be made here is that, although balance appears to be a natural outcome, its implications need not be positive. It may lead to fierce opposition between two sides, which is difficult to resolve.

Interaction patterns of individuals have also been examined for structural balance properties. Szell, Lambiotte and Thurner (2010) analyze a vast data-set from a multiplayer online game (Pardus), encompassing more than 300.000 players. The game allows for six types of interactions, of which some have a positive (friendship, communication, trade) and others have a negative association (hostility, aggression, punishment). The authors provide strong support for structural balance, favoring its weak specification.[4] In a recent paper Facchetti, Iacono and Altafini (2011) analyze signed relationships of three websites (Epinions, Slashdot and WikiElections). They find evidence for structural balance in all of them. Research in sociology has examined the evolution of signed network relations. Doreian and Mrvar (1996) and Doreian and Krackhardt (2001) are two such empirical studies. In both cases a movement towards balance is evident.

This paper proposes a game of strategic network formation, which allows agents to form positive and negative links. The model clarifies the interplay between friendship and alliances on the one hand and antagonism and enmity on the other. My main finding is that every Nash equilibrium of the game obeys (weak) structural balance. That is, Nash equilibrium configurations are such that either all links are positive, or agents can be divided into two or more distinct sets of *different* size, where agents within the same set are friends and agents in different sets are enemies. The asymmetry of equilibria is a salient characteristic of the model.

The setup is simple. Players can either extend a friendly (positive) or an antagonistic (negative) link to each of the remaining agents, at zero cost. A reciprocated positive link constitutes a *friendship* or *alliance*. If at least one link between two agents is negative, then we think of it as an antagonistic relationship. Two types of antagonistic relationships are discerned. First, *coercion*, where one agent extends a positive and the other a negative link. Second, *conflict*, in which case both agents extend a negative link.[5] Under a coercive relationship the agent with more friends or allies extracts payoffs from the agent with fewer friends. The same holds for

[4]For theoretical work on structural balance in the physics and mathematics literature, see Antal, Krapivsky and Redner (2005), Marvel, Strogatz and Kleinberg (2009) and Marvel, Kleinberg, Kleinberg and Strogatz (2011).

[5]For an analogous definition of friendship, coercion and conflict in the social psychology literature, see Willer (1999).

conflict, but now both agents additionally incur a cost of conflict. The assumption that conflict is costly arises naturally, but is also convenient for modeling purposes. When using Nash equilibrium, it ensures that agents are not trapped in a situation where they extend negative links to each other, although both prefer a reciprocated positive link. For related reasons we do not allow for neutral or *zero* links. The latter assumption guarantees that agents do not extend neutral links to each other, although both prefer a reciprocated positive link.[6]

A crucial feature of the model is that the payoffs an agent with more allies (the stronger agent) can extract from an enemy with fewer friends (the weaker agent) is strictly increasing in the own number of allies and strictly decreasing in the number of the respective enemy's allies. This is what drives stronger agents to match their strategies relative to weaker agents. By coordinating on who to coerce, the coerced agents have fewer friends. Note that the rationale of coercion is quite different from cognitive dissonance, as originally conceived by Heider. A model of cognitive dissonance would presumably impose a cost on agents whose friends sustain antagonistic relationships among each other. This paper, in contrast, stresses coercion as an incentive to make friends and enemies and determines the coercive power of an agent endogenously.

In the following the arguments underlying the main result are briefly outlined. Notice first that open conflict is not part of any Nash equilibrium, as each agent involved in conflict can profitably deviate by extending a positive link instead, thereby not incurring the cost of conflict.[7] Next, note that under a coercive link, it is always the agent with more allies who extends the negative link. Otherwise, the weaker agent can profitably deviate by extending a positive link, thus creating an alliance and avoiding negative payoffs under a coercive link (while increasing payoffs on any remaining negative links). Therefore, weaker agents extend positive links to

[6] This is a well known issue, which already arises in Myerson's (1977) link-announcement game and motivates the notion of pairwise stability (Jackson and Wolinsky, 1996). Appendix B presents an alternative model with neutral links, no cost of conflict and bilateral equilibrium (Goyal and Vega-Redondo, 2007) as equilibrium concept. Bilateral equilibrium admits coordinated deviations of pairs of agents and refines pairwise stability. The equilibrium characterization is almost identical to the one presented in the main part of the paper.

[7] In the economics of conflict literature this is known as "settlement in the shadow of conflict". For a more detailed discussion, see the related literature section at the end of the introduction.

stronger agents in any Nash equilibrium. Stronger agents then face a trade-off between coercing a weaker agent and creating an alliance: By creating an alliance, the stronger agent forgoes the coercion payoff from that particular relationship, but increases payoffs on all remaining coercive links. One can easily show that agents with the same number of friends must be friends in any Nash equilibrium. The intuition is that agents with the same number of friends can not coerce payoffs from each other, while entering an alliance increases payoff from coercive links. This is the source of asymmetry of equilibrium configurations. Finally, cliques of positively connected agents arise due to aforementioned incentives of stronger agents to match their strategies relative to weaker agents.

The equilibrium is characterized for a general payoff function, which maps the respective numbers of friends and the number of common friends into an extraction payoff under an antagonistic relationship. Two frequently used functions in the economics of conflict literature are special cases (after a normalization): The contest success function in ratio and in difference form. I obtain simple comparative static results for one of them, the difference form, where the extraction value is a function of the difference in the respective numbers of friends. If the parameter of the contest success function is sufficiently low, and thus coercion relatively less profitable, then Nash equilibria are such that there can be at most two cliques. If the parameter on the contest success function is sufficiently high, and coercion is relatively more profitable, then multiple cliques may arise and relative group size is at least geometrically increasing. The latter result holds if the general payoff function is bounded and therefore applies to both, the contest success function in ratio and in difference form.

In the final part of the paper we allow for heterogeneous agents, i.e., agents may display ex-ante differences in intrinsic coercive strength. Ex-post coercive strength is given by an agent's intrinsic strength, plus the sum of his allies' intrinsic strength. The equilibrium characterization is similar to the homogenous case. Cliques also arise. However, they are now not necessarily of different size, but of different ex-post coercive strength. The main difference relative to the homogenous case is that everyone being friends with everyone is not an equilibrium. This is easy to see, as a stronger agent can then profitably deviate by extending a negative link to any weaker agent.

The paper relates to the economics of conflict literature, which recognizes that property rights may not be perfectly or costlessly enforced. Conflict is modeled in terms of a contest success function (Tullock, 1967, 1980 and Hirshleifer, 1989), where an agent's probability of winning is determined by the resources available for arming. Open conflict, however, does not need to take place and may be settled "in the shadow of conflict". That is, in the absence of asymmetric information and dynamic considerations, the weaker agent hands over resources to the stronger agent in order to avoid the cost of conflict. Part of this research focuses on coalition formation. See, for example, Wärneryd (1998) and Esteban and Sákovics (2003). Based on Chwe's (1994) notion of farsighted stability, group structures are shown to be symmetric. These findings stand in contrast to the results obtained here.[8,9]

The model also contributes to the literature on networks in economics. See, for example, Myerson (1977), Aumann and Myerson (1988), Bala and Goyal (2000) and Jackson and Wolinsky (1996). Two recent papers, which feature contest success functions in a network setting are Goyal and Vigier (2010) and Franke and Öztürk (2009), both with a different focus from mine. Goyal and Vigier (2010) study a design problem and ask how to optimally structure networks, so that they are robust to attacks in the face of an adversary. Franke and Öztürk (2009), in turn, model a setting where agents are embedded in a network of bilateral conflicts. The authors are concerned with conflict intensity on a fixed network and do not consider the possibility of alliances. To the best of my knowledge, this paper presents the first attempt to incorporate friendship or alliance on the one hand and enmity on the other in a network formation context.

The remaining part of the paper is organized as follows: Section 2 introduces the model, Section 3 provides the equilibrium characterization, comparative statics

[8] Jordan (1996) considers coalitional games, where more powerful coalitions can pillage all wealth of weaker coalitions at no cost. The farsighted core allocations, where no coalitions are formed and no acts of pillage occur, depend on a power function. If power is determined by wealth only, then farsighted core allocations may be asymmetric. Interestingly, when extraction values are bounded and coercion is highly effective, then there exists a stable allocation of wealth that is, for corresponding parameter values, reminiscent of the distribution of clique size in my model.

[9] See Rietzke and Roberson (2012) for a setting in which two potential allies face a fixed, common enemy.

and a setting with heterogeneous agents, Section 4 concludes. The proof for the equilibrium characterization is relegated to Appendix A. Appendix B introduces a variant of the model, which allows for neutral links, no cost of conflict and uses bilateral equilibrium as equilibrium concept.

2 Model

Let $N = \{1,2,...,n\}$ be the set of ex-ante identical agents, with $n \geq 3$. A strategy for $i \in N$ is defined as a row vector $\mathbf{g}_i = (g_{i,1}, g_{i,2}, ..., g_{i,i-1}, g_{i,i+1}, ..., g_{i,n})$, where $g_{i,j} \in \{-1,1\}$ for each $j \in N \backslash \{i\}$. Agent i is said to extend a positive link to j if $g_{i,j} = 1$ and a negative link if $g_{i,j} = -1$. The set of strategies of i is defined by G_i and the strategy space by $G = G_1 \times ... \times G_n$. The resulting network of relationships is written as $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n)$. Define the undirected network $\bar{\mathbf{g}}$ in the following way. The link between agents i and j is positive in the undirected network $\bar{\mathbf{g}}$, if both directed links are positive, so that $\bar{g}_{i,j} = 1$ if $g_{i,j} = g_{j,i} = 1$. The link in the undirected network is negative, if one of the two undirected links is negative and the other one is positive, so that $\bar{g}_{i,j} = -1$ if $g_{i,j} * g_{j,i} = -1$. We will call a link in the undirected network $\bar{\mathbf{g}}$ a double negative link, if both agents involved extend a negative link, so that $g_{i,j} = g_{j,i} = -1$ and write $\bar{g}_{i,j} = -2$. Given a network $\mathbf{g}$, $\mathbf{g} + g^+_{i,j}$ and $\mathbf{g} + g^-_{ij}$ have the following interpretation. If $g_{i,j} = -1$ in $\mathbf{g}$, then $\mathbf{g} + g^+_{i,j}$ changes the directed link $g_{i,j} = -1$ into $g_{i,j} = 1$, while if $g_{i,j} = 1$ in $\mathbf{g}$, then $\mathbf{g} + g^+_{i,j} = \mathbf{g}$. Similarly, if $g_{i,j} = 1$ in $\mathbf{g}$, $\mathbf{g} + g^-_{i,j}$ changes the directed link $g_{i,j} = 1$ into $g_{i,j} = -1$, while if $g_{i,j} = -1$ in $\mathbf{g}$, then $\mathbf{g} + g^-_{i,j} = \mathbf{g}$.

Define the following sets: $N^+_i(\mathbf{g}) = \{j \in N \mid \bar{g}_{i,j} = 1\}$ is the set of agents to which agent i reciprocates a positive link and therefore $\bar{g}_{i,j} = 1$ in the undirected network $\bar{\mathbf{g}}$. Denote the set of common friends of agents i and j with $N^+_{i,j}(\mathbf{g}) = N^+_i(\mathbf{g}) \cap N^+_j(\mathbf{g})$. $N^{-1}_i(\mathbf{g}) = \{j \in N \mid \bar{g}_{i,j} = -1\}$ is the set of agents to which i does not reciprocate the sign of a directed link and therefore $\bar{g}_{i,j} = -1$, while $N^{-2}_i(\mathbf{g}) = \{j \in N \mid \bar{g}_{i,j} = -2\}$ is the set of agents to which agent i reciprocates a negative link and $\bar{g}_{i,j} = -2$. Define $N^-_i(\mathbf{g}) = N^{-1}_i(\mathbf{g}) \cup N^{-2}_i(\mathbf{g})$ and denote the following cardinalities with $\eta_i(\mathbf{g}) = \mid N^+_i(\mathbf{g}) \mid$, $\eta_{i,j}(\mathbf{g}) = \mid N^+_{i,j}(\mathbf{g}) \mid$ and $\gamma_i(\mathbf{g}) = \mid N^{-2}_i(\mathbf{g}) \mid$. For ease of notation we sometimes write η_i for $\eta_i(\mathbf{g})$ and $\eta_{i,j}$ for $\eta_{i,j}(\mathbf{g})$.

Links are interpreted in the following way: A reciprocated positive link, i.e., a positive link in the undirected network $\bar{\mathbf{g}}$, establishes a friendship or alliance between i and j. $\bar{g}_{i,j} = -1$ stems from one positive and one negative link and denotes a coercive relationship, whereas a reciprocated negative link, $\bar{g}_{i,j} = -2$, indicates conflict. The coercive power of agent i relative to agent j is determined by the respective number of friends, η_i and η_j, and the number of common friends, $\eta_{i,j}$. Under a coercive link, the agent with more friends coerces payoffs from agents with fewer friends. The same holds for a reciprocated negative link, but now agents additionally incur a cost of conflict, κ.

The payoffs to player i under strategy profile $\mathbf{g}$ are given by

$$\Pi_i(\mathbf{g}) = \sum_{j \in N_i^-(\mathbf{g})} f(\eta_i(\mathbf{g}), \eta_j(\mathbf{g}), \eta_{i,j}(\mathbf{g})) - \gamma_i(\mathbf{g})\kappa,$$

with $\kappa > 0$. Note first that we only sum over all agents to which antagonistic relationships are sustained. That is, direct payoffs from a reciprocated positive link are *zero* and the sole purpose of an alliance is to increase payoffs on antagonistic links. This brings out the tension between friendship/alliance and antagonism in the starkest manner. Under a coercive link, an agent's gain is assumed to be another agent's loss and therefore $f(\eta_i, \eta_j, \eta_{i,j}) + f(\eta_j, \eta_i, \eta_{i,j}) = 0$ for all η_i, η_j, $\eta_{i,j}$. If two agents have the same number of friends, $\eta_i = \eta_j$, then they are not able to coerce payoffs form each other and $f(\eta_i, \eta_j, \eta_{i,j}) = 0$ for all $\eta_{i,j}$. The function f strictly increasing in η_i and strictly decreasing in η_j. Therefore, if $\eta_i > \eta_j$, then $f(\eta_i, \eta_j, \eta_{i,j}) > 0$ and if $\eta_i < \eta_j$, then $f(\eta_i, \eta_j, \eta_{i,j}) < 0$. We also allow for common friends to enter the payoff function. Here we need that $f(\eta_i + 1, \eta_j, \eta_{i,j} + 1) > f(\eta_i, \eta_j, \eta_{i,j})$. That is, the combined effect of an increase in own number of friends and an increase in common friends on payoffs from a coercive link is positive.[10] This assumption ensures, first, that positively connected agents coordinate their actions relative to third agents and, second, that payoffs from remaining negative

[10] Write the forward difference for η_i as $\triangle_{\eta_i} f(\eta_i, \eta_j, \eta_{i,j}) = f(\eta_i + 1, \eta_j, \eta_{i,j}) - f(\eta_i, \eta_j, \eta_{i,j})$ and for $\eta_{i,j}$ as $\triangle_{\eta_{i,j}} f(\eta_i, \eta_j, \eta_{i,j}) = f(\eta_i, \eta_j, \eta_{i,j} + 1) - f(\eta_i, \eta_j, \eta_{i,j})$. If f is such that the effect of an increase in η_i and an increase in $\eta_{i,j}$ are additively separable, then the latter condition can be written as $\triangle_{\eta_i} f(\eta_i, \eta_j, \eta_{i,j}) > -\triangle_{\eta_{i,j}} f(\eta_i, \eta_j, \eta_{i,j})$.

links are strictly increasing when creating a new positive link.

The two most commonly used contest success functions (ratio and difference from) fit the model after a normalization. Both functions assume that an agent's share of a prize is a function of the respective resources available for arming.[11] In the ratio form, relative shares are a function of the ratio of resources, while in the difference form contest success depends on the difference in resources. Denote with c_i and c_j the contest inputs for agent i and j, respectively. In the ratio form $i's$ share of the prize is given by

$$p_{i,j} = \frac{c_i{}^{\phi}}{c_i{}^{\phi} + c_j{}^{\phi}},$$

with $\phi > 0$. This function is a special case of our payoff function f, when defining $c_i = \eta_i^+ + 1$ to avoid dividing by *zero* and subtracting $\frac{1}{2}$, so that the coercion payoff is *zero* for agents with equal strength. The payoff that agent i obtains from agent j under an antagonistic link can then be written as

$$p_{i,j} = \frac{(\eta_i^+ + 1)^{\phi}}{(\eta_i^+ + 1)^{\phi} + (\eta_j^+ + 1)^{\phi}} - \frac{1}{2}.$$

Payoffs for agent j are given by $p_{j,i} = -p_{i,j}$. Note that in this specification the number of common friends does not enter the payoff function. Similarly, for the contest success function in difference form we can write

$$p_{i,j} = \frac{1}{1 + e^{\phi(\eta_j^+ - \eta_i^+)}} - \frac{1}{2}.$$

In both cases the function is parametrized by ϕ, where a higher value of ϕ is favorable for the agent with more friends. Both functions are bounded by $-\frac{1}{2}$ and $\frac{1}{2}$.

[11] In the case of a discrete either-or competition, shares may be interpreted as winning probabilities.

The equilibrium concept used is Nash Equilibrium. A strategy profile $\mathbf{g}^*$ is a Nash Equilibrium (NE) iff

$$\Pi_i(\mathbf{g}_i^*, \mathbf{g}_{-i}^*) \geq \Pi_i(\mathbf{g}_i, \mathbf{g}_{-i}^*), \forall \mathbf{g}_i \in G_i, \forall i \in N.$$

The network resulting from a proposed deviation is denoted with $\mathbf{g}'$.

3 Analysis

This section shows that Nash equilibria are such that either all agents are friends, or agents can be partitioned into sets of different size, where agents within the same set are friends and agents in different sets are enemies. Agents in larger sets coerce payoffs from agents in smaller sets. Lemma 1 rewrites the condition $f(\eta_i+1, \eta_j, \eta_{i,j}+1) > f(\eta_i, \eta_j, \eta_{i,j})$. Lemma 2 shows that, when creating a new undirected positive link, payoffs from remaining negative links are strictly increasing. Lemma 3 and 4 are devoted to equilibrium properties of bilateral relations. Lemma 3 states that outright conflict is not part of any Nash equilibrium, while Lemma 4 shows that in a coercive relationship, it must be the agent with more friends extending the negative link. Proposition 1 presents the equilibrium characterization, while Proposition 2 and Proposition 3 provide existence results.

Lemma 1: $f(\eta_i+1, \eta_j, \eta_{i,j}+1) > f(\eta_i, \eta_j, \eta_{i,j}) \Rightarrow f(\eta_i, \eta_j-1, \eta_{i,j}-1) > f(\eta_i, \eta_j, \eta_{i,j})$.

Proof. From $f(\eta_i, \eta_j, \eta_{i,j}) + f(\eta_j, \eta_i, \eta_{i,j}) = 0$ and $f(\eta_i+1, \eta_j, \eta_{i,j}+1) > f(\eta_i, \eta_j, \eta_{i,j})$ we know that $f(\eta_j, \eta_i+1, \eta_{i,j}+1) < f(\eta_j, \eta_i, \eta_{i,j})$ must hold. The latter condition can be rewritten as $f(\eta_j, \eta_i-1, \eta_{i,j}-1) > f(\eta_j, \eta_i, \eta_{i,j})$. As this is true for arbitrary i and j, we can relabel and write $f(\eta_i, \eta_j-1, \eta_{i,j}-1) > f(\eta_i, \eta_j, \eta_{i,j})$. *Q.E.D.*

Lemma 2 shows that when an undirected positive link is created, then payoffs on remaining negative links are strictly increasing. This is due to the assumption that $f(\eta_i+1, \eta_j, \eta_{i,j}+1) > f(\eta_i, \eta_j, \eta_{i,j})$.

Lemma 2: *If $\bar{g}_{i,k} \neq 1$ and $\hat{\mathbf{g}} = \mathbf{g} + g^{+}_{i,k} + g^{+}_{k,i}$, then $f(\eta_i(\hat{\mathbf{g}}), \eta_j(\hat{\mathbf{g}}), \eta_{i,j}(\hat{\mathbf{g}})) > f(\eta_i(\mathbf{g}), \eta_j(\mathbf{g}), \eta_{i,j}(\mathbf{g}))$ holds for all $j \in N^{-}_{i}(\hat{\mathbf{g}})$.*

Proof. Note that the only difference between networks $\mathbf{g}$ and $\hat{\mathbf{g}}$ is that $\bar{g}_{i,k} \neq 1$, while $\hat{\bar{g}}_{i,k} = 1$. We discern two cases. First, $\bar{g}_{i,j} = -1$, $\bar{g}_{i,k} = -1$ and $\bar{g}_{j,k} = -1$. From $\bar{g}_{j,k} = -1$ we know that $\eta_{i,j}(\mathbf{g}) = \eta_{i,j}(\hat{\mathbf{g}})$. That is, because j and k are enemies, the number of common friends between i and j remains the same when $\hat{\bar{g}}_{i,k} = 1$ is created. Agent i then accrues higher payoffs from j in $\hat{\mathbf{g}}$ than in $\mathbf{g}$, which follows from $f(\eta_i + 1, \eta_j, \eta_{i,j}) > f(\eta_i, \eta_j, \eta_{i,j})$. Second, $\bar{g}_{i,j} = -1$, $\bar{g}_{i,k} = -1$ and $\bar{g}_{j,k} = 1$. From $\bar{g}_{j,k} = 1$ we know that $\eta_{i,j}(\hat{\mathbf{g}}) = \eta_{i,j}(\mathbf{g}) + 1$. That is, because j and k are friends, the number of common friends between i and j increases by *one* when $\hat{\bar{g}}_{i,k} = 1$ is created. Agent i then accrues higher payoffs from j in $\hat{\mathbf{g}}$ than in $\mathbf{g}$, which follows from $f(\eta_i + 1, \eta_j, \eta_{i,j} + 1) > f(\eta_i, \eta_j, \eta_{i,j})$. *Q.E.D.*

Lemma 3: *In any NE $\mathbf{g}^*$, $\bar{g}^*_{i,j} \neq -2$ for all $i, j \in N$.*

Proof. Assume there exists a pair of agents, i and j, with a reciprocated negative link. This can not be part of any Nash equilibrium $\mathbf{g}^*$, as a deviation in the form of extending a positive link to the respective other agent is profitable for either one of the two agents. Assume agent i deviates by extending a positive link to j. The sets of $i's$ friends and enemies are not altered by the deviation and payoffs accruing from any third party k remain the same. However, payoffs from the relationship with j strictly increase by κ. *Q.E.D.*

In Lemma 4 we show that in any Nash equilibrium, for all negative links in place in the undirected network $\bar{\mathbf{g}}^*$, it must be the agent with more friends who extends the directed negative link. This is easy to see, as otherwise the agent with fewer friends can profitably deviate by reciprocating the friendly link, thereby increasing his payoff from this specific link. Moreover, payoffs on any remaining negative links increase by Lemma 2.

Lemma 4: *In any NE $\mathbf{g}^*$, if $\bar{g}^*_{i,j} = -1$ with $\eta_i(\mathbf{g}^*) < \eta_j(\mathbf{g}^*)$, then $g^*_{i,j} = 1$.*

Proof. Assume to the contrary that the link between agents i and j is negative in the undirected network $\bar{\mathbf{g}}^*$, and it is the agent with fewer friends extending the directed negative link. More formally, assume that $\bar{g}^*_{i,j} = -1$ with $g^*_{i,j} = -1$, $g^*_{j,i} = 1$ and $\eta_i(\mathbf{g}^*) < \eta_j(\mathbf{g}^*)$. Then i can profitably deviate by extending a positive link to j with deviation strategy $g^*_i + g^+_{i,j}$, yielding $\bar{g}_{i,j} = 1$. This strictly increases payoffs for i from his link with j, as under a negative undirected link $f(\eta_i(\mathbf{g}^*), \eta_j(\mathbf{g}^*), \eta_{i,j}(\mathbf{g}^*)) < 0$ for $\eta_i(\mathbf{g}^*) < \eta_j(\mathbf{g}^*)$. Furthermore, from Lemma 2 we know that, by reciprocating a positive link from j, agent i increases payoffs on any remaining negative links. *Q.E.D.*

Before presenting the equilibrium characterization, we define the set of agents with k friends, $P_k(\mathbf{g}) = \left\{ j \in N \mid \eta_j(\mathbf{g}) = k \right\}$. Denote the set with the highest number of friends with $P^m(\mathbf{g})$, the one with the second highest subscript $P^{m-1}(\mathbf{g})$ and proceed in this way until the set of agents with the fewest number of friends, $P^1(\mathbf{g})$. Proposition 1 shows that in any Nash equilibrium, agents with the same number of friends are friends and agents with different numbers of friends are enemies. That is, either all agents are positively connected, or the sets $P_k(\mathbf{g}^*)$ constitute maximal cliques of *different* size, with agents in cliques of larger size coercing payoffs from agents in cliques of smaller size.[12] Proposition 2 shows, as part of the existence results, that there always exists a Nash equilibrium where everyone is friends with everyone else, so that $\bar{g}^*_{i,j} = 1$ and $\eta_i(\mathbf{g}^*) = \eta_j(\mathbf{g}^*)$ $\forall i, j \in N$. In Proposition 1 we therefore focus on the case where a pair of agents i and j exists, such that $\eta_i(\mathbf{g}^*) \neq \eta_j(\mathbf{g}^*)$.

Proposition 1: *In any NE* $\mathbf{g}^*$*, if* $\eta_i(\mathbf{g}^*) = \eta_j(\mathbf{g}^*)$*, then* $\bar{g}^*_{i,j} = 1$ *and if* $\eta_i(\mathbf{g}^*) \neq \eta_j(\mathbf{g}^*)$*, then* $\bar{g}^*_{i,j} = -1$.

Proof. *See the Appendix.*

We provide a brief, informal outline of the proof. The proof is by induction. The base case shows in four steps that agents in $P^m(\mathbf{g}^*)$ are positively connected to agents in $P^m(\mathbf{g}^*)$ and negatively to agents not in $P^m(\mathbf{g}^*)$.

Step 1 proves that agents in $P^m(\mathbf{g}^*)$ must be positively connected. Note first, that in any Nash equilibrium, an agent $i \in P^m(\mathbf{g}^*)$ must be negatively connected to some

[12] Formally, a *clique* is a set of agents $C(\mathbf{g}) \subseteq N$, such that $\bar{\mathbf{g}}_{i,j} = 1$ $\forall i, j \in C(\mathbf{g})$. A clique is *maximal* and denoted with $C^m(\mathbf{g})$, if for any $l \notin C^m(\mathbf{g})$, $C^m(\mathbf{g}) \cup \{l\}$ is not a clique.

agent $k \notin P^m(\mathbf{g}^*)$. Otherwise, i can profitably deviate by extending a negative link to $k \notin P^m(\mathbf{g}^*)$. To see that agents in $P^m(\mathbf{g}^*)$ must be positively connected, note that agents within $P^m(\mathbf{g}^*)$ have the same number of friends and cannot extract payoffs from each other, while creating a positive link increases payoffs on all remaining negative links.

Step 2 shows that agents in $P^m(\mathbf{g}^*)$ play the same strategies relative to agents not in $P^m(\mathbf{g}^*)$. Take agent $i \in P^m(\mathbf{g}^*)$ with (weakly) highest payoffs in $P^m(\mathbf{g}^*)$ and an agent $j \in P^m(\mathbf{g}^*)$, such that i and j's sets of friends and enemies differ. Using Lemma 1 we show that by imitating $i's$ strategy, and thereby obtaining the same sets of friends and enemies as i, agent j accrues payoffs that are strictly higher than $i's$ payoffs prior to the deviation. Proposed deviation is therefore profitable.

Step 3 demonstrates that undirected links between all agents in $P^m(\mathbf{g}^*)$ and all agents in $P^{m-1}(\mathbf{g}^*)$ are negative. Assume to the contrary, and in accordance with Step 2, that links between all agents in $P^m(\mathbf{g}^*)$ and some agent $k \in P^{m-1}(\mathbf{g}^*)$ are positive. If agents in $P^m(\mathbf{g}^*)$ and k play the same strategies relative to third agents, then they must have the same number of friends. This yields an immediate contradiction. If i and k play different strategies relative to third agents, then a profitable deviation exists analogous to the one proposed in Step 2.

Finally, Step 4 proves that undirected links between agents in $P^m(\mathbf{g}^*)$ and any agent not in $P^m(\mathbf{g}^*)$ are negative. We start by showing that links between agents in $P^m(\mathbf{g}^*)$ and an agent in $P^{m-2}(\mathbf{g}^*)$ are negative. The reasoning is similar to the one of Step 3. Note that in Step 4, however, an agent in $P^{m-2}(\mathbf{g}^*)$ can not simply imitate the strategy of an agent in $P^m(\mathbf{g}^*)$ to obtain the same sets of friends and enemies, as deviation strategies of agents in $P^{m-2}(\mathbf{g}^*)$ must take into account the possibility that agents in $P^{m-1}(\mathbf{g}^*)$ extend negative links to agents in $P^{m-2}(\mathbf{g}^*)$. With appropriately adapted deviation strategies we then use the argument of Step 3 iteratively to show that agents in $P^m(\mathbf{g}^*)$ extend positive links to all agents in $P^m(\mathbf{g}^*)$ and negative links to all agents not in $P^m(\mathbf{g}^*)$.

For the inductive step we define the super set $\tilde{P}^r(\mathbf{g}^*) = P^m(\mathbf{g}^*) \cup P^{m-1}(\mathbf{g}^*) \cup \ldots \cup P^{m-r-1}(\mathbf{g}^*) \cup P^{m-r}(\mathbf{g}^*)$. Note that $\tilde{P}^0(\mathbf{g}^*) = P^m(\mathbf{g}^*)$. In Step 4 we showed that all agents in $P^m(\mathbf{g}^*)$ are positively connected with all agents in $P^m(\mathbf{g}^*)$ and negatively with all agents not in $P^m(\mathbf{g}^*)$. Assume that the statement holds for all sets in $\tilde{P}^r(\mathbf{g}^*)$. We can then repeat steps 1 through 4 from the base case, relabeling $P^m(\mathbf{g}^*)$ with $P^{m-(r+1)}(\mathbf{g}^*)$, $P^{m-1}(\mathbf{g}^*)$ with $P^{m-(r+2)}(\mathbf{g}^*)$ and so forth, to show that the statement holds for agents in $P^{m-(r+1)}(\mathbf{g}^*)$.

The following two propositions provide existence results. Proposition 2 proves that there always exists an equilibrium where everyone extends positive links to everyone else. To see this, note that no agent can unilaterally deviate and obtain more friends than some other agent. Proposition 3 shows that there always exists an equilibrium where $n-1$ agents are friends of each other and one agent is coerced by everyone else. The argument is similar to the one in Proposition 2, but now we need to also account for deviations where one of the $n-1$ agents enters an alliance with the agent that is coerced by everyone else, while extending negative links to some of the remaining agents.

Proposition 2: *There always exists a NE* $\mathbf{g}^*$, *such that all agents are friends.*

Proof. A deviation for agent i consists of extending negative links to some subset of $N\backslash\{i\}$. If the deviation strategy of i consists of extending a negative link to *one* other agent j, then payoffs remain *zero*, as $\eta_i(\mathbf{g}') = \eta_j(\mathbf{g}') = N-2$. If the deviation strategy of i consists of extending two or more negative links, then $i's$ payoffs will be strictly lower in $\mathbf{g}'$, as $\eta_i(\mathbf{g}') < \eta_j(\mathbf{g}')$ for all $j \in N_i^-(\mathbf{g}')$. *Q.E.D.*

Proposition 3: *There always exists a NE* $\mathbf{g}^*$, *such that* $n-1$ *agents are friends of each other and one agent is an enemy of everyone else.*

Proof. Denote with k the agent that is an enemy of everyone else. First we check for deviations by k. From Lemma 4 we know that under a negative link $\bar{g}^*_{i,k} = -1$, it must be the agent with fewer friends extending the positive link. From $\eta_k(\mathbf{g}^*) < \eta_i(\mathbf{g}^*)\ \forall i \in N\backslash\{k\}$ it then follows that $g^*_{k,i} = 1\ \forall i \in N\backslash\{k\}$, and a deviation for agent k therefore consists of extending negative links to some subset of $N\backslash\{k\}$. For each agent, to which k extends a negative link in a deviation, payoffs decrease by κ and no such deviation is profitable. Next, agent i. There are three types of deviations to consider. First, i extends a positive link to k. This decreases i's payoffs strictly, as $\eta_k(\mathbf{g}^*) < \eta_i(\mathbf{g}^*)$, while payoffs from all other agents remain *zero* in $\mathbf{g}'$.

Second, i extends negative links to some subset of $N\backslash\{k,i\}$. From Lemma 2 we know that i's payoffs will decrease from his link with k. Furthermore, extending negative links to some subset $N\backslash\{i,k\}$ will at most leave payoffs constant from those links, by an argument identical to the one used in Proposition 2. Proposed deviation is therefore not profitable. Third, a combination of the two deviations above. Assume first i extends a positive link to k and *one* negative link to some $j \in N\backslash\{k,i\}$. $i's$ payoffs will remain constant if $n=3$, as then $\eta_j(\mathbf{g}') = \eta_k(\mathbf{g}^*) = 0$ and $\eta_{i,j}(\mathbf{g}') = \eta_{i,k}(\mathbf{g}^*) = 0$. For $n \geq 4$, i will strictly decrease payoffs. To see this, note that $\eta_i(\mathbf{g}') = \eta_i(\mathbf{g}^*) = n-2$, but $\eta_j(\mathbf{g}') = \eta_{i,j}(\mathbf{g}') = n-3$ while $\eta_k(\mathbf{g}^*) = \eta_{i,k}(\mathbf{g}^*) = 0$. The deviation considered is again not profitable, which follows from $f(\eta_i, \eta_j - 1, \eta_{i,j} - 1) > f(\eta_i, \eta_j, \eta_{i,j})$. More precisely, $f(n-2,0,0) > f(n-2, n-3, n-3)$ for $n \geq 4$. Extending more than one negative link to $N\backslash\{k,i\}$ yields even lower payoffs, again by the argument used in Proposition 2. $Q.E.D.$

3.1 Comparative Statics

Next, we present a simple comparative statics analysis for the contest success function in difference form. The decisiveness parameter ϕ determines the coercion or conflict technology. A higher value of ϕ is favorable for the agent with more friends, while a lower value of ϕ is favorable for the agent with fewer friends.

Proposition 4: *For ϕ sufficiently low, there are at most two maximal cliques in any NE $\mathbf{g}^*$.*

Proof. We first show that for ϕ sufficiently low, there exists a profitable deviation for any network configuration that is in accordance with Proposition 1 and displays three sets of different size: $\mid P^1_{x-1}(\mathbf{g}^*) \mid = x$, $\mid P^2_{y-1}(\mathbf{g}^*) \mid = y$ and $\mid P^3_{z-1}(\mathbf{g}^*) \mid = z$ with $1 \leq x < y < z$. More specifically, we check for a deviation of agent $i \in P^2(\mathbf{g}^*)$, consisting of extending positive links to all agents $j \in P^1(\mathbf{g}^*)$. Marginal payoffs are given by

$$\Pi_i(\mathbf{g}') - \Pi_i(\mathbf{g}^*) = z\left(\frac{e^{\phi(x+y-1)}}{e^{\phi(x+y-1)} + e^{\phi(z-1)}}\right) - x\left(\frac{e^{\phi(y-1)}}{e^{\phi(y-1)} + e^{\phi(x-1)}}\right) - z\left(\frac{e^{\phi(y-1)}}{e^{\phi(y-1)} + e^{\phi(z-1)}}\right).$$

Taking the derivative and setting $\phi = 0$ yields

$$\frac{1}{4}(x-y+z)>0.$$

The first derivative is continuous in ϕ and therefore, for $\phi>0$ and ϕ sufficiently close to *zero*, proposed deviation is profitable. Next, note that for any additional clique(s) with $\mid P_{l-1}(\mathbf{g}^*)\mid = l > z$, proposed deviation yields even higher marginal payoffs, as i now also increases his payoffs from any $k \in P_{l-1}(\mathbf{g}^*)$. *Q.E.D.*

Proposition 5: *For ϕ sufficiently high, $\mathbf{g}^*$ is a NE iff*

- $\sum_{i=1}^{j-1} \mid P^i(\mathbf{g}^*)\mid < \mid P^j(\mathbf{g}^*)\mid$,
- $\bar{g}^*_{i,j} \neq -2$ *for all* $i,j \in N$,
- *if* $\bar{g}^*_{i,j} = -1$ *with* $\eta_i(\mathbf{g}^*) < \eta_j(\mathbf{g}^*)$, *then* $g^*_{i,j} = 1$ *and*
- $\forall i,j$ *with* $\eta_i(\mathbf{g}^*) = \eta_j(\mathbf{g}^*)$, *then* $\bar{g}^*_{i,j} = 1$ *and* $\forall i,j$ *with* $\eta_i(\mathbf{g}^*) \neq \eta_j(\mathbf{g}^*)$, *then* $\bar{g}^*_{i,j} = -1$.

Proof. For $\eta_j(\mathbf{g}^*) > \eta_i(\mathbf{g}^*)$ and ϕ sufficiently high, the payoffs that j coerces from i are bounded by and arbitrarily close to $\frac{1}{2}$. Assume now that Lemma 3, Lemma 4 and Proposition 1 hold, but contrary to the claim in Proposition 5, that there exists a Nash equilibrium $\mathbf{g}^*$, with $\sum_{i=1}^{j-1} \mid P^i(\mathbf{g}^*)\mid \geq \mid P^j(\mathbf{g}^*)\mid$ for some $P^j(\mathbf{g}^*)$. Consider a deviation where $k \in P^{j-1}(\mathbf{g}^*)$ extends positive links to a subset of agents in $P^1(\mathbf{g}^*) \cup P^2(\mathbf{g}^*) \cup \ldots \cup P^{j-2}(\mathbf{g}^*)$, such that $\eta_k(\mathbf{g}') = \eta_j(\mathbf{g}^*)$ for $j \in P^j(\mathbf{g}^*)$. This is feasible, as $\sum_{i=1}^{j-1} \mid P^i(\mathbf{g}^*)\mid \geq \mid P^j(\mathbf{g}^*)\mid$. Denote with x the number of agents, to which k needs to extend positive links to, such that $\eta_k(\mathbf{g}') = \eta_j(\mathbf{g}^*)$. To see that proposed deviation is profitable, note that $\mid P^{j-1}(\mathbf{g}^*)\mid + x = \mid P^j(\mathbf{g}^*)\mid$ with $x < \mid P^j(\mathbf{g}^*)\mid$. For ϕ sufficiently high, k forgoes coercion payoffs of arbitrarily close to $\frac{1}{2}$ from each of the x agents to which he extends positive links to. However, he will increase payoffs of arbitrarily close to $\frac{1}{2}$ from the $\mid P^j(\mathbf{g}^*)\mid$ agents in $P^j(\mathbf{g}^*)$. From $\mid P^j(\mathbf{g}^*)\mid > x$ it then follows that proposed deviation is profitable. For sufficiency, note that if $\sum_{i=1}^{j-1} \mid P^i(\mathbf{g}^*)\mid < \mid P^j(\mathbf{g}^*)\mid$ holds, then there is no agent who, by extending positive links to agents in smaller maximal cliques, can obtain at least as many friends as agents in any of the larger sets. Payoffs from a deviation of agent i are negative for ϕ sufficiently high, as the deviating agent will loose out on payoffs of arbitrarily close to $\frac{1}{2}$ for each agent he extends a positive link to, while increasing his payoffs by arbitrarily close to *zero* from any of his remaining negative links. *Q.E.D.*

Note that in Proposition 5 we only need the coercion payoff to be bounded and the result therefore also goes through for the contest success function in ratio form. The condition $\sum_{i=1}^{j-1} \mid P^i(\mathbf{g}^*) \mid < \mid P^j(\mathbf{g}^*) \mid$ implies that group size is at least geometrically increasing with common ratio 2.

3.2 Heterogeneous Agents

In this section we allow for differences in intrinsic strength, which is denoted by λ. An agent i's coercive power in network $\mathbf{g}$ is the sum of his own intrinsic strength λ_i and the intrinsic strength of his friends. That is, $\eta_i(\mathbf{g}) = \lambda_i + \sum_{j \in N_i^+(\mathbf{g})} \lambda_j$. We adapt the assumption $f(\eta_i + 1, \eta_j, \eta_{i,j} + 1) > f(\eta_i, \eta_j, \eta_{i,j})$ to $f(\eta_i + \lambda, \eta_j, \eta_{i,j} + \lambda) > f(\eta_i, \eta_j, \eta_{i,j})$. The proofs for Lemma 1, 2 and 3 are identical to the case with homogenous agents and are omitted here.

Proposition 6: *If there exists a pair of agents i and j such that $\lambda_i \neq \lambda_j$, then there does not exist a NE $\mathbf{g}^*$, such that everyone is friends with everyone.*

Proof. Assume that $\lambda_i > \lambda_j$ and to the contrary to above statement that $\bar{g}^*_{i,j} = 1$ $\forall i, j \in N$. Then i can profitably deviate by extending a negative link to j. *Q.E.D.*

Proposition 7: *In any NE $\mathbf{g}^*$, if $\eta_i(\mathbf{g}^*) = \eta_j(\mathbf{g}^*)$, then $\bar{g}^*_{i,j} = 1$ and if $\eta_i(\mathbf{g}^*) \neq \eta_k(\mathbf{g}^*)$, then $\bar{g}^*_{i,j} = -1$.*

Proof. The proof is identical to the one in Proposition 1 when redefining $\eta_i(\mathbf{g}) = \lambda_i + \sum_{j \in N_i^+(\mathbf{g})} \lambda_j$ and $P_k(\mathbf{g}) = \{i \in N \mid \eta_i(\mathbf{g}) = k\}$. *Q.E.D.*

Note that in the heterogeneous case the sets $P_k(\mathbf{g}^*)$ need not be of different size, but of different ex-post coercive strength. Furthermore, existence is not always guaranteed, as now an agent in $P_k(\mathbf{g}^*)$ may find it profitable to extend a negative link to another agent in $P_k(\mathbf{g}^*)$ with lower intrinsic strength.

4 Conclusion

This paper presents a simple model of *signed* network formation, where agents enter positive (friendship or alliance) and negative (coercion or conflict) relationships. Agents with more friends coerce payoffs from enemies with fewer friends. The coercive power of an agent is determined endogenously.

There are three main insights to be drawn. First, the model shows how in this context self-interested behavior yields the following sharp structural predictions under Nash equilibrium. Either everyone is friends with everyone, or agents can be partitioned into distinct sets, also called cliques, such that agents within the same set are friends and agents in different sets are enemies. This mirrors results on signed networks obtained in sociology, social psychology, international relations and applied physics. Second, cliques are of *different* size. This constitutes a departure from the notion of structural balance, as balanced outcomes allow for cliques of equal size. It also stands in contrast to models of coalition formation in the economics of conflict literature, where stable group structures are typically shown to be symmetric. Third, the game-theoretic approach allows us to address questions concerning the relative size and number of cliques, which could previously not be answered. For the contest success function in difference form we show that, if coercion is relatively less profitable, then there can be at most two cliques in any Nash equilibrium. If, on the other hand, coercion is sufficiently profitable, then multiple cliques may arise and relative group size is at least geometrically increasing.

To the best of my knowledge, this is the first game-theoretic model of *signed* network formation and there are various directions for future research. It appears promising to allow for investment in arming and production. This may provide new insights regarding the so-called trade-off between guns vs. butter. Introducing dynamic considerations and/or incomplete information may yield a model that allows us to study open conflict in the present context.

5 References

1. Antal, T., P. Krapivsky and S. Redner (2005) Dynamics of social balance on networks, *Phys. Rev. E,* 72:036121

2. Antal, T., P. Krapivsky and S. Redner (2006) Social balance on networks: The dynamics of friendship and enmity, *Physica D*, 224(130).

3. Aumann, R. and R. Myerson (1988) Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value, in: *The Shapley Value*, edited by A. Roth , Cambridge University Press, 175-191.

4. Bala, V. and S. Goyal (2000) A non-cooperative model of network formation, *Econometrica*, Vol. 68, 1181-1230.

5. Davis, J. A. (1967) Clustering and structural balance in graphs, *Human Relations*, 20(2), 181-187.

6. Cartwright D. and F. Harary (1956) Structural balance: A generalization of Heider's theory, *Psychological Review*, 63(5), 277-293.

7. Chwe, M. (1994) Farsighted Coalitional Stability, *J. Econ. Theory*, Vol. 63, 299-325.

8. Doreian, P. and D. Krackhardt (2001) Pre-transitive mechanisms for signed networks, *Journal of Mathematical Sociology*, Vol. 25, 43-67.

9. Doreian, P. and A. Mrvar (1996) A partitioning approach to structural balance, *Social Networks*, Vol. 18, 149-168.

10. Easley, D. and J. Kleinberg (2010) *Networks, Crowds and Markets: Reasoning About a Highly Connected World*, Cambridge University Press.

11. Esteban, J.M. and J. Sákovics (2003) Olson vs. Coase: Coalition worth in conflict, *Theory and Decision*, Vol. 55, 339-357.

12. Facchetti, G., G. Iacono and C. Altafini (2011) Computing global structure in large-scale signed social networks, *PNAS*, Vol. 108, 20953-20958

13. Frank, J. and T. Öztürk (2009) Conflict Networks, Ruhr Economic Papers, 116.

14. Goyal, S. and A. Vigier (2010) Robust networks, mimeo.

15. Goyal, S. and F. Vega-Redondo (2007) Structural Holes, *J. Econ. Theory*, Vol. 137, 460 – 492.

16. Harary, F. (1961) A structural analysis of the situation in the Middle East in 1956, *Journal of Conflict Resolution*, Vol. 5, 167-178.

17. Heider, F. (1946) Attitudes and cognitive organization, *Journal of Psychology*, Vol. 21, 107-112.

18. Jordan, J. (2006) Pillage and property, *J. Econ. Theory*, Vol. 131, 26-44

19. Jackson, M.O. and A. Wolinsky (1996) A Strategic Model of Social and Economic Networks, *J. Econ. Theory*, Vol. 71, 44-74.

20. Marvel S. A., S. H. Strogatz and J. M. Kleinberg (2009) Energy landscape of social balance, *Phys. Rev. Lett.* 103:198701.

21. Marvel S. A., J. M. Kleinberg, R. D. Kleinberg and S. H. Strogatz (2011) Continuous-time model of structural balance, *PNAS*, Vol. 108, 1771-1776.

22. Moore, M. (1979) Structural balance and international relations, *European Journal of Social Psychology*, Vol. 9, 323-326.

23. Myerson, R. B. (1977) Graphs and Cooperation in Games, *Mathematics of Operations Research* 2, 225–229.

24. Rietzke, D. and B. Roberson (2012) The robustness of 'enemy-of-my-enemy-is-my-friend' alliances, *Social Choice and Welfare*, Springer, published online January 2012

25. Szell, M., R. Lambiotte and S. Thurner (2010) Multirelational organization of large-scale social networks in an online world, *PNAS*, Vol. 107, 13636-13641.

26. Tullock, G. (1967) The Welfare Cost of Tariffs, Monopoly, and Theft, *Western Economic Journal*, 224-232.

27. Tullock, G. (1980) Efficient Rent Seeking, in: *Toward a Theory of Rent Seeking Society,* edited by J. Buchanan, R. Tollison and G. Tullock, Texas A&M University Press, 97-112.

28. Wärneryd, K. (1998) Distributional conflict and jurisdictional organization, *J. Public Econ.*, Vol. 69, 435-450.

29. Willer, D., (1999) Actors in Relations, in: *Network Exchange Theory*, edited by D. Willer, Praeger Publishers, 23-48.

Appendix A

Recall that in Proposition 1 we focus on the case where a pair of agents i and j exists, such that $\eta_i(\mathbf{g}^*) \neq \eta_j(\mathbf{g}^*)$, while the case where $\eta_i(\mathbf{g}^*) = \eta_j(\mathbf{g}^*)$ for all $i, j \in N$ is covered in Proposition 2 as part of the existence results.

Proposition 1: *In any NE* $\mathbf{g}^*$, *if* $\eta_i(\mathbf{g}^*) = \eta_j(\mathbf{g}^*)$, *then* $\bar{g}^*_{i,j} = 1$ *and if* $\eta_i(\mathbf{g}^*) \neq \eta_j(\mathbf{g}^*)$, *then* $\bar{g}^*_{i,j} = -1$.

Proof. The proof uses an induction argument and we start by proving the base case in four steps.

Base Case: *In any NE* $\mathbf{g}^*$, $\bar{g}^*_{i,j} = 1$ $\forall i, j \in P^m(\mathbf{g}^*)$ *and* $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ *and* $\forall k \notin P^m(\mathbf{g}^*)$.

Step 1: *In any NE* $\mathbf{g}^*$, $\bar{g}^*_{i,j} = 1$ $\forall i, j \in P^m(\mathbf{g}^*)$.

Assume to the contrary that $\bar{g}^*_{i,j} = -1$ for some pair of agents $i, j \in P^m(\mathbf{g}^*)$ and, without loss of generality, that $g^*_{i,j} = -1$. If there exists an agent k, such that $k \notin P^m(\mathbf{g}^*)$ and $k \in N^-_i(\mathbf{g}^*)$, then $g^*_i + g^+_{i,j}$ is a profitable deviation, as payoffs from $i's$ link with j remain *zero*, while payoffs from all of $i's$ negative links strictly increase by Lemma 2. If there does not exist an agent k, such that $k \notin P^m(\mathbf{g}^*)$ and $k \in$

$N_i^-(\mathbf{g}^*)$, then i can profitably deviate with $g_i^* + g_{i,j}^+ + g_{i,k}^-$. Agent i obtains a strictly positive payoff after proposed deviation because $\eta_k(\mathbf{g}') < \eta_i(\mathbf{g}')$, compared with a payoff of *zero* prior to it.

Step 2: *In any NE* $\mathbf{g}^*$, $N_i^+(\mathbf{g}^*)\backslash\{j\} = N_j^+(\mathbf{g}^*)\backslash\{i\} \wedge N_i^-(\mathbf{g}^*) = N_j^-(\mathbf{g}^*)$ $\forall i, j \in P^m(\mathbf{g}^*)$.

The statement holds trivially for $\mid P^m(\mathbf{g}^*) \mid = 1$. Assume $\mid P^m(\mathbf{g}^*) \mid \geq 2$ and, contrary to the above, that $\exists i, j \in P^m(\mathbf{g}^*) : N_i^+(\mathbf{g}^*)\backslash\{j\} \neq N_j^+(\mathbf{g}^*)\backslash\{i\} \vee N_i^-(\mathbf{g}^*) \neq N_j^-(\mathbf{g}^*)$. That is, there exists a pair of agents i and j, such that their respective sets of friends and enemies are different. From Step 1 and Lemma 4 we know that $g_{k,i} = 1$ $\forall k \in N\backslash\{i\}$. Therefore, for $i, j \in P^m(\mathbf{g}^*) : N_i^+(\mathbf{g}^*)\backslash\{j\} \neq N_j^+(\mathbf{g}^*)\backslash\{i\} \vee N_i^-(\mathbf{g}^*) \neq N_j^-(\mathbf{g}^*)$ to hold, it must be that i and j play different strategies relative to third agents, which we denote with $g_{i\backslash j}^* \neq g_{j\backslash i}^*$. Without loss of generality, assume $\Pi_j(\mathbf{g}^*) \geq \Pi_i(\mathbf{g}^*)$. We show that i can strictly increase his payoffs by imitating $j's$ strategy (while keeping his positive link to j), so that $g'_{i\backslash j} = g_{j\backslash i}^*$. More specifically, $\Pi_i(\mathbf{g}') > \Pi_j(\mathbf{g}^*) \geq \Pi_i(\mathbf{g}^*)$. There are two types of agents to consider when comparing the payoffs of i after proposed deviation, $\Pi_i(\mathbf{g}')$, with payoffs of j prior to it, $\Pi_j(\mathbf{g}^*)$. First, the agents that are i and $j's$ enemies prior to the deviation, $k \in N_j^-(\mathbf{g}^*) \cap N_i^-(\mathbf{g}^*)$. Second, the agents that are $j's$ enemies, but $i's$ friends prior to the deviation, $l \in N_j^-(\mathbf{g}^*) \cap N_i^+(\mathbf{g}^*)$. We start by showing that i's payoffs from links to agents $k \in N_j^-(\mathbf{g}^*) \cap N_i^-(\mathbf{g}^*)$ in $\mathbf{g}'$ are equal to the payoffs that j obtains from these agents in $\mathbf{g}^*$. Note first that, as i is imitating $j's$ strategy, $\eta_i(\mathbf{g}') = \eta_j(\mathbf{g}^*)$. Because i does not change his strategy relative to $k \in N_j^-(\mathbf{g}^*) \cap N_i^-(\mathbf{g}^*)$, $\eta_k(\mathbf{g}^*) = \eta_k(\mathbf{g}')$ also holds. Finally, the number of $i's$ common friends with k, $\eta_{i,k}(\mathbf{g}')$, is the same for i in $\mathbf{g}'$ as the number of $j's$ common friends with k, $\eta_{j,k}(\mathbf{g}^*)$, in $\mathbf{g}^*$. To see this, note that k's sets of friends and enemies are the same in $\mathbf{g}'$ and $\mathbf{g}^*$. Therefore, from $\eta_i(\mathbf{g}') = \eta_j(\mathbf{g}^*)$, $\eta_k(\mathbf{g}') = \eta_k(\mathbf{g}^*)$ and $\eta_{i,k}(\mathbf{g}') = \eta_{j,k}(\mathbf{g}^*)$ we know that i's payoffs from all links with agents $k \in N_j^-(\mathbf{g}^*) \cap N_i^-(\mathbf{g}^*)$ are the same for i in $\mathbf{g}'$ as for j in $\mathbf{g}^*$. Next, payoffs from agents that are $j's$ enemies, but $i's$ friends prior to the deviation, i.e., $l \in N_j^-(\mathbf{g}^*) \cap N_i^+(\mathbf{g}^*)$. From $g_{i\backslash j}^* \neq g_{j\backslash i}^*$ and $i, j \in P^m(\mathbf{g}^*)$ we know that at least one such agent l exists. Again, as i imitates j's strategy, $\eta_i(\mathbf{g}') = \eta_j(\mathbf{g}^*)$. Since i extends a negative link to l in $\mathbf{g}'$ and a positive one in $\mathbf{g}^*$,

$\eta_l(\mathbf{g}') = \eta_l(\mathbf{g}^*) - 1$. Next, notice that while i is a common friend of j and l in $\mathbf{g}^*$, j is not a common friend of i and l in $\mathbf{g}'$ and therefore $\eta_{i,l}(\mathbf{g}') = \eta_{j,l}(\mathbf{g}^*) - 1$. From $f(\eta_i, \eta_j - 1, \eta_{i,j} - 1) > f(\eta_i, \eta_j, \eta_{i,j})$ (Lemma 1) we then know that i's payoffs in $\mathbf{g}'$ from all $l \in N_j^-(\mathbf{g}^*) \cap N_i^+(\mathbf{g}^*)$ are higher than j's payoffs from these agents in $\mathbf{g}^*$. We can now conclude that $\Pi_i(\mathbf{g}') > \Pi_j(\mathbf{g}^*) \geq \Pi_i(\mathbf{g}^*)$. Proposed deviation is profitable.

Step 3: *In any NE* $\mathbf{g}^*$, $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ *and* $\forall k \in P^{m-1}(\mathbf{g}^*)$.

Assume to the contrary that there exists an agent $k \in P^{m-1}(\mathbf{g}^*)$ such that $\bar{g}^*_{i,k} = 1$ $\forall i \in P^m(\mathbf{g}^*)$. From $k \in P^{m-1}(\mathbf{g}^*)$ it follows that $N_k^-(\mathbf{g}^*) \neq \emptyset$ and, by an argument analogous to the one used in Step 1, $\bar{g}^*_{j,k} = 1$ $\forall j, k \in P^{m-1}(\mathbf{g}^*)$. By Lemma 4 we know that $g_{h,k} = 1$ $\forall h \in P^{m-x}(\mathbf{g}^*)$ for $x \geq 2$. Therefore, $g^*_{l,k} = 1$ $\forall l \in N \backslash \{k\}$. We can now discern two cases. If $g^*_{k \backslash i} \neq g^*_{i \backslash k}$, then we can use the same argument as in Step 2 to show that either k or i (or both) can strictly increase payoffs by imitating the respective other agent's strategy. If, on the other hand, $g^*_{k \backslash i} = g^*_{i \backslash k}$, then we reach an immediate contradiction, as $\eta_k(\mathbf{g}^*) = \eta_i(\mathbf{g}^*)$ for some $k \in P^{m-1}(\mathbf{g}^*)$ and $i \in P^m(\mathbf{g}^*)$.

Step 4: *In any NE* $\mathbf{g}^*$, $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ *and* $\forall k \notin P^m(\mathbf{g}^*)$.

If there are only two sets of agents with different numbers of friends, $P^m(\mathbf{g}^*)$ and $P^{m-1}(\mathbf{g}^*)$, then we are done by Step 3. Assume that there are at least three such sets. We first show that $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ and $\forall k \in P^{m-2}(\mathbf{g}^*)$. Assume to the contrary that there exists a pair of agents $i \in P^m(\mathbf{g}^*)$ and $k \in P^{m-2}(\mathbf{g}^*)$ such that $\bar{g}^*_{i,k} = 1$. From Step 2 we know that then $\bar{g}^*_{i,k} = 1$ $\forall i \in P^m(\mathbf{g}^*)$ must hold. The argument used here is similar to Step 2. Recall from Step 2 that an agent i in $P^m(\mathbf{g}^*)$ is able to obtain the same sets of friends and enemies as another agent j in $P^m(\mathbf{g}^*)$, simply by imitating j's strategy. For an agent k in $P^{m-2}(\mathbf{g}^*)$, however, it may be the case that some agents in $P^{m-1}(\mathbf{g}^*)$ extend negative links to k. That is, when k is to obtain the same sets of friends and enemies in $\mathbf{g}'$ as agent $i \in P^m(\mathbf{g}^*)$ in $\mathbf{g}^*$, then the deviation strategy will depend on the linking behavior of agents in $P^{m-1}(\mathbf{g}^*)$. As before we discern two cases. Assume first that $\Pi_i(\mathbf{g}^*) > \Pi_k(\mathbf{g}^*)$. Agent k can profitably deviate with the following strategy. For all agents with no more friends than k, that is, for all $l \in P^{m-x}(\mathbf{g}^*)$ with $x \geq 2$ (and therefore $g^*_{l,k} = 1$ by Lemma 4 and the argument used in Step 1), imitate the sign of the directed link played by agent $i \in P^m(\mathbf{g}^*)$. If $g^*_{i,l} = -1$, then $g'_{k,l} = -1$ and if $g^*_{i,l} = 1$, then $g'_{k,l} = 1$. For $l \in P^{m-1}(\mathbf{g}^*)$ it may be the case that $g^*_{l,k} = -1$. Recall from Step 3 that $\bar{g}^*_{i,l} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ and $\forall l \in P^{m-1}(\mathbf{g}^*)$. Therefore, relative to $l \in P^{m-1}(\mathbf{g}^*)$, the deviation strategy is as follows. If $g^*_{l,k} = -1$, then $g'_{k,l} = 1$ and if $g^*_{l,k} = 1$ then $g'_{l,k} = -1$. We have constructed a deviation such that $i's$ friends in $\mathbf{g}^*$ are $k's$ friends in $\mathbf{g}'$ and $i's$ enemies in $\mathbf{g}^*$ are $k's$ enemies in $\mathbf{g}'$. Next, we use the underlying argument of Step 2 to show that proposed deviation is, in fact, profitable. Note, however, since k has strictly less friends than i in $\mathbf{g}^*$ (and therefore strictly more enemies), it may be that $N^-_i(\mathbf{g}^*) \subsetneq N^-_k(\mathbf{g}^*)$ and there then does not exist an agent $l \in N^-_i(\mathbf{g}^*) \cap N^+_k(\mathbf{g}^*)$. In this case $\Pi_k(\mathbf{g}') = \Pi_i(\mathbf{g}^*)$ holds, because payoffs from all agents in $N^-_i(\mathbf{g}^*) \cap N^-_k(\mathbf{g}^*)$ are the same for k in $\mathbf{g}'$ as for i in $\mathbf{g}^*$, as shown in Step 2. If an agent $l \in N^-_i(\mathbf{g}^*) \cap N^+_k(\mathbf{g}^*)$ does exist, then $\Pi_k(\mathbf{g}') > \Pi_i(\mathbf{g}^*)$, again as shown in Step 2. Therefore, $\Pi_k(\mathbf{g}') \geq \Pi_i(\mathbf{g}^*) > \Pi_k(\mathbf{g}^*)$ and proposed deviation is profitable. Assume next that $\Pi_k(\mathbf{g}^*) \geq \Pi_i(\mathbf{g}^*)$. Agent i can profitably deviate with the following strategy. If $\bar{g}^*_{k,l} = -1$, then $g'_{i,l} = -1$ and if $\bar{g}^*_{k,l} = 1$, then $g'_{i,l} = 1$ $\forall l \in N \backslash \{i\}$. From $i \in P^m(\mathbf{g}^*)$ we know that all remaining agents extend positive links to i and proposed deviation yields i the same sets of friends and enemies in $\mathbf{g}'$ as agent k in $\mathbf{g}^*$. From the argument in Step 2 it follows that $\Pi_i(\mathbf{g}') > \Pi_k(\mathbf{g}^*) \geq \Pi_i(\mathbf{g}^*)$. Therefore, in any Nash equilibrium $\mathbf{g}^*$, $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ and $\forall k \in P^{m-1}(\mathbf{g}^*) \cup P^{m-2}(\mathbf{g}^*)$. We can repeat the above argument $\forall k \in P^{m-3}(\mathbf{g}^*)$. Doing so iteratively yields $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ and $\forall k \notin P^m(\mathbf{g}^*)$.

Define the super set $\tilde{P}^r(\mathbf{g}^*) = P^m(\mathbf{g}^*) \cup P^{m-1}(\mathbf{g}^*) \cup \ldots \cup P^{m-r-1}(\mathbf{g}^*) \cup P^{m-r}(\mathbf{g}^*)$. Note that $\tilde{P}^0(\mathbf{g}^*) = P^m(\mathbf{g}^*)$.

Inductive Step: *In any NE $\mathbf{g}^*$, if $\bar{g}^*_{i,j} = 1$ $\forall i,j \in P^{m-x}(\mathbf{g}^*)$ and $\bar{g}^*_{i,k} = -1$ $\forall i \in P^{m-x}(\mathbf{g}^*)$ and $\forall k \notin P^{m-x}(\mathbf{g}^*)$ holds $\forall x \in \mathbb{N} : 0 \leq x \leq r$, then $\bar{g}^*_{i,j} = 1$ $\forall i,j \in P^{m-(r+1)}(\mathbf{g}^*)$ and $\bar{g}^*_{i,k} = -1$ $\forall i \in P^{m-(r+1)}(\mathbf{g}^*)$ and $\forall k \notin P^{m-(r+1)}(\mathbf{g}^*)$.*

In Step 4 we show that $\bar{g}^*_{i,j} = 1$ $\forall i,j \in P^m(\mathbf{g}^*)$ $\wedge$ $\bar{g}^*_{i,k} = -1$ $\forall i \in P^m(\mathbf{g}^*)$ and $\forall k \notin P^m(\mathbf{g}^*)$. Assume the statement holds for all sets $P^{m-x}(\mathbf{g}^*)$ with $x \in \mathbb{N} : 0 \leq$

$x \leq r$. From Lemma 4 we know that $g^*_{i,k} = -1$ and $g^*_{k,i} = 1$ $\forall i \in \tilde{P}^r(\mathbf{g}^*)$ and $\forall k \notin \tilde{P}^r(\mathbf{g}^*)$, while from Lemma 3 we know that $\nexists \bar{g}^*_{i,k} = -2$. We can now repeat steps 1 through 4 from the base case, relabeling $P^m(\mathbf{g}^*)$ with $P^{m-(r+1)}(\mathbf{g}^*)$ and $P^{m-1}(\mathbf{g}^*)$ with $P^{m-(r+2)}(\mathbf{g}^*)$ to establish the above result. *Q.E.D*.

Appendix B - For Online Publication

This section presents a variation of the model, which allows for neutral (or *zero*) links and assumes no cost of conflict. Let $N = \{1, 2, ..., n\}$ be the set of ex-ante identical agents, with $n \geq 3$. A strategy for $i \in N$ is defined as a row vector $\mathbf{g}_i = (g_{i,1}, g_{i,2}, ..., g_{i,i-1}, g_{i,i+1}, ..., g_{i,n})$, but now instead of $g_{i,j} \in \{-1, 1\}$, we assume $g_{i,j} \in \{-1, 0, 1\}$ for each $j \in N \setminus \{i\}$. Agent i is said to extend a positive link to j, if $g_{i,j} = 1$, a negative link, if $g_{i,j} = -1$, and to extend a neutral (or *zero*) link, if $g_{i,j} = 0$. Define the undirected network $\bar{\mathbf{g}}$ in the following way. The link between agent i and j is positive in the undirected network $\bar{\mathbf{g}}$, if both directed links are positive, so that $\bar{g}_{i,j} = 1$, if $g_{i,j} = g_{j,i} = 1$. The link in the undirected network is negative, if at least one of the two directed links is negative, so that $\bar{g}_{i,j} = -1$, if $g_{i,j} = -1$ or $g_{j,i} = -1$. No link is created in the undirected network $\bar{\mathbf{g}}$, if either both agents involved extend a *zero* link or one agent extends a *zero* link and the other extends a positive link. That is, $\bar{g}_{i,j} = 0$, if either $g_{i,j} = g_{j,i} = 0$ or, without loss of generality, $g_{i,j} = 0$ and $g_{j,i} = 1$.

The sets of friends, enemies and neutral agents are defined accordingly: $N^+_i(\mathbf{g}) = \{j \in N \mid \bar{g}_{i,j} = 1\}$, $N^-_i(\mathbf{g}) = \{j \in N \mid \bar{g}_{i,j} = -1\}$ and $N^0_i(\mathbf{g}) = \{j \in N \mid \bar{g}_{i,j} = 0\}$. Denote the following cardinality with $\eta_i(\mathbf{g}) = \mid N^+_i(\mathbf{g}) \mid$. For ease of notation we will sometimes write η_i for $\eta_i(\mathbf{g})$.

The payoffs to player i under strategy profile $\mathbf{g}$ are given by

$$\Pi_i(\mathbf{g}) = \sum_{j \in N^-_i(\mathbf{g})} f(\eta_i(\mathbf{g}), \eta_j(\mathbf{g})).$$

Again assume that f is such that $f(\eta_i, \eta_j) + f(\eta_j, \eta_i) = 0$ for all η_i, η_j. If $\eta_i = \eta_j$, then $f(\eta_i, \eta_j) = 0$. $f(\eta_i, \eta_j)$ is increasing in η_i and decreasing in η_j. Write the forward difference as $\triangle_{\eta_i} f(\eta_i, \eta_j) = f(\eta_i + 1, \eta_j) - f(\eta_i, \eta_j)$. We now make the additional assumption that $\triangle_{\eta_i} f(\eta_i, \eta_j)$ is decreasing in η_i for given η_j. Note also that the number of common friends, $\eta_{i,j}$, does not enter the payoff function. Before commenting on the latter two assumptions, we define bilateral equilibrium (Goyal and Vega-Redondo, 2007), which allows for coordinated deviations of pairs of agents and refines Nash equilibrium.

Definition 1: *A strategy profile $\tilde{\mathbf{g}}^*$ is a bilateral equilibrium (BE) if*

- *for any $i \in N$ and every $g_i \in G_i$, $\Pi_i(\tilde{g}^*) \geq \Pi_i(g_i, \tilde{g}^*_{-i})$;*
- *for any pair of players $i, j \in N$ and every strategy pair g_i, g_j,*

  $\Pi_i(g_i, g_j, \tilde{g}^*_{-i-j}) > \Pi_i(\tilde{g}^*_i, \tilde{g}^*_j, \tilde{g}^*_{-i-j}) \Rightarrow \Pi_j(g_i, g_j, \tilde{g}^*_{-i-j}) < \Pi_j(\tilde{g}^*_i, \tilde{g}^*_j, \tilde{g}^*_{-i-j})$.

A strategy profile is a *BE* if no player can deviate unilaterally and no pair of players can deviate bilaterally and benefit from the deviation (in the case of bilateral deviations, at least one of them strictly). Bilateral equilibrium also refines the notion of pairwise stability (Jackson and Wolinsky, 1996), as it allows pairs of agents to form and delete (positive) links simultaneously.

Next, we describe why the assumption that $\triangle_{\eta_i} f(\eta_i, \eta_j)$ is decreasing in η_i is needed, and why the number of common friends, $\eta_{i,j}$, does not enter the payoff function in the current specification. Note first, that without cost of conflict, configurations may arise, where pairs of agents extend negative links to each other, i.e., $g_{i,j} = -1$ and $g_{j,i} = -1$. In this case both agents involved in the undirected negative link need to deviate (bilaterally) in order to create an alliance/friendship. This implies that an agent, even if he has more friends than any other agent, may only be able to switch *one* undirected negative link from a negative into a positive one (through a bilateral deviation). We may therefore not be able to rely on the argument, used extensively in the main part of the paper, that certain agents can profitably deviate by playing strategies, which yield the same sets of friends and enemies as some other agent with weakly higher payoffs. The additional assumption that $\triangle_{\eta_i} f(\eta_i, \eta_j)$ is decreasing in η_i for given η_j is needed to show that changing the sign of only *one* negative undirected link into a positive one is sufficient for the deviation in question to be profitable.

For similar reasons the number of common friends, $\eta_{i,j}$, does not enter the payoff function. Recall that in the main part of the paper we showed that an agent $i \in P^m(\mathbf{g}^*)$ could profitably deviate by imitating the strategy of an agent $j \in P^m(\mathbf{g}^*)$ with weakly higher payoffs. However, an analogous deviation where only *one* negative undirected link is switched into a positive one, and one positive undirected link is switched into a negative one, say to agent k, may not be profitable. This will be the case, if after the deviation the difference between $\eta_{i,k}$ and $\eta_{j,k}$ is sufficiently large.

Lemma 1 shows that there are no undirected neutral links in a bilateral equilibrium. The intuition is that, in the absence of linking cost, a pair of agents either finds it profitable to create a positive undirected link to increase payoffs from their negative links, or one of the agents wants to create a negative link and extract payoffs from the respective other agent. If extending positive and negative links was costly, then undirected neutral links may be part of a bilateral equilibrium. Introducing linking cost is, however, out of the scope of this paper and left for future research.

Lemma 1: *In any BE,* $\tilde{\tilde{g}}^*_{i,j} \neq 0 \; \forall i, j \in N$.

Proof. Assume to the contrary that a link $\tilde{\tilde{g}}^*_{i,j} = 0$ exists in a *BE* for some $i, j \in N$. We discern two cases. First, without loss of generality, $\eta_i(\tilde{\mathbf{g}}^*) > \eta_j(\tilde{\mathbf{g}}^*)$. Then i can profitably deviate by extending a negative link to j, thereby extracting payoffs from j. Second, $\eta_i(\tilde{\mathbf{g}}^*) = \eta_j(\tilde{\mathbf{g}}^*)$. If i and j sustain negative links, then i and j can profitably deviate by creating the positive link $\tilde{\tilde{g}}^*_{i,j} = 1$, thereby increasing payoffs on all remaining negative links, while the payoff from the link $\tilde{\tilde{g}}^*_{i,j}$ remains *zero*. If i and j do not sustain negative links and there exists a third agent k, such that $\eta_i(\tilde{\mathbf{g}}^*) = \eta_j(\tilde{\mathbf{g}}^*) \geq \eta_k(\tilde{\mathbf{g}}^*)$, then a profitable deviation consists of i and j creating the link $\tilde{\tilde{g}}^*_{i,j} = 1$ and extending a negative link to agent k. Deviation strategies are given by $\tilde{g}^*_i + g^+_{i,j} + g^-_{i,k}$ and $\tilde{g}^*_j + g^+_{j,i} + g^-_{j,k}$. Agents i and j then extract payoffs from k under $\mathbf{g}'$, while payoffs prior to the deviation are *zero*. If only one of the agents has a negative link, say agent i, then a profitable deviation consists of $\tilde{g}^*_i + g^+_{i,j}$ and $\tilde{g}^*_j + g^+_{j,i} + g^-_{j,k}$, by an analogous argument as above. Next, the case where i and j do not have any negative links and there does not exist a third agent k such that $\eta_i(\tilde{\mathbf{g}}^*) = \eta_j(\tilde{\mathbf{g}}^*) \geq \eta_k(\tilde{\mathbf{g}}^*)$. That is, $\eta_i(\tilde{\mathbf{g}}^*) = \eta_j(\tilde{\mathbf{g}}^*) < \eta_k(\tilde{\mathbf{g}}^*)$ $\forall k \in N \backslash \{i, j\}$. Take agent m with the weakly highest number of friends in $\tilde{\mathbf{g}}^*$. If all links of m are positive, then a profitable deviation consists of $\tilde{g}_m^* + g_{m,j}^-$. Agent m obtains a strictly positive payoff after proposed deviation, as opposed to a payoff of *zero* prior to it. If m has a negative link to some agent k (other than i and j), then a profitable deviation consists of creating a positive link with k and extending, without loss of generality, a negative link to agent j. Deviation strategies are given by $\tilde{g}_m^* + g_{m,k}^+ + g_{m,j}^-$ and $\tilde{g}_k^* + g_{k,m}^+ + g_{k,j}^-$. This is profitable for agent m as $\eta_m(\mathbf{g}') \geq \eta_m(\tilde{\mathbf{g}}^*)$ and $\eta_j(\mathbf{g}') < \eta_k(\tilde{\mathbf{g}}^*)$. For agent k the deviation is profitable from $\eta_k(\mathbf{g}') \geq \eta_k(\tilde{\mathbf{g}}^*)$ and $\eta_j(\mathbf{g}') < \eta_m(\tilde{\mathbf{g}}^*)$. If only one of the agents i and j has a negative link, say agent i, then $\tilde{g}_m^* + g_{m,k}^+ + g_{m,j}^-$ and $\tilde{g}_k^* + g_{k,m}^+ + g_{k,j}^-$ is profitable by the above argument. *Q.E.D.*

For the equilibrium characterization, define again the set of agents with k friends, $P_k(\mathbf{g}) = \left\{ j \in N \mid \eta_j(\mathbf{g}) = k \right\}$. Denote the set with the highest number of friends with $P^m(\mathbf{g})$, the one with the second highest subscript $P^{m-1}(\mathbf{g})$ and proceed in this way until the set of agents with the fewest number of friends, $P^1(\mathbf{g})$. We again focus first on the case where there are at least two agents such that $\eta_i(\tilde{\mathbf{g}}^*) \neq \eta_j(\mathbf{g}^*)$. Note that the statement of Proposition 1 in Appendix B is identical to the one of Proposition 1 in the main part of the paper.

Proposition 1: *In any BE* $\tilde{\mathbf{g}}^*$, *if* $\eta_i(\tilde{\mathbf{g}}^*) = \eta_j(\tilde{\mathbf{g}}^*)$ *then* $\tilde{\tilde{g}}_{i,j}^* = 1$, *and if* $\eta_i(\tilde{\mathbf{g}}^*) \neq \eta_j(\mathbf{g}^*)$ *then* $\tilde{\tilde{g}}_{i,j}^* = -1$.

Proof. The proof uses an induction argument and we start by proving the base case in five steps.

Base Case: *In any BE* $\tilde{\mathbf{g}}^*$, $\tilde{\tilde{g}}_{i,j}^* = 1$ $\forall i, j \in P^m(\tilde{\mathbf{g}}^*)$ *and* $\tilde{\tilde{g}}_{i,k}^* = -1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ *and* $\forall k \notin P^m(\tilde{\mathbf{g}}^*)$

Step 1: *In any BE* $\tilde{\mathbf{g}}^*$, $\tilde{\tilde{g}}_{i,j}^* = 1$ $\forall i, j \in P^m(\tilde{\mathbf{g}}^*)$.

Assume to the contrary that $\tilde{\tilde{g}}_{i,j}^* \neq 1$ for some pair of agents $i, j \in P^m(\tilde{\mathbf{g}}^*)$. If i and j sustain negative links to a third agent, i.e., if the sets $N_i^-(\tilde{\mathbf{g}}^*) \backslash \{j\}$ and $N_j^-(\tilde{\mathbf{g}}^*) \backslash \{i\}$ are non-empty, then creating the link $\tilde{\tilde{g}}'_{i,j} = 1$ (with deviation strategies $\tilde{\tilde{g}}_i^* + g_{i,j}^+$ and

$\tilde{\bar{g}}^*_j + g^+_{j,i}$) is profitable. Payoffs from the link between i and j remain *zero*, while they strictly increase on all remaining negative links. If one of the two agents, say i, does not sustain any negative links, then a profitable deviation consists of $\tilde{\bar{g}}^*_i + g^+_{i,j} + g^-_{i,k}$ with $k \in N\backslash\{i,j\}$. Payoffs prior to the deviation are *zero*, while they are strictly positive after it, since $\eta_k(\tilde{\mathbf{g}}^*) < \eta_i(\mathbf{g}')$. If j also does not sustain any negative links in $\tilde{\mathbf{g}}^*$, then j can profitably deviate with an analogous strategy, $\tilde{\bar{g}}^*_j + g^+_{i,j} + g^-_{i,k}$.

Step 2: *In any BE* $\tilde{\mathbf{g}}^*$, $N^+_i(\tilde{\mathbf{g}}^*)\backslash\{j\} = N^+_j(\tilde{\mathbf{g}}^*)\backslash\{i\}$ *and* $N^-_i(\tilde{\mathbf{g}}^*) = N^-_j(\tilde{\mathbf{g}}^*)$ $\forall i,j \in P^m(\tilde{\mathbf{g}}^*)$.

From Step 1 we know that $\tilde{\bar{g}}^*_{i,j} = 1$ $\forall i,j \in P^m(\tilde{\mathbf{g}}^*)$. Next, assume $\mid P^m(\tilde{g}^*) \mid \geq 2$ (the second part of above statement holds trivially for $\mid P^m(\tilde{g}^*) \mid = 1$) and, contrary to the above, that $\exists i,j \in P^m(\tilde{\mathbf{g}}^*) : N^+_i(\tilde{\mathbf{g}}^*)\backslash\{j\} \neq N^+_j(\tilde{\mathbf{g}}^*)\backslash\{i\} \vee N^-_i(\tilde{\mathbf{g}}^*) \neq N^-_j(\tilde{\mathbf{g}}^*)$. Note that there then must exists a pair of agents $k,l \notin P^m(\tilde{\mathbf{g}}^*)$, such that agent k is an enemy of i, but a friend of j and agent l is a friend of i, but an enemy of j. That is, $k \in N^-_i(\tilde{\mathbf{g}}^*) \cap N^+_j(\tilde{\mathbf{g}}^*)$ and $l \in N^+_i(\tilde{\mathbf{g}}^*) \cap N^-_j(\tilde{\mathbf{g}}^*)$. Assume without loss of generality that $\eta_l(\tilde{\mathbf{g}}^*) \leq \eta_k(\tilde{\mathbf{g}}^*)$. Agents i and k can profitably deviate by creating the undirected positive link $\bar{g}'_{i,k} = 1$ and i extending a negative link to agent l. The corresponding deviation strategies are given by $\tilde{g}^*_i + g^+_{i,k} + g^-_{i,l}$ and $\tilde{g}^*_k + g^+_{k,i}$. Agent $i's$ number of friends remains the same, $\eta_i(\mathbf{g}') = \eta_i(\tilde{\mathbf{g}}^*)$, and therefore payoffs from agents other than k and l remain unchanged. However, $\eta_l(\tilde{\mathbf{g}}^*) \leq \eta_k(\tilde{\mathbf{g}}^*)$ implies that $\eta_l(\mathbf{g}') < \eta_k(\tilde{\mathbf{g}}^*)$ holds. Therefore, i obtains higher payoffs from l under $\mathbf{g}'$ than what he obtains from k under $\tilde{\mathbf{g}}^*$. Proposed deviation is also profitable for k, as $k's$ payoffs from his link with i are negative in $\tilde{\mathbf{g}}^*$, while they are *zero* in $\mathbf{g}'$. Furthermore, payoffs from any negative links increase because $\eta_k(\mathbf{g}') > \eta_k(\tilde{\mathbf{g}}^*)$.

Step 3: *Any BE* $\tilde{\mathbf{g}}^*$ *must be such that either i)* $\tilde{\bar{g}}^*_{i,k} = -1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ *and* $\forall k \in P^{m-1}(\tilde{\mathbf{g}}^*)$ *or ii)* $\tilde{\bar{g}}^*_{i,k} = 1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ *and* $\forall k \in P^{m-1}(\tilde{\mathbf{g}}^*)$.

Assume to the contrary, and in accordance with Step 2, that there exists an agent $k \in P^{m-1}(\tilde{\mathbf{g}}^*)$ such that $\tilde{\bar{g}}^*_{i,k} = 1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ and another agent $l \in P^{m-1}(\tilde{\mathbf{g}}^*)$ such that $\tilde{\bar{g}}^*_{i,l} = -1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$. Then agents i and l can deviate profitably by creating the undirected positive link $\bar{g}'_{i,l} = 1$ and agent i extending a negative link to agent k. The corresponding deviation strategies are given by $\tilde{g}^*_i + g^+_{i,l} + g^-_{i,k}$ and $\tilde{g}^*_l + g^+_{l,i}$. Agent

$i's$ number of friends remains the same after proposed deviation, while $\eta_k(\mathbf{g}') < \eta_l(\tilde{\mathbf{g}}^*)$. For agent l proposed deviation is profitable, as payoffs from his link with i are *zero* after proposed deviation, while they are negative under $\tilde{\mathbf{g}}^*$. Furthermore, payoffs from any negative links increase because $\eta_l(\mathbf{g}') > \eta_l(\tilde{\mathbf{g}}^*)$.

Step 4: *In any BE* $\tilde{\mathbf{g}}^*$, $\tilde{\bar{g}}^*_{i,k} = -1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ *and* $\forall k \in P^{m-1}(\tilde{\mathbf{g}}^*)$.

Assume to the contrary, but in accordance with Step 3, that $\tilde{\bar{g}}^*_{i,k} = 1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ and $\forall k \in P^{m-1}(\tilde{\mathbf{g}}^*)$. If there are only two sets, $P^m(\tilde{\mathbf{g}}^*)$ and $P^{m-1}(\tilde{\mathbf{g}}^*)$, then all links of $i \in P^m(\tilde{\mathbf{g}}^*)$ are positive. Agent i can then profitably deviate with $\tilde{g}^*_i + g^-_{i,k}$. Agent i's payoffs are positive under $\mathbf{g}'$ from $\eta_k(\mathbf{g}') < \eta_i(\mathbf{g}')$, while they are *zero* under $\tilde{\mathbf{g}}^*$. Assume next that there are more than two sets and the set $P^{m-2}(\tilde{\mathbf{g}}^*)$ therefore exists. We first show that i's enemies must be a strict subset of k's enemies, $N^-_i(\tilde{\mathbf{g}}^*) \subsetneq N^-_k(\tilde{\mathbf{g}}^*)$. Assume the contrary. Then there exists a pair of agents l and m such that agent l is a friend of i, but an enemy of k and agent m is an enemy of i, but a friend of k. That is, $l \in N^+_i(\tilde{\mathbf{g}}^*) \cap N^-_k(\tilde{\mathbf{g}}^*)$ and $m \in N^-_i(\tilde{\mathbf{g}}^*) \cap N^+_k(\tilde{\mathbf{g}}^*)$. For $\eta_l(\tilde{\mathbf{g}}^*) \leq \eta_m(\tilde{\mathbf{g}}^*)$ the following deviation is profitable: $\tilde{g}^*_i + g^+_{i,m} + g^-_{i,l}$ and $\tilde{g}^*_m + g^+_{m,i}$. Agent i increases payoffs as $\eta_i(\mathbf{g}') = \eta_i(\tilde{\mathbf{g}}^*)$ while $\eta_l(\mathbf{g}') < \eta_i(\tilde{\mathbf{g}}^*)$. The proposed deviation is profitable for m, as payoffs from his link with i are negative prior to the deviation and *zero* after it. Moreover, payoffs on any negative links increase. For $\eta_m(\tilde{\mathbf{g}}^*) \leq \eta_l(\tilde{\mathbf{g}}^*)$, $\tilde{g}^*_k + g^+_{k,l} + g^-_{k,m}$ and $\tilde{g}^*_l + g^+_{l,k}$ is profitable by an analogous argument. We have so far established that in any bilateral equilibrium $\tilde{\mathbf{g}}^*$, $N^-_i(\tilde{\mathbf{g}}^*) \subsetneq N^-_k(\tilde{\mathbf{g}}^*)$ for $i \in P^m(\tilde{\mathbf{g}}^*)$ and $k \in P^{m-1}(\tilde{\mathbf{g}}^*)$. To show that $\tilde{\bar{g}}^*_{i,k} = -1$ $\forall i \in P^m(\tilde{\mathbf{g}}^*)$ and $\forall k \in P^{m-1}(\tilde{\mathbf{g}}^*)$ we discern two cases. First, $\Pi_k(\tilde{\mathbf{g}}^*) \geq \Pi_i(\tilde{\mathbf{g}}^*)$. Agent i can profitably deviate with the following strategy. If $\tilde{\bar{g}}^*_{k,l} = -1$ then $g'_{i,l} = -1$ and if $\tilde{\bar{g}}^*_{k,l} = 1$ then $g'_{i,l} = 1$ $\forall l \in N \backslash \{i\}$. Note that after proposed deviation i has the same sets of friends and enemies in $\mathbf{g}'$ as agent k in $\tilde{\mathbf{g}}^*$. To see this, recall that $N^-_i(\tilde{\mathbf{g}}^*) \subsetneq N^-_k(\tilde{\mathbf{g}}^*)$ and proposed deviation therefore only consists changing $\tilde{\bar{g}}^*_{i,l} = 1$ to $\bar{g}'_{i,l} = -1$ by extending negative links to agents $l \in N^+_i(\tilde{\mathbf{g}}^*) \cap N^-_k(\tilde{\mathbf{g}}^*)$. Next, we show that proposed deviation is profitable. Note first that i's payoffs from agents $l \in N^-_i(\tilde{\mathbf{g}}^*) \cap N^-_k(\tilde{\mathbf{g}}^*)$ are the same in $\mathbf{g}'$ as k's payoffs in $\tilde{\mathbf{g}}^*$, because $\eta_l(\mathbf{g}') = \eta_l(\tilde{\mathbf{g}}^*)$ and $\eta_i(\mathbf{g}') = \eta_k(\tilde{\mathbf{g}}^*)$. However, payoffs from agents $l \in N^+_i(\tilde{\mathbf{g}}^*) \cap N^-_k(\tilde{\mathbf{g}}^*)$ are strictly higher for i in $\mathbf{g}'$ than for k in $\tilde{\mathbf{g}}^*$, because $\eta_l(\mathbf{g}') < \eta_l(\tilde{\mathbf{g}}^*)$ and $\eta_i(\mathbf{g}') = \eta_k(\tilde{\mathbf{g}}^*)$.

At least one such agent l exists, which follows from $N_i^-(\tilde{\mathbf{g}}^*) \subsetneq N_k^-(\tilde{\mathbf{g}}^*)$. Therefore, $\Pi_i(\mathbf{g}') > \Pi_k(\tilde{\mathbf{g}}^*) \geq \Pi_i(\tilde{\mathbf{g}}^*)$ holds and proposed deviation is profitable. Second, $\Pi_i(\tilde{\mathbf{g}}^*) > \Pi_k(\tilde{\mathbf{g}}^*)$. Recall that $N_i^-(\tilde{\mathbf{g}}^*) \subsetneq N_k^-(\tilde{\mathbf{g}}^*)$. If there exists only one agent $l \in N_i^+(\tilde{\mathbf{g}}^*) \cap N_k^-(\tilde{\mathbf{g}}^*)$, then k and l can deviate with $\tilde{g}_k^* + g_{k,l}^+$ and $\tilde{g}_l^* + g_{l,k}^+$ such that $N_k^-(\tilde{\mathbf{g}}^*) = N_i^-(\mathbf{g}')$ and $N_k^+(\tilde{\mathbf{g}}^*) \backslash \{i\} = N_i^+(\mathbf{g}') \backslash \{k\}$. The deviation is profitable for k, as $\Pi_i(\tilde{\mathbf{g}}^*) = \Pi_k(\mathbf{g}') > \Pi_k(\tilde{\mathbf{g}}^*)$. For l the deviation is profitable, because payoffs from his link with k are negative under $\tilde{\mathbf{g}}^*$ and *zero* under $\mathbf{g}'$, while payoffs from any negative links increase due to $\eta_l(\mathbf{g}') > \eta_l(\tilde{\mathbf{g}}^*)$. If there are two or more agents $l, m \in N_i^+(\tilde{\mathbf{g}}^*) \cap N_k^-(\tilde{\mathbf{g}}^*)$ such that a negative link is reciprocated with a negative link, i.e., $\tilde{g}_{k,l}^* = \tilde{g}_{l,k}^* = -1$ and $\tilde{g}_{k,m}^* = \tilde{g}_{m,k}^* = -1$, then there is no bilateral deviation such that $N_k^-(\tilde{\mathbf{g}}^*) = N_i^-(\mathbf{g}')$ and $N_k^+(\tilde{\mathbf{g}}^*) \backslash \{i\} = N_i^+(\mathbf{g}') \backslash \{k\}$. However, in the following we show that, if it is profitable for $i \in P^m(\tilde{\mathbf{g}}^*)$ to sustain positive links to the agents in $N_i^+(\tilde{\mathbf{g}}^*) \cap N_k^-(\tilde{\mathbf{g}}^*)$, then a profitable bilateral deviation exists nonetheless. Take an agent $l \in N_i^+(\tilde{\mathbf{g}}^*) \cap N_k^-(\tilde{\mathbf{g}}^*)$ and compare $i's$ marginal payoffs from the positive link to l in $\tilde{\mathbf{g}}^*$ with $k's$ marginal payoffs when creating a positive link with l in a bilateral deviation. Other than the payoffs that i and k forgo by linking positively to l, there are two sets of agents to consider. Marginal payoffs from agents in $N_i^-(\tilde{\mathbf{g}}^*) \cap N_k^-(\mathbf{g}')$ and from agents in $N_i^+(\tilde{\mathbf{g}}^*) \cap N_k^-(\mathbf{g}')$. First, agents $j \in N_i^-(\tilde{\mathbf{g}}^*) \cap N_k^-(\mathbf{g}')$. Marginal payoffs from agents j are larger for k than for i, which follows from $\eta_j(\tilde{\mathbf{g}}^*) = \eta_j(\mathbf{g}')$, $\eta_i(\tilde{\mathbf{g}}^*) > \eta_k(\mathbf{g}')$ and the assumption that $\triangle_{\eta_i} f(\eta_i, \eta_j)$ is decreasing in η_i for given η_j. Second, marginal payoffs from agents such that $j \in N_i^+(\tilde{\mathbf{g}}^*) \cap N_k^-(\mathbf{g}')$. These are positive for k, while they are *zero* for i. Finally, payoffs that i and k forgo by linking positively to l. For agent i these are given by $f(\eta_i(\tilde{\mathbf{g}}^*) - 1, \eta_l(\tilde{\mathbf{g}}^*) - 1)$ while for k they are $f(\eta_k(\tilde{\mathbf{g}}^*), \eta_l(\tilde{\mathbf{g}}^*))$. Payoffs forgone are higher for i than for k, as $\eta_i(\tilde{\mathbf{g}}^*) > \eta_k(\mathbf{g}')$ and $\eta_l(\tilde{\mathbf{g}}^*) - 1 < \eta_l(\mathbf{g}')$. Therefore, if the positive link to l is profitable for agent i, then a bilateral deviation of k and l is profitable for k. Proposed deviation is profitable for l, as $\eta_l(\tilde{\mathbf{g}}^*) < \eta_k(\tilde{\mathbf{g}}^*)$. If, on the other hand, the positive link to l is not profitable for agent i, then i can profitably deviate by extending a negative link instead.

Step 5: *In any BE* $\tilde{\mathbf{g}}^*$, $\tilde{\bar{g}}^*_{i,k} = -1 \; \forall i \in P^m(\tilde{\mathbf{g}}^*)$ *and* $\forall k \notin P^m(\tilde{\mathbf{g}}^*)$.

From Step 4 we know that $\tilde{\bar{g}}^*_{i,k} = -1 \; \forall i \in P^m(\tilde{\mathbf{g}}^*)$ and $\forall k \in P^{m-1}(\tilde{\mathbf{g}}^*)$. Assume to the contrary that there exists a positive link $\tilde{\bar{g}}^*_{i,l} = 1$ to some agent $l \in P^{m-x}(\tilde{\mathbf{g}}^*)$ with $x \geq 2$. Then agents i and k can deviate with the following deviation strategies: $\tilde{g}^*_i + g^+_{i,k} + g^-_{i,l}$ and $\tilde{g}^*_k + g^+_{k,i}$. This is profitable for i, as $\eta_l(\mathbf{g}') < \eta_k(\tilde{\mathbf{g}}^*)$ and it is profitable for k, as $\eta_k(\tilde{\mathbf{g}}^*) < \eta_i(\tilde{\mathbf{g}}^*)$.

Define the super set $\tilde{P^r}(\tilde{g}^*) = P^m(\tilde{g}^*) \cup P^{m-1}(\tilde{g}^*) \cup \ldots \cup P^{m-r-1}(\tilde{g}^*) \cup P^{m-r}(\tilde{g}^*)$. Note that $\tilde{P}^0(\tilde{g}^*) = P^m(\tilde{g}^*)$.

Inductive Step: *In any BE* $\tilde{g}^*$, *if* $\tilde{\bar{g}}^*_{i,j} = 1 \; \forall i,j \in P^{m-x}(\tilde{\mathbf{g}}^*)$ *and* $\tilde{\bar{g}}^*_{i,k} = -1 \; \forall i \in P^{m-x}(\tilde{\mathbf{g}}^*)$ *and* $\forall k \notin P^{m-x}(\tilde{\mathbf{g}}^*)$ *holds* $\forall x \in \mathbb{N} : 0 \leq x \leq r$, *then* $\tilde{\bar{g}}^*_{i,j} = 1 \; \forall i,j \in P^{m-(r+1)}(\tilde{\mathbf{g}}^*)$ *and* $\tilde{\bar{g}}^*_{i,k} = -1 \; \forall i \in P^{m-(r+1)}(\tilde{\mathbf{g}}^*)$ *and* $\forall k \notin P^{m-(r+1)}(\tilde{\mathbf{g}}^*)$.

In the Step 5 we showed that $\tilde{\bar{g}}^*_{i,j} = 1 \; \forall i,j \in P^m(\tilde{\mathbf{g}}^*) \wedge \tilde{\bar{g}}^*_{i,k} = -1 \; \forall i \in P^m(\tilde{\mathbf{g}}^*)$ and $\forall k \notin P^m(\tilde{\mathbf{g}}^*)$. Assume the statement holds for all sets $P^{m-x}(\tilde{\mathbf{g}}^*) \; \forall x \in \mathbb{N} : 0 \leq x \leq r$. Note that for a link $\tilde{\bar{g}}^*_{i,k} = -1$ with $\eta_i(\tilde{\mathbf{g}}^*) > \eta_k(\tilde{\mathbf{g}}^*)$, $\tilde{g}^*_{i,k} = -1$ holds. Otherwise, for the undirected link $\tilde{\bar{g}}^*_{i,k} = -1$ to be negative, it must be that k extends the directed negative link $\tilde{g}^*_{i,k} = -1$. But then k can profitably deviate unilaterally by extending $\tilde{g}^*_{i,k} \neq -1$. We therefore know that $\tilde{g}^*_{i,k} = -1 \; \forall i \in \tilde{P^r}(\tilde{\mathbf{g}}^*)$ and $\forall k \notin \tilde{P^r}(\tilde{\mathbf{g}}^*)$. (This, of course, still allows for $\tilde{g}^*_{k,i} = -1$). We can now repeat steps 1 through 5 from the base case, relabeling $P^m(\tilde{\mathbf{g}}^*)$ with $P^{m-(r+1)}(\tilde{\mathbf{g}}^*)$ and $P^{m-1}(\tilde{\mathbf{g}}^*)$ with $P^{m-(r+2)}(\tilde{\mathbf{g}}^*)$ to establish the above result. Note that in Step 4, when comparing marginal payoffs of agents $i \in P^{m-(r+1)}(\tilde{\mathbf{g}}^*)$ and $k \in P^{m-(r+2)}(\tilde{\mathbf{g}}^*)$ we have to take into account marginal payoffs from agents in $P^{m-x}(\tilde{\mathbf{g}}^*) \forall x \in \mathbb{N} : 0 \leq x \leq r$. Marginal payoffs are again higher for k than for i, which follows from $\eta_j(\tilde{\mathbf{g}}^*) = \eta_j(\mathbf{g}') \; \forall j \in \tilde{P^r}(\tilde{\mathbf{g}}^*)$, $\eta_i(\tilde{\mathbf{g}}^*) > \eta_k(\mathbf{g}')$ and the assumption that $\triangle_{\eta_i} f(\eta_i, \eta_j)$ is decreasing in η_i for given η_j. *Q.E.D.*

Lemma 2 shows that there is no bilateral equilibrium such that either all undirected links are positive or all undirected links are negative. This is easy to see. If all links are positive, then two agents have an incentive to extend negative links to a third agent in a coordinated deviation. If all links are negative, then two agents can create a positive undirected link instead, thereby extracting payoffs from all remaining agents.

Lemma 2: *There does not exist a BE such that either $\bar{g}_{i,j} = 1$ or $\bar{g}_{i,j} = -1$ $\forall i, j \in N$.*

Proof. First, $\bar{g}_{i,j} = 1 \ \forall i, j \in N$. This is not a *BE* as any pair of agents i, j can profitably deviate by both extending a negative link to a third agent k. Second, $\bar{g}_{i,j} = -1 \ \forall i, j \in N$. Agents i and j can profitably deviate with the following strategies $\tilde{g}_i^* + g_{i,j}^+$ and $\tilde{g}_j^* + g_{j,i}^+$. This strictly increases payoffs, as i and j have now more friends than all remaining agents, to which i and j sustain negative links. $Q.E.D.$

A bilateral equilibrium may not always exist. The reason for this is that now pairs of agents may find it profitable to simultaneously extend negative links to a third agent within the same clique. Below we provide a simple example where a bilateral equilibrium exists.

Example 1: There are $n = 5$ players. Assume f to be the normalized contest success function in ratio form with $\phi = 1$. An agent $i's$ payoffs from a coercive link with agent j, denoted by $p_{i,j}$, is then given by $p_{i,j} = \frac{\eta_i+1}{(\eta_i+1)+(\eta_j+1)} - \frac{1}{2}$. The following configuration is a bilateral equilibrium. A clique of positively connected agents 1, 2, and 3, who extend negative links to agents in a second clique, consisting of players 4 and 5. There are only two relevant candidates for a profitable deviation. First, a deviation where two agents in the larger clique, say agents 1 and 2, extend negative links to a third agent within the clique, say agent 3. Second, a deviation where an agent in the larger clique, say agent 1, creates a positive link with an agent in the smaller clique, say agent 4. Payoffs for agents in the larger clique prior to a deviation are given by $2 \cdot \frac{3}{3+2} = \frac{6}{5}$. Agent 1 and $2's$ payoffs after the first deviation are given by $\frac{2}{2+1} - \frac{1}{2} + 2 \cdot (\frac{2}{2+2} - \frac{1}{2}) = \frac{1}{5}$. This is less than the payoffs prior to the deviation of $\frac{6}{5}$ and proposed deviation is therefore not profitable. Agent $1's$ payoffs after the second deviation are given by $\frac{4}{4+1} = \frac{4}{5}$. Again this is less than payoffs prior to the deviation of $\frac{6}{5}$ and proposed deviation is therefore not profitable.