Equilibrium configurations in the heterogeneous model of signed network formation

Sunjin Kim

March 13, 2019

Abstract

The paper is an extension of Hiller (2017). An exhaustive classification obtains for the particular case where there are two weak players and two strong players: there are three possible network formations, and all of these involve segregation or discrimination. When there is heterogeneity between the players, a utopia network formation where all players are friends with each other cannot be a Nash equilibrium. Among the equilibrium formations, a particular case is named ‘positive assortative matching’, where players of the same type coalesce. For an arbitrary number of players and two types, I study the conditions that are either necessary or sufficient for positive assortative matching to arise in equilibrium. Moreover, when players are all homogeneous, I characterize the conditions under which segregation into two uneven groups can be sustained in equilibrium. The results give some insight into discrimination and collusion.

JEL Codes: D74, D85
Hiller (2017) proposed a signed network formation model incorporating both friendship and conflict relationships. In his model, players form friendships and extract payoff from weaker enemies. He established some essential properties in the signed network formation model. Some of the properties apply to the general case with heterogeneous players, but most of the properties discussed the particular case with homogeneous players. His analysis gave excellent sights in situations where the players are similar.

Hiller’s key findings are as follows. First, the network formation is weakly structurally balanced in any Nash equilibria (Nash equilibrium). A graph is called weakly structurally balanced when the graph consists of some cliques. Players have positive relationships with the players in the same clique but have negative relationships with the players in the different cliques. As a result, a friend of my friend is always my friend in any weakly structurally balanced graph. Second, in the case of homogeneous players, there exist two specific equilibrium configurations. In the first one all players are friends, and in the second configuration, one player is the enemy of everyone else but the other players are all friends. Third, he characterized a condition under which there are only two cliques in any Nash equilibrium. In his model, each player’s payoff consists of two parts, an “extraction”, and the costs of negative links. The extraction is positive if the player’s strength is larger than his enemy’s power; it is negative otherwise. If second derivatives of the extraction due to a change in the player’s or the enemy’s power is sufficiently close to zero, there exist only two cliques in the Nash equilibrium.

When the players are heterogeneous, however, there are unaccountably many kinds of Nash equilibrium and it is hard to sort them. Due to this high number of cases, in this research, I assume there are only two types of players in the network formation game. Though this particular case cannot include all of the heterogeneous cases, the specific case can explain many situations.

In this paper, I check which network formations can be Nash equilibria when there are two strong players and two weak players. All Nash equilibrium in the model are specified with conditions. All of N.E show discrimination phenomena. A utopia formation, where all players are friends to each other, can not be sustained in a Nash equilibrium. In section 3.2 I focus on conditions for a positive assortative matching formation to be a Nash equilibrium, where players of the same type join together to form a group, with arbitrary numbers of players. The result shows that if strong group’s aggregate power is larger than the weak group’s aggregate power, then the positive assortative matching strategy can be a Nash equilibrium. However, a small difference between the power is not enough to make the formation a Nash equilibrium.

The paper unfolds as follows. Section 1 describes the related literature. Section 2 includes
1 Literature Review

Conflict is one of the important topics in economics because people usually have a conflicting interest in their economic activities. If their interest is consistent with the others’, they can cooperate with each other, and it is not a big deal. Hence the remaining issue is how to deal with the conflicting interest. Some researchers studied how players behave on any exogenously given network by using the game-theoretic approach. Bramoullé and Kranton (2007) and Bramoullé, Kranton and D’amours (2014) analyzed an outcome on network when there is substitutability between linked players’ production. Goyal, Heidari and Kearns (2014) and Kovenock and Roberson (2018) analyzed what optimal strategies to attack and to defend network is. König et al. (2017) derived a Nash equilibrium fighting effort when there are alliance and enmity network given.

On the other hand, there is another attempt to understand the conflict phenomena concerning an endogenous network formation process. Hiller (2012) and Hiller (2017) studied a signed network formation model. In this model, players could choose friendship and enmities. Their choices formed the network which determined their payoff. Jackson and Nei (2015) showed that trade and high war cost decreased conflict in the network formation game. Grandjean, Tellone and Vergote (2017) modeled a sequential game where players formed a network in the first stage and had a contest in the second stage.

This paper is an extension of Hiller (2012) and Hiller (2017) using the signed network formation model. He suggested an essential property that all Nash equilibrium network formations were weakly structurally balanced, but analyzed the model focusing on cases where the players are homogeneous. In this study, I focus on heterogeneity among the players. When there are heterogeneous players in the model, there can be many other network formations to be Nash equilibrium. Among many kinds of Nash equilibrium, there is a particular network formation named a positive assortative matching. In the formation, the players form the same group only with the same type players. The positive assortative matching gives intuition for social phenomena such as discrimination, segregation, and bullying. The results can give an intuition when discrimination happens.
2 Model

There are two types of players, \( t \in \{ s, w \} \). \( s \) means strong, and \( w \) means weak. \( n \) players \( S_1, S_2, ... S_n \) are type \( s \) and \( m \) players \( W_1, W_2, ..., W_m \) are type \( w \). \( n \) and \( m \) are larger than or equal to 2. The player \( S_1 \)'s intrinsic power is \( \lambda_s \), and the player \( W_1 \)'s intrinsic power is \( \lambda_w \). \( \lambda_s \) is supposed to be bigger than \( \lambda_w \). Define \( N \) a set of all players, \( N_s \) a set of the type \( s \) players, and \( N_w \) a set of the type \( w \) players.

The players’ strategy profile is a directed network \( \mathbf{g} = (g_{S_1}, g_{S_2}, ..., g_{S_n}, g_{W_1}, g_{W_2}, ..., g_{W_m}) \), where \( g_i \) is the player \( i \)'s strategy, \( g_i \) is a row vector consisting of \( n + m - 1 \) elements. Each element represents the relationship from the player \( i \) to the other players. For example, the player \( S_1 \)'s strategy \( g_{S_1} = (g_{S_1,s_2}, ..., g_{S_1,S_n}, g_{S_1,W_1}, g_{S_1,W_2}, ..., g_{S_1,W_m}) \). \( g_{i,j} \) is either -1 or 1 for all \( j \in N \setminus \{ i \} \). 1 is a positive link and -1 is a negative link. The directed network \( \mathbf{g} \) forms an undirected network \( \bar{\mathbf{g}} \). If \( g_{i,j} = g_{j,i} = 1 \), then \( \bar{g}_{i,j} = 1 \). Otherwise, that is if at least one of \( g_{i,j} \) and \( g_{j,i} \) is -1, then \( \bar{g}_{i,j} = -1 \). Changes of directed link \( g_{i,j} \) is denoted as follows given a network \( \mathbf{g} \). If \( g_{i,j} = -1 \) in \( \mathbf{g} \), then \( \mathbf{g} + \mathbf{g}_{i,j}^+ \) changes the directed link \( g_{i,j} = -1 \) to \( g_{i,j} = 1 \). If \( g_{i,j} = 1 \) in \( \mathbf{g} \), then \( \mathbf{g} + \mathbf{g}_{i,j}^- \) changes the directed link \( g_{i,j} = 1 \) to \( g_{i,j} = -1 \). However, if \( g_{i,j} = 1 \) in \( \mathbf{g} \), then \( g + g_{i,j}^- = g \). Similarly, if \( g_{i,j} = -1 \) in \( \mathbf{g} \), then \( g + g_{i,j}^+ = g \). The summation sign \( \Sigma \) can be used to denote multiple link changes.

The utility of player \( i \) under a strategy profile \( \mathbf{g} \) is given by

\[
   u_i(\mathbf{g} \cdot \mathbf{g}_{-i}) = \sum_{j \in N_i^+(\mathbf{g})} f(n_i(\mathbf{g} \cdot \mathbf{g}_{-i}), n_j(\mathbf{g} \cdot \mathbf{g}_{-i})) - e_i(\mathbf{g} \cdot \mathbf{g}_{-i})\varepsilon - c_i(\mathbf{g} \cdot \mathbf{g}_{-i})\kappa. \tag{1}
\]

\( \mathbf{g}_{-i} \) is a set of all the players’ strategies except for the player \( i \)'s strategy. \( N_i^+(\mathbf{g}) = \{ j \in N | \bar{g}_{ij} = 1 \} \), meaning a set of players to which the player \( i \) reciprocates a positive link. \( N_i^-(\mathbf{g}) = \{ j \in N | \bar{g}_{ij} = -1 \} \), meaning a set of players such that \( i \) extends and/or receives a negative link. \( c_i(\mathbf{g}) = |N_i^-(\mathbf{g})| \), that is the cardinality of \( N_i^-(\mathbf{g}) \). \( N_i^{e-} = \{ j \in N | g_{ij} = -1 \} \), a set of players to which the player \( i \) extends a negative link. \( c_i(\mathbf{g}) \) is a cardinality of \( N_i^{e-} \), \( |N_i^{e-}(\mathbf{g})| \). Strength or power of a player is determined endogenously and consists of a player \( i \)'s intrinsic strength, \( \lambda_i > 0 \), and the intrinsic strength of all players that reciprocate the player \( i \)'s positive link. The strength of the player \( i \) in network \( \mathbf{g} \) is given by

\[
   n_i(\mathbf{g} \cdot \mathbf{g}_{-i}) = \lambda_i + \sum_{j \in N_i^+(\mathbf{g})} \lambda_j. \tag{2}
\]

Extending a negative link initiates an antagonistic relationship and may be thought of as picking a fight, where picking a fight incurs a cost of \( \varepsilon > 0 \). Once at least one negative link is extended, then both players are assumed to engage in conflict, and a cost of conflict
is $\kappa \geq 0$. Gross payoffs from a negative link are determined by a function $f(n_i(g), n_j(g))$, which depends on the players’ respective strengths $n_i(g)$ and $n_j(g)$. Direct payoffs from a reciprocated positive link are zero. The value a player extracts through an antagonistic link is assumed to be another player’s loss and, therefore, $f(n_i, n_j) + f(n_j, n_i) = 0 \ \forall n_i, n_j$. If a pair of players of equal strength is negatively connected, then no payoff extraction takes place: $f(n_i, n_j) = 0 \ \forall n_i, n_j : n_i = n_j$. $f(n_i, n_j)$ is strictly increasing in $n_i$ and decreasing in $n_j$. In this paper, when a closed form solution is required for intuitions, I will use a simple normalized contest success function as follows

$$f(n_i, n_j) = \frac{n_i}{n_i + n_j} - \frac{1}{2}. $$

The normalized contest success function satisfies the properties mentioned above. By using the specific functional form, I will sacrifice generality but gain some closed form solution allowing comparative static analysis. The total linking cost of a player $i$ in network $g$ is determined by $c_i(g)$ extended negative links, in which each incurs a cost of $\varepsilon > 0$, and $c_i(g)$ undirected negative links, in which each incurs a cost of $\kappa \geq 0$. Hiller (2017) assumed that $\varepsilon + \kappa < f(\sum_{j \in N \setminus \{k\}} \lambda_j, \lambda_k)$ where $\lambda_k \leq \lambda_j \ \forall j \in N$. In similar, in this report, $\varepsilon$ and $\kappa$ are supposed to be small enough to be ignored compared to any extracted value, so $\varepsilon + \kappa < \min\{f(n_i(g), n_j(g))\}$ for all $g$. However, I also discuss large size of $\varepsilon$ and $\kappa$ in some sections. The small cost assumption allows us to derive closed-form conditions for Nash equilibrium network formation ignoring the cost factor. At the same time, the costs are still positive. Therefore the players prefer the positive relationship to the negative relationship if the utilities except for the costs are the same. A deviation strategy profile including $g'_i$, that is a network after a proposed deviation, is denoted by $g' = (g'_i, g_{-i})$.

3 Analysis

3.1 Equilibrium Characterization with four players

Hiller (2017) gave an important proposition characterizing any Nash Equilibria(Nash equilibrium) in the model.

Proposition 1 (Hiller (2017)) In any NE, if $n_i(g) = n_j(g)$, then $\bar{g}_{ij} = 1$, and if $n_i(g) \neq n_j(g)$, then $\bar{g}_{ij} = -1$.

The proposition implies that if some players are friends with each other in a Nash equilibrium, the players’ powers are the same. Hence, if they are on the same team, their powers
are the same. Also, all of the Nash equilibrium formations are weakly structurally balanced, so all players form some cliques in any Nash equilibrium. For example, my friend’s friend is always my friend, and my enemy’s friend is always my enemy. Proposition is essential to exclude not weakly structurally balanced formations from Nash equilibrium, so it decreases cases that we should check. However, it is the fact that there are still too many Nash equilibrium when we deal with the general number of players. Let’s consider models with 4 players to see which Nash equilibrium network formations are possible and when they are Nash equilibrium.

In the first model, $n = 2$ and $m = 2$ so there are two players respectively for each strong and weak type. The total number of players is four. Considering Proposition there are nine kinds of possible formations as presented in Figure. Among them, three kinds of formations can be Nash equilibrium. The three possible formations are marked with red lines.

First, let’s call formation (a) a positive assortative matching (PAM). It is easy to imagine a situation where similar people assemble. In the positive assortative matching, the strong type players $S_1$ and $S_2$ extend positive links to each other and make negative links to the weak type players $W_1$ and $W_2$. $W_1$ and $W_2$ extend positive links to all of the players in the formation.

**Result 1** When there are two strong type players and two weak type players, the positive assortative matching (a) is a Nash equilibrium if and only if

\[ f(2\lambda_s + \lambda_w, 2\lambda_w) - f(2\lambda_s, 2\lambda_w) \leq f(2\lambda_s, 2\lambda_w) - \kappa - \varepsilon, \]
\[ f(2\lambda_s, 2\lambda_w) \geq \kappa + \varepsilon, \]
\[ 2f(2\lambda_s, 2\lambda_w) \geq f(\lambda_s + 2\lambda_w, \lambda_s) + \kappa + \varepsilon. \]

In formation (b), the strong type players $S_1$ and $S_2$ and one of the weak type players $W_1$ extend the positive links to each other and extend negative links to the other weak type player $W_2$. Let’s call this formation weak-type victim formation.

**Result 2** When there are two strong type players and two weak type players, weak-type victim formation (b) is a Nash equilibrium if and only if

\[ f(2\lambda_s + \lambda_w, \lambda_w) - f(2\lambda_s, \lambda_w) \geq f(2\lambda_s, \lambda_s + \lambda_w) - \kappa - \varepsilon \]
\[ f(2\lambda_s + \lambda_w, \lambda_w) \geq \kappa + \varepsilon. \]

In formation (c), one strong type player $S_1$ and weak type players $W_1$ and $W_2$ extend positive links to each other and extend negative links to the other strong type player $S_2$. $S_2$ extends positive links to all other players. This can be called strong-type victim formation.

**Result 3** When there are two strong type players and two weak type players, strong type victim formation (c) is a Nash equilibrium if and only if

\[ f(\lambda_s + 2\lambda_w, \lambda_s) - f(\lambda_s + \lambda_w, \lambda_s) \geq f(\lambda_s + \lambda_w, 2\lambda_w) - \kappa - \varepsilon, \]
\[ f(\lambda_s + 2\lambda_w, \lambda_s) \geq \text{Max}(\kappa + \varepsilon, f(2\lambda_s + \lambda_w, 2\lambda_w), 2f(2\lambda_s, 2\lambda_w) - \kappa - \varepsilon). \]
The rest of the formations are not possible to be Nash equilibrium.

**Result 4** When there are two strong type players and two weak type players, the other network formations from (d) to (i) in Figure 1 cannot be a Nash equilibrium.

The proof for the condition of Nash equilibrium are in the appendix.

Result 4 shows that the dominant form of Nash equilibrium network formation is a bipartite network formation where there are two groups. The conditions in Result 1 to 3 are paraphrases that the suggested strategy should produce larger payoff than the other deviations. Even though they look trivial, we can derive the following results combining the conditions.
Result 5 When there are two strong type players and two weak type players, if \(2f(2\lambda_s, 2\lambda_w) \geq f(\lambda_s + 2\lambda_w, \lambda_s) + \kappa + \varepsilon\), PAM (a) can be a Nash equilibrium, but the formation (c) cannot be. If \(2f(2\lambda_s, 2\lambda_w) = f(\lambda_s + 2\lambda_w, \lambda_s) + \kappa + \varepsilon\), both (a) and (c) can be Nash equilibrium. If \(2f(2\lambda_s, 2\lambda_w) \leq f(\lambda_s + 2\lambda_w, \lambda_s) + \kappa + \varepsilon\), PAM (a) cannot be but the formation (c) can be a Nash equilibrium.

Result 6 If the difference between \(\lambda_s\) and \(\lambda_w\) is large, strong-type victim formation (a) cannot be a Nash equilibrium.

Proof. If \(\lambda_s \geq 2\lambda_w\), then \(2f(2\lambda_s, 2\lambda_w) > f(2\lambda_s, 2\lambda_w) \geq f(\lambda_s + 2\lambda_w, \lambda_s)\). If \(\varepsilon + \kappa\) is so small that \(\varepsilon + \kappa < f(2\lambda_s, 2\lambda_w)\), then \(2f(2\lambda_s, 2\lambda_w) \geq f(\lambda_s + 2\lambda_w, \lambda_s) + \kappa + \varepsilon\).

Result 6 shows that the size of the difference between \(\lambda_s\) and \(\lambda_w\) determines whether strong-type victim formations can be Nash equilibrium. However, the reverse condition that the difference is small enough cannot exclude the possibility for PAM to be a Nash equilibrium.

Until now, we have the necessary and sufficient conditions for the formations to be Nash equilibrium, but it is not enough to drive intuitions from the conditions. At this point, we can apply the contest success function for \(f(\cdot)\) to compare the size of the strength ratio ensuring the formations to be Nash equilibrium to get numerical descriptions of the network formations. Let’s define \(a = \frac{\lambda_s}{\lambda_w}\) and assume that \(\varepsilon\) and \(\kappa\) are small enough to be ignored. As a result of calculation, the formation (a), PAM, can be a Nash equilibrium when \(a \geq 1.5\). The formation (b) can be a Nash equilibrium when approximately \(a < 1.5\).\(^1\) The formation (c) can be a Nash equilibrium when \(a \leq 1.5\). The calculation is consistent with the above results. Figure 2 shows conditions for each network formation graphically.

\(^1\)I calculated the value with Mathematica.
Note that the criteria to be Nash equilibrium of (c) is less than that of (b). It means that bullying a weak player with the other weak player is more likely to happen than bullying a strong player with the other strong player. Hence, there is an interval of strength ratio allowing the formation (b) to be Nash equilibrium but not the formation (c). That is, the size of $a$ is between 1.5 and 1.54, the phenomenon we typically regard unjustifiable can be a Nash equilibrium, but bullying a strong player cannot be a Nash equilibrium.

The calculation shows that when $a$ is high, segregation happens, and when $a$ is low, bullying occurs. Then, is a utopia where everyone is friendly with everyone impossible in the model? According to Hiller (2012), when there is heterogeneity, the utopia network formation such as (e) in Figure 1 is always not a Nash equilibrium.

**Proposition 2 (Hiller (2012))**  
If there exists a pair of agents $i$ and $j$ such that $\lambda_i \neq \lambda_j$, then there does not exist a Nash equilibrium $g^*$, such that everyone is fiends with everyone.

The utopia impossible even though the players are only slightly different from each other if $\varepsilon + \kappa < f(\sum_{j \in N \setminus \{k\}} \lambda_j, \lambda_k)$ where $\lambda_k \leq \lambda_j \forall j \in N$. On the other hand, if $\varepsilon + \kappa$ is large, the utopia formation is possible even with the heterogeneity. In particular, if $\varepsilon + \kappa > f(\sum_{j \in N \setminus \{k\}} \lambda_j, \lambda_k)$ where $\lambda_k \leq \lambda_j \forall j \in N$, the utopia formation is the only possible Nash equilibrium outcome regardless of heterogeneity.

From now, as Hiller (2017) did, I maintain the assumption of small cost to allow negative links in any Nash equilibrium. Then the utopia network formation is not possible to be a Nash equilibrium in any heterogeneous cases. Considering the rest of possible Nash equilibrium, each segregated network formation has its pro and cons. The formation (a) shows extreme segregation. The players only care of the same type player and have negative relationships with the different kind of players The formation looks like real-world phenomena such as discrimination between genders and races. On the other hand, the formations (b) and (c) looks like bullying minority in the real world. The different kinds of the players form the same team to maximize their payoff. It is also a matter in human society because the majority bullies minority. But it is hard to interpret it as discrimination because the bullying phenomena happen irrespective of a bullied player’s type. As the formation (c) shows, the strong type player can be also bullied when his intrinsic power is not strong enough relative to the weak type players. The segregation in (b) and (c) are an ex-post event, and any players can be victims. However, the segregation in (a) has an ex-ante property, because the weak type is inherently given and they will be definitely bullied in the Nash equilibrium.

Now let’s consider the second four players model where one type of players are three and the other type of player is one. In the model, PAM can be a Nash equilibrium.
**Result 7** When there are 3 strong type players and 1 weak type players, PAM is always a Nash equilibrium with any $f(\cdot)$ and small enough $\kappa$ and $\varepsilon$.

**Result 8** When there are 1 strong type players and 3 weak type players, PAM is a Nash equilibrium with any $f(\cdot)$, if

i) $\lambda_s > 3\lambda_w$, and $3f(\lambda_s, 3\lambda_w) - 3\kappa - 3\varepsilon \geq \text{Max}[2f(\lambda_s + \lambda_w, 3\lambda_w) - 2\kappa - 2\varepsilon, f(\lambda_s + 2\lambda_w, 3\lambda_w) - \kappa - \varepsilon]$

ii) $\lambda_s < 3\lambda_w$ and $f(3\lambda_w, \lambda_s) - \kappa - \varepsilon \geq \text{Max}[f(2\lambda_w + \lambda_s, 2\lambda_w) - \kappa - \varepsilon, 2f(\lambda_w + \lambda_s, 2\lambda_w) - 2\kappa - 2\varepsilon]$.

Result 7 can be generalized to $n$ strong type players case.

**Result 9** When there are $n$ strong type players and one weak type players, PAM is a Nash equilibrium with any $f(\cdot)$ and small enough $\kappa$ and $\varepsilon$.

The proof for Result 9 is in Appendix. Result 8 is just a listing of conditions which any Nash equilibrium strategy should satisfy. If the normalized contest function is applied, then $a \geq \frac{1}{2}(-1 + \sqrt{97}) \approx 4.42443$ or $a \leq -1 + \sqrt{7} \approx 1.64575$ makes PAM a Nash equilibrium. In two strong and two weak type players cases, the strong type players always have stronger strength and extend negative links to the weak type players in Nash equilibrium. On the other hand, in one strong and three weak type players case, when $\lambda_s < 3\lambda_w$, the weak type players also can extend negative links in the Nash equilibrium.

While Result 9 generalizes Result 7, it is hard to generalize 8. This is because we should find the biggest payoff by deviating from PAM to check whether PAM is a Nash equilibrium, but the biggest payoff is changed depending on the functional form of $f(\cdot)$ and the number of players. If we use the normalized contest function, given a fixed number of players $n$, it is possible to find the biggest deviation payoff, but it is not easy to derive a general necessary and sufficient condition for PAM to be a Nash equilibrium Instead, it is possible to derive a sufficient condition for PAM to be a Nash equilibrium.

**Result 10** When there are $n$ weak type players and one strong type players and $\lambda_s > n\lambda_w$, there exists $\tilde{a}$ such that PAM is a Nash equilibrium with the normalized contest function and small enough $\kappa$ and $\varepsilon$ if and only if $a \geq \tilde{a}(n, 1, \varepsilon, \kappa)$.

When $\lambda_s < n\lambda_w$, there exists $a^s$ such that PAM is a Nash equilibrium with the normalized contest function and small enough if $a \leq a^s(n, 1, \varepsilon, \kappa)$.

It is in the appendix how to derive $a^1$ and $a^2$. 
3.2 Generalized conditions when the same type players form a clique.

As we saw in the model with the four players, multiple kinds of Nash equilibrium network formation is possible. Furthermore, if we consider a general case with arbitrary numbers of strong players and weak players, there should also be too many kinds of Nash equilibrium, so it will be hard to check the conditions for all Nash equilibrium. Therefore, it is inevitable to focus on some significant network formations.

In this section, I derive conditions for PAM to be a Nash equilibrium when there are an arbitrary number of players. Thus $n$ and $m$ can be any numbers. We can derive closed-form conditions for PAM to be Nash equilibrium by using the normalized contest function and the assumption that $\kappa$ and $\varepsilon$ are small enough. PAM is worth to check because first, besides the utopia formation and PAM, there are many network formations where the different types of players are mixed in the groups. It is difficult to subcategorize the formations except the utopia formation and PAM in detail. Second, as Proposition 2 shows, the utopia network formation where everyone is friends to each other is not a Nash equilibrium. Because of the difficulty in categorization and significance of the formations, it is meaningful to study PAM.

Before the main discussion, let's define one more term. As Proposition 1 mentioned, in any Nash equilibrium formations, the players form some cliques which are a partition of the set of player $N$. In PAM where the same type of players form a clique, the same type of players shares equal network power. When there are only two cliques in the formations, we can call a set, in which the players’ network power is higher than them in the other cliques, a strong group(team). The other clique is called weak group(team). Note that regardless of the group players’ type, the strong group and the weak group are determined by the players’ network power. Sometimes, the weak type players can consist of the strong group in a Nash equilibrium if their number is large enough relative to the strong type players. Then the weak type players also can extend negative links to the strong type players and extract some payoff in PAM formation. Therefore, who will be the strong(weak) team player is depending on the parameters $\lambda_s, \lambda_w, n$ and $m$. Thus, I should distinguish players who have stronger(weaker) intrinsic power $\lambda_s(\lambda_w)$ from players who have stronger(weaker) power $n_s(n_w)$. Players with $n_s(n_w)$ can be called strong(weak) team players. Thus, the size of $n_s$ is supposed to be larger than $n_w$. If $n_s = n_w$ in PAM, that is $n\lambda_s = m\lambda_w$, PAM is not possible to be a Nash equilibrium. Therefore, the case does not need to be considered, hence $n_s$ is always larger than $n_w$.

Now, let’s formally define PAM strategy. In PAM, there are only positive (negative) undirected relationships within (between) groups. That is, the strong team players extend
the negative links to the weak team players. Except for those links, the other directed links are positive. In particular, all strong group players extend negative links to all weak group players, and the rest of the directed links are all positive. As mentioned before, the strong group players are not always the strong type players. Therefore, let’s classify the cases.

First, the strong (weak) type players can form the strong (weak) group. The condition is $n\lambda_s > m\lambda_w$. The network formation strategy is formally, $\forall i \in N_s$, $g_{i,k} = 1 \ \forall k \in N_s \setminus \{i\}$, $g_{i,j} = -1 \ \forall j \in N_w$, and $\forall j \in N_w$, $g_{j,i} = g_{j,l} = 1 \ \forall i \in N_s$, $\forall l \in N_w \setminus \{j\}$. Conditions for the suggested strategy profile $g$ to be a Nash equilibrium are as follows.

**Result 11** Suppose $f$ is the normalized contest function and $\varepsilon$ and $\kappa$ are small enough. When $n\lambda_s > m\lambda_w$, type $s(w)$ players become the strong(weak) team player. There exists $\bar{a}(n,m,\varepsilon,\kappa)$ such that PAM is a Nash equilibrium if and only if $a \geq \bar{a}$.

When the strong type players become the strong group players, there is the condition for the Nash equilibrium when the number of players is many enough. $a \geq \bar{a}$ is the condition not to make one weak type new friend. If the condition is satisfied, all of the other possible deviations are not profitable. It means that, if the weak type players are weak enough relative to the strong type players, then there is no incentive for the strong type players from PAM to deviate by embracing some of the weak players. However, as $a$ decreases and the weak type players get stronger, the weak type players become so attractive to be new friends. The smaller $a$ is, the more friendships it is attractive for the strong type players to make. The other strategy with any positive links is always less attractive than the deviation strategy making only negative links.

Second, there is a special case when the players’ network power in two groups are the same. It happens when $n\lambda_s = m\lambda_w$. As mentioned above, the suggested bipolar formation always cannot be a Nash equilibrium. The proof is also in Appendix.

Third, there is the last case where the weak type players form the strong team, and they extend negative links to the strong type players. It happens when $m\lambda_w > n\lambda_s$. The strategy profile is formally, $\forall i \in N_w$, $g_{i,k} = 1 \ \forall k \in N_w \setminus \{i\}$, $g_{i,j} = -1 \ \forall j \in N_s$, and $\forall j \in N_s$, $g_{j,i} = g_{j,l} = 1 \ \forall i \in N_w$, $\forall l \in N_s \setminus \{j\}$. $g$ is a Nash equilibrium when the following conditions are satisfied.

**Result 12** Suppose $f$ is the normalized contest function and $\varepsilon$ and $\kappa$ are small enough. When $n\lambda_s < m\lambda_w$, type $w(s)$ players become the strong(weak) team player. There exists $\underline{a}(n,m,\varepsilon,\kappa)$ and $\bar{a}(n,m,\varepsilon,\kappa)$ such that PAM is a Nash equilibrium if $a \leq \underline{a}$, and PAM is not a Nash equilibrium if $a \geq \bar{a}$.

The results can be transformed in respect of $n_s$ and $n_w$ instead of $\lambda_s$ and $\lambda_w$. 
The results have some implication. First, PAM is a Nash equilibrium when the ratio of network powers $\frac{n_s}{n_w}$ should be strictly larger than a certain value. If it is less than it, PAM is not a Nash equilibrium even though $n_s > n_w$. It means some engagement policies can be originated from rational behavior, not from the warm heart. Without the network factor, the result is different. To understand the difference, let’s consider a case when there is only one strong team player and one weak team player. When there are only two players, it is always the Nash equilibrium where the strong team player extends a negative link and the weak team player extends a positive link for all $\lambda_s$ and $\lambda_w$. However, as $n$ or $m$ is larger than 1, PAM can fail to be a Nash equilibrium even though $n_s > n_w$. Second, when the difference between network powers between the types get less, the strong type players have an incentive to deviate from PAM. When the strong type players are the strong team players, it occurs when the intrinsic powers $\lambda_s$ and $\lambda_w$ get similar, while when the weak type players are the strong team players, it happens when the difference between the intrinsic powers $\lambda_s$ and $\lambda_w$ get larger. Third, when the number of players is not many, the condition for PAM to be a Nash equilibrium is strengthened. For example, when $n\lambda_s < m\lambda_w$, $a^s = \min\left[\sqrt[3]{2n^2 - 4n + 1}m + m - 2n^2 + 2\sqrt{n^3 + n^3m + n^3 - n}\right]$. If $n$ is small, $a^s = -2n^2 + 2\sqrt{n^3 + n^3m + n^3 - n} < \frac{\sqrt[3]{2n^2 - 4n + 1}m + m}{2(3n^2 - n)}$. However, when $n$ is not small enough relative to $m$, $a^s = \frac{\sqrt[3]{2n^2 - 4n + 1}m + m}{2(3n^2 - n)}$. Therefore, $a$ should be smaller relative to the case where $n$ is large. It is because when the number of the weak team (but strong type) players are few, some strong team (but weak type) player can utilize their strength to maximize their payoff. Embracing the weak team player, the deviating player will also initiate other conflicts with the strong team players were friends before.

Before finishing the section, let’s consider an application of the generalized conditions. Instead of the heterogeneous player setting, we can degenerate it to the model with homogeneous players. In the degenerated model, all players have the same intrinsic power. The only variables in the model are the number of the strong group players $n_1$ and the number of the weak group players $n_2$.

As mentioned before, the cost from the negative link $\varepsilon + \kappa$ is positive but negligible. In the suggested formation, there are $n_1$ bullying players and $n_2$ bullied players and $n_1 + n_2 = n$. Arbitrarily, let’s divide the set of players $N$ to $N_1 = \{1, 2, ..., n_1\}$ and $N_2 = \{n_1 + 1, n_1 + 2, ..., n\}$, so $N_1 + N_2 = N$. The strategy profile of PAM $g$ is that $g_{ij} = 1, g_{ik} = -1$ for $\forall i \in N_1, \forall j \in N_1 \setminus \{i\}, \forall k \in N_2, g_{kl} = 1$ for $\forall l \in N$. As a result of the strategy, the players in $N_1$ have a stronger network power $n_1 = n_1$ than the other players in $N_2$ whose network power is $n_2$. To ensure a positive extraction for the players in $N_1$, $n_1 > n_2$. Therefore, there
does not exist a case where fewer players extract a larger number of players. Applying the similar way of proof, we can derive a condition for any network formations with two cliques to be a Nash equilibrium.

**Result 14** Suppose all players have the same \( \lambda \). With the generalized contest success function and small enough \( \varepsilon \) and \( \kappa \), there exists \( \bar{n}_1(n_2, \varepsilon, \kappa) \) such that any network formations with two cliques are Nash equilibrium if and only if \( n_1 \geq \bar{n}_1 \).

The proof is in Appendix.

Note that all structurally balanced network formation consists of two cliques. If there are more than two cliques in the network, it is not structurally balanced by definition. Therefore, it is the characterization of Nash equilibrium for any structurally balanced network formation when the players are homogeneous.

### 4 Discussions of Extensions and Applications

The paper analyzes some heterogeneous player cases in the signed network formation model. Hiller (2017) analyzed some equilibria in the homogeneous players’ case with the general form of extraction function \( f(\cdot) \). In this paper, I use the normalized contest function to \( f(\cdot) \) if the normalized contest function drives the closed form solutions. The closed form conditions are useful when we need the comparative static analysis. On the other hand, using this function trades off generality of the model describing more complicated situation.

The analysis of the heterogeneous players’ cases gives some intuitions to discrimination. The focus of this paper is to analyze segregation originated from type which means the ex-ante difference. According to the analysis, some parameters can make PAM be a Nash equilibrium. For example, when the strong players are more than the weak players and the difference between intrinsic powers are large enough, people can expect PAM. Then discrimination to all weak type players is also anticipated. Policymakers can actively consider some measures to alleviate the discrimination before the network formation because targets of the policy is clear. Except for PAM, people cannot anticipate who is bullied in any other Nash equilibrium formation, as shown in the four players model. In this case, the policymakers cannot help performing any policies after observing the persecution.

This research gives some intuitions to collusion, too. Let’s assume that players choose a deviation utility which gives the maximum utility. Then cooperation crashes down in two ways by the strong group players. The first is for the strong group players to admit new different type players to their strong group, and the second is for the strong group player to betray some of their group players and to join to the other group. The result shows that the
first way drives the necessary and sufficient condition for a Nash equilibrium in most of the cases. If the first way is not attractive, then the second way is also not a profitable choice for the strong group players. However, the second way sometimes gives the condition for a Nash equilibrium instead of the first way. It happens for two cases. First, when the betrayal effectively weakens their colleagues, and the number of players on the other side is small enough. Second, when the number of weak team players is small, but each weak team player is more competitive than each strong team player, the strong team players have an incentive to change their partners.

Another intuition from the heterogeneous players model is that the utopia network formation is not possible to be a Nash equilibrium. This phenomenon was mentioned in Hiller (2012) and I checked the result once again. In the homogeneous players model, the utopia network formation is not a strong Nash equilibrium, but a Nash equilibrium. But in the heterogeneous model, if there is a slight difference between the players’ intrinsic strength and the cost from conflict is significantly small, there is always an incentive to extract the weak type players. Thus the utopia strategy is not a Nash equilibrium. For the policymakers or the principals who regard the segregation negative, the result is not desirable. They will try to make the players homogeneous or increase the cost of conflict. The policymaker can also design a mechanism to remove the incentive for the players to extract others.

5 Conclusion

The paper has built on Hiller (2017) by studying different equilibrium configurations. The model with two strong players and two weak players is of particular interest: there are three possible network formations in equilibrium and all of these involve segregation and discrimination. When there is heterogeneity between the players, the utopia network formation cannot be a Nash equilibrium. I also examine the four player model with one weak (strong) player and three strong (weak) players. When the one player is weak-type and the three players are strong-type, the bullying formation by the three players is always a Nash equilibrium. When the one player is strong-type and the three players are weak-type, there are two bullying formations by the three weak type players and by the one strong player. These formations are Nash equilibria when the bullying group’s aggregate power is sufficiently larger than the victim’s aggregate power.

In the case of an arbitrary number of players and two types, I derive the generalized condition for existence of an equilibrium exhibiting positive assortative matching. In similar to the four-player model, the positive assortative matching is a Nash equilibrium if the strong group’s aggregate power is sufficiently larger than the weak group’s aggregate power. Finally
I characterize the condition under which homogeneous players form an equilibrium network configuration with two cliques.

References


Appendix A. Proof of Result 1 to 4: The two strong and two weak players model

Result 1: A positive assortative matching: The same type players are friends.

The condition ii) is a condition for $S_1$ and $S_2$ not to extend the positive links to both $W_1$ and $W_2$. If the condition is satisfied, there are only two deviations which can be more profitable than the positive assortative matching $g$. First, one of the strong type players can extend a positive link to one of the weak type players. Second, one of the strong type players can change his friend from the strong type player to the weak type players. That is, he extends the negative link to the other strong type player and extends the positive links to the weak type players. Lastly, the other deviations are always not profitable.

The condition $f(2\lambda_s + \lambda_w, 2\lambda_w) - f(2\lambda_s, 2\lambda_w) \leq f(2\lambda_s, \lambda_w) + f(2\lambda_s, \lambda_s + \lambda_w) - 2\kappa - 2\varepsilon$ makes the first deviation unattractive. The condition is another expression of $2f(2\lambda_s, 2\lambda_w) - 2\kappa - 2\varepsilon \geq f(2\lambda_s + \lambda_w, 2\lambda_w) - \kappa - \varepsilon$. It is exactly the condition for a strong type player does not have an incentive to choose the first deviation. For the same reason, $S_1$ and $S_2$ do not have an incentive to deviate by using the second strategy if and only if the condition $2f(2\lambda_s, 2\lambda_w) > f(\lambda_s + 2\lambda_w, \lambda_s) + \kappa + \varepsilon$ is satisfied. If $\lambda_s \geq 2\lambda_w$, then the condition holds when $\kappa + \varepsilon$ is small enough. Lastly, the condition $f(2\lambda_s, 2\lambda_w) \geq \kappa + \varepsilon$ excludes a deviation changing all negative links to the positive links.

Result 2: One of the weak players is the strong players’ friend.

First, for $S_1$ and $S_2$, extending the negative link to $W_1$ is the only possible deviation which can be more profitable than $g$. It is trivial to show that the other deviations are always non-profitable than $g$. The strong type players $S_1$ and $S_2$ do not have an incentive to deviate from the formation if and only if $f(2\lambda_s + \lambda_w, \lambda_w) - \kappa - \varepsilon \geq f(2\lambda_s, \lambda_w) + f(2\lambda_s, \lambda_s + \lambda_w) - 2\kappa - 2\varepsilon$.

When one of them extends a negative link to the weak type player $W_1$, the new payoff from the negative link $f(2\lambda_s, \lambda_s + \lambda_w)$ and the new cost $\kappa + \varepsilon$ occurs. At the same time, the payoff from the existing positive link is decreased from $f(2\lambda_s + \lambda_w, \lambda_w)$ to $f(2\lambda_s, \lambda_w)$. Considering these factors, $u_{S_i}(g')$ with the deviation should be more profitable than $u_{S_i}(g)$. The condition $f(2\lambda_s + \lambda_w, \lambda_w) - \kappa - \varepsilon \geq f(2\lambda_s, \lambda_w) + f(2\lambda_s, \lambda_s + \lambda_w) - 2\kappa - 2\varepsilon$ makes the deviation not more profitable and the strong players stick to $g$.

Second, the weak type player $W_1$ always does not have an incentive to deviate from the formation $g$. If he extends any negative links to the strong type players, payoff from the
negative link is negative because his network power \( n_{W1}(g') = \lambda_w + \lambda_s \) or \( \lambda_w \) is always less than \( n_{S1}(g') = 2\lambda_s \). If he extends a positive link to the weak type player \( W_2 \), then \( u_{W1}(g) = 0 \).

Third, it is trivial to show that the weak type player \( W_2 \) does not have an incentive to deviate from the formation. Even though he changes his positive directed link to the negative one, any undirected link does not change, but the only cost for the negative link happens.

Lastly, \( f(2\lambda_s + \lambda_w, \lambda_s) - \kappa - \varepsilon > 0 \) is a condition for \( S_1, S_2, \) and \( W_1 \) not to extend the positive link to \( W_2 \).

Result 3: One of strong players is weak players’ friend

The first condition i) is from \( f(\lambda_s + 2\lambda_w, \lambda_s) - \kappa - \varepsilon \geq f(\lambda_s + \lambda_w, \lambda_s) + f(\lambda_s + \lambda_w, 2\lambda_w) - 2\kappa - 2\varepsilon \).

If the condition is satisfied, \( S_1 \) does not have an incentive to deviate by extending the negative links to one of the weak players.

The second condition ii) ensures all of the other deviations are not more profitable than \( g \). First, if \( f(\lambda_s + 2\lambda_w, \lambda_s) - \kappa - \varepsilon \geq 0 \), then \( S_1, W_1, \) and \( W_2 \) will not extend the positive link to \( S_2 \). Second, if \( f(\lambda_s + 2\lambda_w, \lambda_s) - \kappa - \varepsilon \geq f(2\lambda_s + \lambda_w, \lambda_w) - \kappa - \varepsilon \), \( S_1 \) does not have an incentive to switch his friend from one of the players to \( S_1 \). Third, if \( f(\lambda_s + 2\lambda_w, \lambda_s) - \kappa - \varepsilon \geq 2f(2\lambda_s, 2\lambda_w) - 2\kappa - 2\varepsilon \), \( S_1 \) does not have an incentive to change his friend set from \( \{W_1, W_2\} \) to \( \{S_1\} \). If the third condition is satisfied, a deviation where \( S_1 \) only extends the negative links to \( W_1 \) and \( W_2 \) are also not profitable.

Result 4: Non-Nash equilibrium formations

From (d) to (i) in Figure 1, the included formations cannot be Nash equilibrium. I will simply mention a profitable deviation from each suggested formations. In (d), there should be at least one directed negative link between the teams. However, both team players have the same strengths. Therefore, the players who are extending the negative links always have an incentive to extend positive links to save the cost. In (e), the strong players always have an incentive to extend negative links to the weak players. In (f), the isolated strong player’s payoff is \( 2f(\lambda_s, \lambda_s + \lambda_w) + f(\lambda_s, \lambda_w) < 0 \). If the strong player extends a positive link to the isolated player, then his payoff from the new formation is 0. In (g), the weak players always have an incentive to extend a positive link to form the same team. In similar, the strong players in (h) always have an incentive to extend a positive link to form the same team. Lastly, in (i), the strong players and the weak players have an incentive to form the same team with the same type of player.
Appendix B. when there are three strong type players and one weak type players.

Proof of Result 9
Suppose there are $n$ players, $N = \{1, 2, ..., n\}$, with $n \geq 3$, such that Player 1’s intrinsic power $\lambda_1$ is strictly weaker than the others’ and the others’ are the same, $\lambda_1 < \lambda_i = \lambda$ for all $i \in N \setminus \{1\}$. Show that a strategy profile $g$ such that the players except 1 are friends with each other and Player 1 is the enemy to everyone else is a Nash equilibrium.

Answer
Consider the following strategy profile $g = (g_1, g_2, ..., g_n)$, where $g_i$ is Player $i$’s strategy for all $i \in N$. $g_i$ is a row vector $(g_{i,1}, g_{i,2}, ..., g_{i,i-1}, g_{i,i+1}, ..., g_{i,n})$ where $g_{i,j} \in \{-1, 1\}$ for each $j \in N \setminus \{i\}$

- Player 1: $g_{1,j} = 1, \forall j \in N \setminus \{1\}$
- Player $k \in N \setminus \{1\}$: $g_{k,1} = -1$, and $g_{k,l} = 1, \forall l \in N \setminus \{1, k\}$

We will prove that $g$ is a Nash equilibrium by showing that the payoff of any given player $i$’s unilateral deviation $g'_i$ is not strictly higher than her payoff under $g$. The deviation strategy profile including $g'_i$ is denoted by $g' = (g'_i, g_{-i})$ where $g_{-i}$ is a set consisting of the other players’ strategies except Player $i$’s strategy.

Step 1: Let us show Player 1’s strategy $g_1$ such that $g_{1,k} = 1, \forall k \in N \setminus \{1\}$ is the best response to the other players strategy $g_{-1}$. Suppose player 1 deviates from $g_{1,k} = 1, \forall k \in N \setminus \{1\}$. Then any possible deviation $g'_1 \neq g_1$ is such that $\exists K \neq \emptyset$ and $K \subseteq N \setminus \{1\}$, where $g'_{1,k} = -1$ for all $k \in K$, and $g'_{1,k} = 1$ otherwise. Before the deviation, the payoff is

$$u_1(g_1, g_{-1}) = \sum_{j \in N \setminus \{1\}} f(n_1(g_1, g_{-1}), n_j(g_1, g_{-1})) - (N - 1)\kappa.$$  \hspace{1cm} (3)

On the other hand, the deviation strategy derives

$$u_1(g'_1, g_{-1}) = \sum_{j \in N \setminus \{1\}} f(n_1(g'_1, g_{-1}), n_j(g'_1, g_{-1})) - |K|\varepsilon - (N - 1)\kappa.$$  \hspace{1cm} (4)

There is another way to prove step 1 using a lemma in Timo Hiller(2017). As long as Lemma 2 is valid, Player 1 with $n_1 < n_k$ for $\forall k \in N \setminus \{1\}$ makes positive links to the stronger players. $n_1 < n_k$ is true because $n_1 = \lambda_1$ but $n_k = \sum_{j \in N \setminus \{1\}} \lambda_j = (n - 1)\lambda$. According to the assumption $\lambda_1 < \lambda_i = \lambda$, $n_1 = \lambda_1 < (n - 1)\lambda = n_k$.  

18
where |K| is the cardinality of the set K.

Note that Equation (2) comes from the fact that even though Player 1 chose the deviation strategy, there is no change on the undirected network g because the other players already have initiated the negative link to player 1. As a result, $n_1(g'_1, g_1) = n_1(g_1, g_1)$ and $n_j(g'_1, g_1) = n_j(g_1, g_1)$, thus $f(n_1(g'_1, g_1), n_j(g'_1, g_1)) = f(n_1(g_1, g_1), n_j(g_1, g_1))$. Hence $u_1(g'_1, g_1) > u_1(g'_1, g_1)$ and the deviation strategy is not profitable for Player 1.

Step 2: We will show that, for any player $k \in N \setminus \{1\}$, the strategy $g'_k$ is such that $g'_{k,1} = -1$, $g'_{k,l} = 1$, $\forall l \in N \setminus \{1,k\}$, and it is the best response against the other players’ given strategies $g_{-k}$. There are two cases of possible deviation strategy $g'_k$. (a) $\exists L \subset N \setminus \{1,k\}$: $g + \sum_{l \in L} g_{k,l}$. It changes the positive directed links among the friends to the negative link, maintaining the existing negative link to Player 1. (b) $\exists L \subset N \setminus \{1,k\}$: $g + g_{k,1} + \sum_{l \in L} g_{k,l}$. It also changes the positive directed links among the friends to the negative link and turns the negative link to Player 1 to the positive link at the same time.

In the first deviation strategy (a), $g' = g + \sum_{l \in L} g_{k,l}$ and any possible deviation $g'_k \neq g_k$ is such that $\exists L \neq \emptyset$ and $L \subseteq N \setminus \{1,k\}$, where $g'_{k,1} = -1$, $g'_{k,l} = -1$ for all $l \in L$, and $g'_{k,l} = 1$ otherwise. Before the deviation,

$$u_k(g_k, g_{-k}) = f(n_k(g_k, g_{-k}), n_1(g_k, g_{-k})) - \varepsilon - \kappa. \quad (5)$$

After the deviation,

$$u_k(g'_k, g_{-k}) = f(n_k(g'_k, g_{-k}), n_1(g'_k, g_{-k})) + \sum_{j \in L} f(n_k(g'_k, g_{-k}), n_j(g'_k, g_{-k}))$$

$$- (|L| + 1)\varepsilon - (|L| + 1)\kappa. \quad (6)$$

where |L| is the cardinality of the set L. It is not profitable because Player k gets 0 extraction value at most from the other player $l \in L$ whose intrinsic power $\lambda_l = \lambda_k = \lambda$. If Player $k$ changes one link ($|L| = 1$), $f(n_k(g'_k, g_{-k}), n_l(g'_k, g_{-k})) = 0$ because $n_k(g'_k, g_{-k}) = n_l(g'_k, g_{-k}) = (n-2)\lambda$. If he changes more than two links ($|L| \geq 2$), then $f(n_k(g'_k, g_{-k}), n_l(g'_k, g_{-k})) < 0$ because $n_k(g'_k, g_{-k}) = (n-m-1)\lambda$ but $n_l(g'_k, g_{-k}) = (n-2)\lambda$. Therefore, $\sum_{j \in L} f(n_k(g'_k, g_{-k}), n_j(g'_k, g_{-k})) < 0 \forall L \neq \emptyset$. Moreover, Player $k$ extracts less after the deviation from Player 1. Before the deviation, $n_k(g_k, g_{-k}) = (n-1)\lambda$, but after the deviation, $n_k(g'_k, g_{-k}) = (n-|L|-1)\lambda$. According to the assumption in the model, $f$ is strictly increasing in $n_k$. $n_1(g_k, g_{-k}) = n_1(g'_k, g_{-k})$ because there is no change in the undirected links related to Player 1. Therefore, $f(n_k(g_k, g_{-k}), n_1(g'_k, g_{-k})) < f(n_k(g_k, g_{-k}), n_1(g_k, g_{-k}))$. As a result of the deviation, the gross payoff $f(n_k(g_k, g_{-k}), n_1(g_k, g_{-k}))$ is decreased to $\sum_{j \in L} f(n_k(g'_k, g_{-k}), n_j(g'_k, g_{-k})) +
\(f(n_k(g'_k, g_{-k}), n_1(g'_k, g_{-k}))\) and the costs are increased. Therefore, the first deviation is not profitable.

(b) In the second way, Player \(k\) can make new positive link and negative links at the same time, \(g + g'_{k,1} + \sum_{l \in L} g_{k,l}\). Any possible deviation \(g'_k \neq g_k\) is such that \(\exists L\) and \(L \subseteq N \setminus \{1, k\}\), where \(g'_{k,1} = 1, g'_{k,l} = -1\) for all \(l \in L\), and \(g'_{k,l} = 1\) otherwise. We will show that the deviation is also not profitable for Player \(k\) compared to the given strategy \(g_k\). The payoff before the deviation is the same to the equation (3). After the deviation,

\[
u_k(g'_k, g_{-k}) = \sum_{j \in L} f(n_k(g'_k, g_{-k}), n_j(g'_k, g_{-k})) - |L|\varepsilon - |L|\kappa. \tag{7}
\]

Here, \(n_k(g'_k, g_{-k}) = (n - |L| - 1)\lambda + \lambda_1\) and \(n_j(g'_k, g_{-k}) = (n - 2)\lambda\). Then we can rewrite \(u_k(g'_k, g_{-k})\) that

\[u_k(g'_k, g_{-k}) = |L|f((n - |L| - 1)\lambda + \lambda_1, (n - 2)\lambda) - |L|\varepsilon - |L|\kappa.
\]

If \(L = \emptyset\), \(|L| = 0\), so \(g'_k\) is not profitable deviation. When \(L \neq \emptyset\) so \(|L| \geq 1\), \((n - 1)\lambda > (n - |L| - 1)\lambda + \lambda_1\) and \((n - 2)\lambda > \lambda_1\), hence \(((n - 1)\lambda, \lambda_1) > f((n - |L| - 1)\lambda + \lambda_1, (n - 2)\lambda)\). Hence the deviation is not profitable. Therefore, there is no incentive for Player \(k\) to choose any possible deviation strategy \(g'_k\).

Therefore, all players have no incentive to deviate from the strategy profile. Thus, the strategy profile is Nash equilibrium.

**Appendix C. Proof for Result 10 to 13: The generalized conditions when the same type players form a clique**

\(n\lambda_s > m\lambda_w\)

**Proof strategy**

Beginning the proof, I want to mention that we need to check only the strong group players’ incentive to deviate from the strategy \(g\). It is because players in the weak group do not have any profitable deviating strategy. In any deviations, the players in the weak group can only make new enemies. However, extending any negative link to the strong type enemy cannot change the relationship but adds more cost. Similarly, when they extend the negative links to the weak type friends, there is no or negative extraction, but the deviation brings more cost. Therefore, it is enough to check the strong type players’ deviation possibilities.
Generally, in Hiller’s model, we should check three kinds of deviations to determine whether it is Nash equilibrium or not. First, the strong group players can extend positive links where there were negative links (making new friends). Second, the players can extend negative links where there were positive links (making new enemies). Lastly, the players can extend positive links and negative links at the same time (switching their friends). In PAM, we do not need to worry about the first case when we check the strong group players’ incentives. It is because, as a result of the deviation, there is no positive extraction but costs happened. He will have the same or less power than the new enemy after the deviation because any deviating player and his new enemy has the same intrinsic power. I will derive the condition in which the second and the third deviations are not more profitable than the suggested strategy profile.

When \( n\lambda_s > m\lambda_w \), the strong type players become the strong group players. They extend the positive links to the same group players and extend the negative links to the weak type players. The weak type players are the weak group players and extend the positive links to all other players.

For most of the cases, there is a common condition making the second deviation non-profitable. However, the condition for the third deviation to be unattractive varies with the number of players.

\( n\lambda_s > m\lambda_w; \ n > m \)

First, I will check the condition that the strong players’ deviation of making new friends is not profitable. Formally, the strong player \( i \)’s deviation strategy is \( g'_i = g_i + \sum g_{i,j}^+, \ \exists j \in N_w \).

Let’s denote the player \( i \)’s utility from the deviation strategy \( g'_i \) \( u'_i \), and the number of new friends (or the number of new negative links) \( b \).

Before the deviation,

\[
  u_i = mf(n\lambda_s, m\lambda_w) - m\kappa - m\varepsilon \\
  = \frac{mn\lambda_s}{n\lambda_s + m\lambda_w} - \frac{m}{2} - m\kappa - m\varepsilon
\]
After the deviation,

\[ u_i^d = (m - b)f(n\lambda_s + b\lambda_w, m\lambda_w) - (m - b)(\kappa + \varepsilon) \]
\[ = \frac{(m - b)(n\lambda_s + b\lambda_w)}{n\lambda_s + (m + b)\lambda_w} - \frac{m - b}{2} - (m - b)(\kappa + \varepsilon) \]

If \( u_i \geq u_i^d \forall b \in \{1, 2, ..., m\} \), then there is no incentive for the second type of deviation. As \( \kappa \& \varepsilon \) go to 0, we have

\[ \frac{mn\lambda_s}{n\lambda_s + m\lambda_w} - \frac{m}{2} \geq \frac{(m - b)(n\lambda_s + b\lambda_w)}{n\lambda_s + (m + b)\lambda_w} - \frac{m - b}{2} \]
\[ \frac{mn\lambda_s}{n\lambda_s + m\lambda_w} \geq \frac{(m - b)(n\lambda_s + b\lambda_w)}{n\lambda_s + (m + b)\lambda_w} - \frac{b}{2} \]

For convenience in calculation, let some of the terms be normalized, \( \hat{b} = \frac{b}{m}, \lambda_w = 1, \lambda_s = a\lambda_w = a, \hat{n} = \frac{n}{m} \). Then,

\[ \frac{\hat{n}a}{\hat{n}a + 1} \geq \frac{(1 - \hat{b})(\hat{n}a + \hat{b}\lambda_w)}{\hat{n}a + (1 + \hat{b})\lambda_w} - \frac{\hat{b}}{2} \]

If we solve it, we can get,

\[ \hat{n}a \geq \frac{1}{2}(1 - \sqrt{12 - 4\hat{b} + \hat{b}^2}) \]

For \( g_i \) to be a Nash equilibrium, it should hold \( \forall \hat{b} \in \{\frac{1}{m}, \frac{2}{m}, ..., 1\} \subset [0, 1] \). If we differentiate the right-hand side about \( \hat{b} \),

\[ \frac{dRHS}{d\hat{b}} = \frac{1}{2}(-1 + \frac{1}{2}(-4 + 2\hat{b})(12 - 4\hat{b} + \hat{b}^2)^{-\frac{1}{2}}) \leq 0 \text{ for } \hat{b} \in [0, 1] \]

So as \( \hat{b} \) increases, RHS is decreased. RHS will be maximized with the smallest \( \hat{b} = \frac{1}{m} \). Hence,
RHS at $\hat{b} = \frac{1}{m}$ is the critical point for the second type of deviations.

\[ \hat{n}a \geq \frac{1}{2} \left( -\frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}} \right) \]
\[ = \frac{1}{2m} \left( -1 + \sqrt{12m^2 - 4m + 1} \right) \]
\[ a \geq \frac{1}{2n} \left( -1 + \sqrt{12m^2 - 4m + 1} \right) \]

Therefore, if and only if $a = \frac{\lambda_s}{\lambda_w} \geq \frac{1}{2n} \left( -1 + \sqrt{12m^2 - 4m + 1} \right)$, the second deviation is not profitable for the strong type players.

Next, I will check the condition that the third type deviation switching their friends is not profitable. Formally, the strong player $i$’s deviation strategy $g'_i = g_i + \sum g^+_{i,j} + \sum g^-_{i,k}$, $\exists j \in N_w$, $\exists k \in N_s \setminus \{i\}$. After the deviation, the player $i$’s utility is

\[ u'_i = (m - b) f((n - d)\lambda_s + b\lambda_w, m\lambda_w) + df((n - d)\lambda_s + b\lambda_w, (n - 1)\lambda_s) - (m - b + b)(\kappa + \varepsilon) \]
\[ = \frac{(m - b)((n - d)\lambda_s + b\lambda_w)}{(n - d)\lambda_s + (m + b)\lambda_w} + \frac{d((n - d)\lambda_s + b\lambda_w)}{(n - d)\lambda_s + b\lambda_w + (n - 1)\lambda_s} - \frac{m - b + d}{2} - (m - b)(\kappa + \varepsilon) \]

where $b$ is the number of new positive links, and $d$ is the number of new negative links.

In this section it is supposed that the strong type players are more than the weak type players, $n > m$. Because $n$ and $m$ are natural numbers, $n > m$ is the same to $n - 1 \geq m$. Here, I will show that when $n - 1 \geq m$, the number of my friends are larger than the number of enemies in $g$, the deviation strategy making a positive number of new friends and a positive number of new enemies cannot be more profitable deviation strategy than the deviation only making non-zero friends. Then even though we do not check the incentive for the third type of deviations the third-type deviation, if the second type deviations are not profitable, the third type deviations are also not profitable, thus the suggested strategy profile is a Nash equilibrium.

Suppose there exist $\tilde{b}$ and $\tilde{d}$ which attain the most profitable deviation among the deviations across the second and third types. Formally, $\exists \tilde{b} > 0, \tilde{d} > 0$ s.t. $u'_i(\tilde{b}, \tilde{d}) > u'_i(b, d)$ $\forall b \in \{1, 2, ..., m\}$, $d \in \{1, 2, ..., n - 1\}$. Let’s consider an alternative strategy with $\tilde{b} = \tilde{b} - 1$ and $\tilde{d} = \tilde{d} - 1$. It means that the player $i$ makes one less weak type friend and one more strong type friend than the $(\tilde{b}, \tilde{d})$ strategy. It is the same that the player $i$ makes one more weak type enemy and one less strong type enemy. After the deviation $(\tilde{b}, \tilde{d})$, the player $i$’s
strength in the network is
\[ \tilde{n}_i = (n - \bar{d})\lambda_s + \bar{b}\lambda_w. \]

With the alternative deviation \((\tilde{b}, \tilde{d})\), the player \(i\)'s strength is
\[ \tilde{n}_i = (n - \bar{d})\lambda_s + \tilde{b}\lambda_w = (n - \bar{d} + 1)\lambda_s + (\tilde{b} - 1)\lambda_w \]
\[ = (n - \bar{d})\lambda_s + \tilde{b}\lambda_w + \lambda_s - \lambda_w > \bar{n}_i. \]

The strong type enemies have the same strengths in both networks \((n - 1)\lambda_s\). And the weak type enemies in both networks have the strength \(m\lambda_w\). \((n - 1)\lambda_s > m\lambda_w\) is derived by the assumptions \(n - 1 \geq m\) and \(\lambda_s > \lambda_w\). Thus, the new weak type enemy has less power than the strong type enemy. As a result, in the alternative strategy with \((\tilde{b}, \tilde{d})\), the player \(i\) can increase his extraction from the new weak type enemy compared to the extraction with the strong type enemy in the \((\bar{b}, \bar{d})\) strategy. For the rest of the enemies who are common in both \((\bar{b}, \bar{d})\) and \((\tilde{b}, \tilde{d})\), he can also extract more because his strength \(n_i\) is increased with the strategy \((\tilde{b}, \tilde{d})\).

We can also show the above by using utility functional form as follows.
\[ \bar{u}_i = (m - \tilde{b})f((n - \bar{d})\lambda_s + \tilde{b}\lambda_w, m\lambda_w) + \tilde{d}f((n - \bar{d})\lambda_s + \tilde{b}\lambda_w, (n - 1)\lambda_s) - (m - \tilde{b} + \tilde{d})(\kappa + \varepsilon) \]
\[ = (m - \tilde{b})f((n - \bar{d} + 1)\lambda_s + (\tilde{b} - 1)\lambda_w, m\lambda_w) + \tilde{d}f((n - \bar{d} + 1)\lambda_s + (\tilde{b} - 1)\lambda_w, (n - 1)\lambda_s) \]
\[ + f((n - \bar{d} + 1)\lambda_s + (\tilde{b} - 1)\lambda_w, m\lambda_w) - f((n - \bar{d} + 1)\lambda_s + (\tilde{b} - 1)\lambda_w, (n - 1)\lambda_s) \]
\[ - (m - \tilde{b} + \tilde{d})(\kappa + \varepsilon) \geq \]
\[ \bar{u}_i = (m - \tilde{b})f((n - \bar{d})\lambda_s + \tilde{b}\lambda_w, m\lambda_w) + \tilde{d}f((n - \bar{d})f((n - \bar{d})\lambda_s + \tilde{b}\lambda_w, (n - 1)\lambda_s) - (m - \tilde{b} + \tilde{d})(\kappa + \varepsilon) \]

Therefore, any strategy with \((\tilde{b}, \tilde{d})\) s.t. \(\tilde{b} > 0\) and \(\tilde{d} > 0\) cannot be the most profitable deviation. Note that if \(\tilde{b} = 0\) and \(\tilde{d} > 0\), it means the first type of deviations, if \(\tilde{b} > 0\) and \(\tilde{d} = 0\), it means the second type of deviations. It was mentioned above that the first deviations are always not profitable. Therefore, if the second type of deviations is not profitable, the first and third types of deviations are also automatically unprofitable. Therefore, the condition to check the second type deviation possibility becomes the necessary and sufficient condition regardless of the first and third deviation conditions.

This proof is desirable because it holds not only for the normalized contest function but also for any extraction functions. Hence when we try to check the possibility of the third
type of deviation, we can use the same logic. It will be helpful when I generalize the result with other functions in future research.

\[ n \lambda_s > m \lambda_w; \ n = m \]

This subsection will cover the case when the number of the strong type players is the same to that of the weak type players. Then the strong type enemies have stronger strength than the weak type enemies in any deviations. As mentioned before, it is not profitable to make the new strong type enemies by using the first type of deviations. To check the incentive for the second type of deviations we can reuse the above calculation by inserting \( n \) into \( m \). As a result,

\[
a \geq \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right)
\]

In similar to discussed above, the third type of deviations is unprofitable when \((n - 1)\lambda_s > m\lambda_w = n\lambda_w\). Therefore,

\[
a > \frac{n}{n - 1}
\]

is another sufficient condition to exclude the third-type deviation. When \( n = 4 \), \( \frac{n}{n - 1} = \frac{4}{3} = 1.3333 \ldots \), \( a \geq \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right) = \frac{1}{8} \left( -1 + \sqrt{177} \right) \), approximately 1.53801. So the larger value \( \frac{1}{8} \left( -1 + \sqrt{177} \right) \) becomes the necessary and sufficient condition. For \( n > 4 \), the condition \( \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right) \) is increased but \( \frac{n}{n - 1} \) is decreased. Therefore, \( \forall n \geq 4 \), \( a \geq \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right) \) is the sufficient and necessary condition. When \( n = 2 \) and \( n = 3 \), \( \frac{n}{n - 1} > \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right) \). It implies that although \( a \geq \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right) \), \( a \) can be less than \( \frac{n}{n - 1} \). Thus we cannot depend on the condition \( a \geq \frac{1}{2n} \left( -1 + \sqrt{12n^2 - 4n + 1} \right) \) as the necessary and sufficient condition like mentioned above. So, I calculated the third condition respectively. As a result of calculation, when \( n = 2 \), the second condition is \( a \geq \frac{1}{6} \left( -1 + \sqrt{41} \right) \approx 1.35078 \). The most profitable deviation among the third type is \( b = 1 \) and \( d = 2 \), and the condition is \( a \geq 1.5 \), and \( \frac{n}{n - 1} = \frac{2}{1} \). Therefore, the necessary and sufficient condition when \( n = 2 \) is \( a \geq 1.5 \). When \( n = 3 \), the second condition is \( a \geq \frac{1}{6} \left( -1 + \sqrt{97} \right) \approx 1.4748096 \). The most profitable deviation among the third type deviations is \( b = 2 \) and \( d = 1 \), and the condition is approximately \( a > 1.39384 \). Hence, in the case when \( n = 3 \), it is the necessary and sufficient condition that \( a \geq \frac{1}{2n} \left( -1 + \sqrt{12m^2 - 4m + 1} \right) \). In conclusion, the necessary and sufficient condition for Nash equilibrium when \( n = m \) is
\[ a \geq 1.5 \quad \text{when } n = 2 \]
\[ \geq \frac{1}{2n} (-1 + \sqrt{12m^2 - 4m + 1}) \quad \text{when } n \geq 3 \]

\[ n\lambda_s > m\lambda_w: n < m \text{ and } (n-1)\lambda_s \geq m\lambda_w \]

This subsection will discuss the case when the number of strong type players is less than that of the weak type players, but the strong type players can still form the strong team(group). This allows them to extract some payoff from the weak players. When \((n - 1)\lambda_s > m\lambda_w\) so \(a > \frac{m}{n-1}\), we can apply the same method used in the previous subsections to check the third type of deviations. Therefore, if the strength ratio \(a = \frac{\lambda_s}{\lambda_w} \geq \frac{1}{2n} (-1 + \sqrt{12m^2 - 4m + 1})\), it excludes the second type deviation possibility. When \(n = 2, \frac{1}{2n} (-1 + \sqrt{12m^2 - 4m + 1}) < \frac{m}{n-1}\) for all \(m\). Because

\[
\frac{1}{2n} (-1 + \sqrt{12m^2 - 4m + 1}) = \frac{m}{n-1} \frac{n-1}{2n} (- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}})
\]

when \(n = 2, \frac{n-1}{2n} = \frac{1}{4}\), but the supremum of \(- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}}\) for \(m > n = 2\) is \(\sqrt{12} \approx 3.4641\). Therefore, when \(n = 2\), \(g\) is a Nash equilibrium with any intrinsic strength ratio \(a\).

On the other hand, when \(n = 3, \frac{n-1}{2n} = \frac{2}{5}\). When \(m = 4, (- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}}) = \frac{\sqrt{127}}{4} - \frac{1}{4} \approx 3.07603\). So \(\frac{n-1}{2n} (- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}}) > 1, \frac{n-1}{2n} \) increases in \(n\), and \(- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}} \) increases in \(m\). As a result, for any \(n \geq 3, \frac{1}{2n} (-1 + \sqrt{12m^2 - 4m + 1}) > \frac{m}{n-1}\), and the necessary and sufficient condition is \(a \geq \frac{1}{2n} (-1 + \sqrt{12m^2 - 4m + 1})\).

\[ n\lambda_s > m\lambda_w: n < m \text{ and } (n-1)\lambda_s < m\lambda_w < n\lambda_s \]

In this subsection, I will check a special case where \((n - 1)\lambda_s < m\lambda_w < n\lambda_s\). The inequalities mean that the strong type players can form the strong group. However, if any strong type player betrays some players in the same group, and extends the negative link, then the
betrayed player’s network power is weaker than the weak type players’ network power. The condition \((n-1)\lambda_s < m\lambda_w < n\lambda_s\) can be transformed to \(\frac{m}{n} < a = \frac{\lambda_s}{\lambda_w} < \frac{m}{n-1}\) regarding \(a\). Recall that to exclude the second type of deviations, \(a \geq \frac{1}{2n}(-1 + \sqrt{12m^2 - 4m + 1}) = \frac{m}{n-1} \frac{n-1}{2n}(\frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}})\). When \(n = 2\), \(\frac{m}{n-1} = \frac{1}{4}\) and the supremum of \(\left(- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}}\right)\) is \(\sqrt{12} \approx 3.464\). Thus, \(\frac{m}{n-1} > \frac{m}{n-1} \frac{n-1}{2n} \left(- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}}\right) = \frac{1}{2n}(-1 + \sqrt{12m^2 - 4m + 1})\). On the other hand, \(\frac{m}{n} < \frac{m}{n} \left(-1 + \sqrt{12m^2 - 4m + 1}\right)\) because \((-1 + \sqrt{12m^2 - 4m + 1}) > 2\) for all \(m\). Therefore, when \(n = 2 < m\), \(\frac{1}{4}(-1 + \sqrt{12m^2 - 4m + 1}) \leq a < m\) guarantees \(g\) to be a Nash equilibrium.

However, if \(n \geq 3\), there does not exist \(a\) making \(g\) be a Nash equilibrium on this interval. When \(n \geq 3\), \(\frac{m}{n-1} \frac{n-1}{2n} \left(- \frac{1}{m} + \sqrt{12 - \frac{4}{m} + \frac{1}{m^2}}\right) > \frac{m}{n-1} \cdot a \geq \frac{1}{2n}(-1 + \sqrt{12m^2 - 4m + 1})\) to exclude the second type of deviations, but \(a < \frac{1}{n-1}\) by the assumption. It is a contradiction. Therefore, for the range of \((n-1)\lambda_s < m\lambda_w < n\lambda_s\), only \(n = 2\) and \(\frac{1}{4}(-1 + \sqrt{12m^2 - 4m + 1}) \leq a < m\) are the necessary and sufficient condition for \(g\) to be a Nash equilibrium.

Overall, When \(n\lambda_s > m\lambda_w\), type \(s(w)\) players become the strong(weak) team member. PAM is a Nash equilibrium if and only if i) When \(n = 2\) and \(m = 2\), \(a \geq 1.5\), ii) When \(n > 3\) or \(m > 3\), \(a \geq \frac{1}{2n}(-1 + \sqrt{12m^2 - 4m + 1})\). Thus,

\[
\bar{a}(n, m, \varepsilon, \kappa) \geq 1.5 \quad \text{when } n = 2, m = 2
\]
\[
\geq \frac{1}{2n}(-1 + \sqrt{12m^2 - 4m + 1}) \quad \text{Otherwise}
\]

\(n\lambda_s = m\lambda_w\)

It is simple to argue that there is no PAM Nash equilibrium strategy when \(n\lambda_s = m\lambda_w\). In any PAM \(g\), the extraction value is 0, and any players utility, who extend the negative links, is negative due to the negative link cost \(\kappa\) and \(\varepsilon\). Moreover, to extend the positive link to the existing enemy is always profitable. Therefore, there is always an incentive to deviate from \(g\) when \(n\lambda_s = m\lambda_w\).
In this section, I will check the situation when the weak type players form the strong group, and the strong type players form the weak group because the number of strong players is few. In similar, the first type of deviations only making new enemies is not profitable. Here, the new enemies are the weak type players in the same strong group. In the following subsection, I will check the second and third types of deviations.

\[ m\lambda_w > n\lambda_s : (m-1)\lambda_w \leq n\lambda_s < m\lambda_w \]

The condition \((m-1)\lambda_w < n\lambda_s < m\lambda_w\) means that if the weak type strong group player \(i\) extends a negative link to the same group weak player \(j\), the player \(j\)'s network power \((m-1)\lambda_w\) is lower than any strong type players' network power \(n\lambda_s\) in the weak group. This condition is changed to \(\frac{m-1}{n} < a = \frac{\lambda_s}{\lambda_w} < \frac{m}{n}\). It implies that the intrinsic power ratio \(a\) is quite large. When this condition holds, there is no Nash equilibrium PAM. It is because there is always more profitable deviation for any weak type player.

With the suggested strategy formation \(g\), the utility \(u\) is

\[ u_i = n\left(\frac{m\lambda_w}{m\lambda_w + n\lambda_s} - \frac{1}{2}\right) - n\kappa - n\varepsilon. \]

Now, let's consider the deviation strategy \(g'\) that \(b = 1\) and \(d = 1\). Comparing with \(g\), the player \(i\) switches his one weak type friend and one strong type enemy. The utility \(u'_i\) from the deviation strategy \(g'\) is

\[ u'_i = (n-1)\left(\frac{(m-1)\lambda_w + \lambda_s}{(m-1)\lambda_w + \lambda_s + n\lambda_s} - \frac{1}{2}\right) + \left(\frac{(m-1)\lambda_w + \lambda_s}{(m-1)\lambda_w + \lambda_s + (m-1)\lambda_w} - \frac{1}{2}\right) - n\kappa - n\varepsilon. \]

Note that

\[ \frac{(m-1)\lambda_w + \lambda_s}{(m-1)\lambda_w + \lambda_s + n\lambda_s} > \frac{m\lambda_w}{m\lambda_w + n\lambda_s} \]

and

\[ \frac{(m-1)\lambda_w + \lambda_s}{(m-1)\lambda_w + \lambda_s + (m-1)\lambda_w} > \frac{m\lambda_w}{m\lambda_w + n\lambda_s}. \]

Therefore, \(u'_i > u_i\). It shows that the player \(i\) have an incentive to deviate from \(g\) by choosing \(g'\), and \(g\) is not a Nash equilibrium.

\[ m\lambda_w > n\lambda_s : n\lambda_s < (m-1)\lambda_w \]

Proof strategy
For any weak type player $i$, there are three kinds of deviation strategies from $g_i$. First, $g_i^1 = g_i + \sum_{j \in J_i} g_{ij}^+ + \sum_{k \in K_1} g_{ik}^-$ where nonempty sets $J_1 \subset N_s$, $K_1 \subset N_w$. Second, $g_i^2 = g_i + \sum_{j \in J_2} g_{ij}^+$ where nonempty set $J_2 \subset N_s$. Third, $g_i^3 = g_i + \sum_{k \in K_3} g_{ik}^-$ where nonempty set $K_3 \subset N_w$. The proof consists of three parts. First, I will show that if $u(g) \geq u(g_i^1, g_{-i})$ and $u(g) \geq u(g_i^3, g_{-i})$ for all $J_2, K_3$, then $u(g) \geq u(g_i^1, g_{-i})$ for all $J_1$ and $K_1$. Second, I will derive a condition not to choose the deviation strategies $g_i^2$. Third, I will explain how to derive a sufficient condition not to choose $g_3$. Also, another sufficient condition to choose $g_3$ will be presented, too.

First, if $u(g) \geq u(g_i^1, g_{-i})$ and $u(g) \geq u(g_i^3, g_{-i})$ for all $J_2, K_3$, then $u(g) \geq u(g_i^1, g_{-i})$ for all $J_1$ and $K_1$. To prove it, suppose there exist some $J_1$ and $K_1$ such that $u(g) < u(g_i^1, g_{-i})$ when $u(g) \geq u(g_i^3, g_{-i})$ and $u(g) \geq u(g_i^3, g_{-i})$ for all $J_2, K_3$. Then

$$(1)(n-b')f(m+b'a-d', na) + d'f(m+b'a-d', m-1) > nf(m, na)$$

for some $b'$ and $d'$ such that $0 < b' \leq n$ and $0 < d' \leq m-1$

$$(2)nf(m, na) \geq (n-b) f(m+ba, na)$$ for all $b$ such that $0 < b \leq n$

$$(3)nf(m, na) \geq df(m-d + na, m-1)$$ for all $d$ such that $0 < d \leq m-1$

Then by using (1) and (2),

$$(n-b')f(m+b'a-d', na) + d'f(m+b'a-d', m-1) - (n-b)f(m+ba, na) > 0$$

for all $b$ and $d$ and some $b'$ and $d'$.

$$n(f(m+b'a-d', na)-f(m+ba, na)) + d'f(m+b'a-d', m-1) - b'f(m+b'a-d', na) + nf(m+ba, na) > 0$$

When $ba = b'a - d'$, $b' - b = \frac{d'}{a}$,

$$d'f(f(m+b'a-d', m-1) - \frac{d'}{a}f(m+b'a-d', na) > 0$$

$$a > \frac{f(m+b'a-d', na)}{f(m+b'a-d', m-1)}$$

By using (1) and (3),

$$(n-b')f(m+b'a-d', na) + d'f(m+b'a-d', m-1) - df(m-d+na, m-1) > 0$$

29
for all $b$ and $d$ and some $b'$ and $d'$. When $ba' - d' = -d + na, d' - d = -(n - b')a$.

\[(n - b')f(m + b'a - d', na) - (n - b')af(m + b'a - d', m - 1) > 0\]

\[a < \frac{f(m + b'a - d', na)}{f(m + b'a - d', m - 1)}\]

It is a contradiction.

Second, if $a \leq \frac{\sqrt{12n^2 - 4n + 1}m + m}{2(3n^2 - n)}$, then $u(g) \geq u(g_i^2, g_{-i})$ for all $g^2$. The condition $a \leq \frac{\sqrt{12n^2 - 4n + 1}m + m}{2(3n^2 - n)}$ is derived from $u(g) \geq u(g_i', g_{-i})$ where $g_i' = g_i + g_{ij}$, $j \in N_s$. When $u(g) \geq u(g_i^2, g_{-i})$, then

\[n - b m + ba - na \geq \frac{n m - na}{2} \frac{m + ba + na}{2 m + na}\]

for all $b \in \{1, 2, ..., n\}$. Then $a \leq \frac{mn^2 + 2\sqrt{2n^4 + n^3m + n^2 - n}}{n^2}$ is a sufficient condition not to choose all $g^3$, and $-4n^2 + 2\sqrt{2n^4 + n^3m + n^2 - n}$ is a necessary condition not to choose all $g^3$. The utility from $g^3$ is

\[u(g^3) = \frac{d m - d + na - (m - 1)}{2} \frac{m + na}{m + na} - (m - 1)\]

There is an upper bound of $u(g^3)$ that

\[\frac{d m - d + na - (m - 1)}{2} \frac{m + na}{m + na} > u(g^3)\]

The maximum of LHS is achieved when $d = \frac{na + 1}{2}$, hence

\[\frac{(\frac{na + 1}{2})^2}{2(m + na)} > u(g^3)\]

If

\[u(g) = \frac{n m - na}{2} \frac{m + na}{m + na} > \frac{(\frac{na + 1}{2})^2}{2(m + na)}\]

Then $u(g) > u(g^3)$ for all $g^3$. It holds when $a \leq \frac{-2n^2 + 2\sqrt{2n^4 + n^3m + n^2 - n}}{n^2}$.

**Appendix D. Proof for Result 14**

First, any player $k \in N_2$ does not have an incentive to deviate from $g$. All players in $N_2$ are extending the positive links in $g$, so the only possible deviation is changing some links from the positive to the negative ones, $g' = g + \sum_{j \in S} g_{ij}$ where $S \subset N \setminus \{i\}$. If the changed
link $g_{ki}$ is directed to $i \in N_1$, the undirected link $\bar{g}_{ki} = -1$ is not changed because $g_{ik} = -1$ already in both $g$ and $g'$. In addition, the cost $\kappa + \varepsilon > 0$ also occurs, so the player $k$ will not choose the deviation. If the changed link $g_{km}$ is heading to $m \in N_2 \setminus \{k\}$, the extraction from the link $f(n_k(g'), n_m(g'))$ is negative because $n_k(g') = n_2 - b$ where $b$ is the number of the negative links to $m \in N_2 \setminus \{k\}$ in $g'$ but $n_m(g') = n_2 - 1$. Because $1 \leq b \leq n_2 - 1$, $n_k(g') \leq n_m(g')$. Moreover, the cost $\kappa + \varepsilon > 0$ occurs too in this case.

Next, the player $i \in N_1$ deviates from $g$ when $n_1 \geq \frac{1}{2}(-1 + \sqrt{1 - 4n_2 + 12n_2^2})$. In $g$, he has the positive links to the players in $N_1$ and the negative links to the players in $N_2$. His possible deviation strategy $g' = g + \sum_{k \in S_2} g_{ik}^+ + \sum_{j \in S_1} g_{ij}^-$ where $S_1 \subset N_1$, $S_2 \subset N_2$. Among the deviations, $g'' = g + g_{ik}^+$, where $k \in N_2$, gives the condition. If we check the condition, we do not need to check the other conditions. It is because some deviations are always less profitable than $g$, and the other deviations are excluded if the condition hold. First of all, let’s check the less profitable deviations. $g^* = g + \sum_{j \in S_1} g_{ij}^-$ where $S_1 \subset N_1$. Let $s_1 = |S_1|$. Before the deviation,

$$u_i(g) = \sum_{k \in N_2} f(n_i(g), n_k(g)) - |N_2|(\kappa + \varepsilon)$$

$$= n_2 f(n_1, n_2) - n_2(\kappa + \varepsilon),$$

and after the deviation by choosing $g^*$,

$$u_i(g') = \sum_{k \in N_2} f(n_i(g'), n_k(g')) + \sum_{j \in S_1} f(n_i(g'), n_j(g')) - |N_2 + S_1|(\kappa + \varepsilon)$$

$$= n_2 f(n_1 - s_1, n_2) + s_1 f(n_1 - s_1, n_1 - 1) - (n_2 + s_1)(\kappa + \varepsilon).$$

When we assume $\kappa + \varepsilon$ small enough, $u_i(g^*) < u_i(g)$ because $f(n_1, n_2) > f(n_1 - s_1, n_2)$ and $f(n_1 - s_1, n_1 - 1) \leq 0$. Now, let’s check the case where the other deviations are excluded if the condition $n_1 \geq \frac{1}{2}(-1 + \sqrt{1 - 4n_2 + 12n_2^2})$ is met. Let’s name all of the other deviations $g = g + \sum_{k \in S_2} g_{ik}^+$, $g^* = g + \sum_{k \in S_2} g_{ik}^+ + \sum_{j \in S_1} g_{ij}^-$ where $S_1 \subset N_1$, $S_2 \subset N_2$ and $S_1 \neq \emptyset$, $S_2 \neq \emptyset$. Then $u_i(g) \leq u_i(g)$ for $g = g + g_{ij}^+ + g_{ik}^-$ where $j \in N_1 \setminus \{i\} \setminus S_1$, $k \in N_2 \setminus S_2$. Let’s
The right hand side is maximized when $n_2$ is the smallest. The smallest $n_2$ is 1, so

$$n_1 \geq \bar{n}_1(n_2, \varepsilon, \kappa) = \frac{1}{2} \left(-1 + \sqrt{1 - 4n_2 + 12n_2^2}\right)$$

With the generalized contest success function, when players are homogeneous, any
network formations with two cliques are Nash equilibrium if and only if \( n_1 \geq \frac{1}{2}(-1 + \sqrt{1 - 4n_2 + 12n_2^2}) \).

The condition is equivalent to

\[
n_1 \geq \frac{1}{2} \left( 3n - \sqrt{3n^2 + 2n} \right)
\]

by using \( n_1 + n_2 = n \). It is easy to check that \( \frac{2n}{3} \) is an upper bound of the RHS. So \( n_1 \geq \frac{2n}{3} \) is a sufficient condition for \( g \) to be a Nash equilibrium which people are easy to memorize.