Social Clubs and Social Networks

Chaim Fershtman[1] Dotan Persitz[2]

[1]Department of Economics
Tel Aviv University

[2]Coller School of Management
Tel Aviv University

April 10, 2018

The Common Strategic Network Formation Setup

- Agents need to decide whether they form links with other agents (e.g. Jackson and Wolinsky (1996) and Bala and Goyal (2000)).
- Each link is costly.
- Benefits from direct links as well as indirect links.
- Common assumption: every pair of agents can potentially link.

Social Environments

- In modern life most of the social contacts are formed within social context.
- Examples - family, gym, neighborhood, alumni, department, conference, interest group, workplace, scouts, army unit, synagogues, churches.
- Within the club, agents may connect with one another and form their social ties.
- Rivera et al. (2010, p. 106): “If networks are the fabric of inter-personal interaction, social foci are the looms in which they are woven”.
- We suggest a model of **simultaneous** formation of the social network and the social clubs’ system.

Our View

Literature

- Links are formed in a context (club).
- The clubs have certain characteristics: e.g. acceptance and leaving policies, formation rules, size effects, participation costs (time and fees).
- The characteristics of the club determine the existence and quality of the links.
- Agents choose an affiliation portfolio but care about their position in the induced network.
- Hence, club memberships and network formation are determined simultaneously.
- The goal - provide insights on network formation that cannot be captured by the standard strategic link formation models.

The General Story

- Agents choose clubs.
  - Membership is a decision of the candidate only.
  - Membership is costly.
  - An agent can have multiple memberships.
- Two agents that share a club are connected by a weighted link.
- The weight is determined by some function:
  - Main Model: Club congestion function.
  - Second Model: Individual congestion function.
- Indirect connections are depreciated (endogenous depreciation).
- The agents benefit from their position in the induced weighted network.
- We are interested in the stable and efficient environments.

An Environment

- $N = \{1, ..., n_a\}$ is a finite set of identical agents.
- $S = \{1, ..., n_s\}$ is a finite set of identical clubs.
- The pair $\{i, s\}$ implies that individual i is affiliated with club s.
- $A^c \equiv \{\{i, s\} : i \in N, s \in S\}$ is the set of all possible affiliations.
- An environment is a triplet $G \equiv < N, S, A >$ where $A \subseteq A^c$.
- Equivalent representations: bipartite graphs and hyper-graphs.
- Sociologists refer to an environment as an “Affiliation Network” (starting with Davis et al. (1941)).

Additional Notation

- Individual i's affiliations: $S_G(i) \equiv \{s \in S | \{i,s\} \in A\}$ ($s_G(i) \equiv |S_G(i)|$).
- Club s's members: $N_G(s) \equiv \{i \in N | \{i,s\} \in A\}$ ($n_G(s) \equiv |N_G(s)|$).
- Additional affiliation: $G + \{i,s\} \equiv < N, S, A \cup \{\{i,s\}\} >$ ($G - \{i,s\} \equiv < N, S, A \backslash \{\{i,s\}\} >$).
- Additional club: let $m \subseteq N$ and let s be a vacant club, $G + m \equiv < N, S, A \cup \bigcup_{i \in m} \{\{i,s\}\} >$ (we assume that one vacant club always exists).
- $\mathcal{G}_n$ is the set of all environments with n agents.

The Induced Network

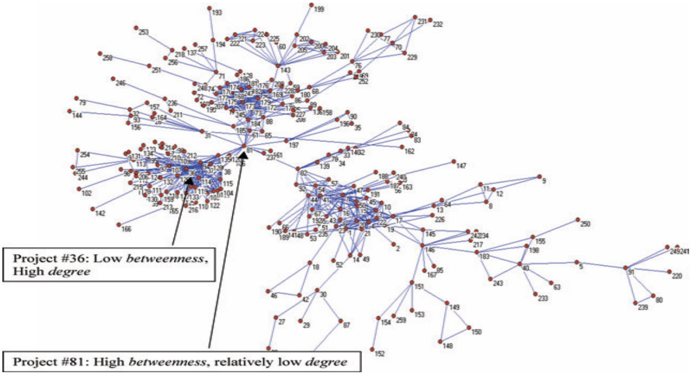
- We are interested in the underlying network where the nodes are the agents.
- Potentially, we could have also analyzed the underlying clubs' network:
  - The nodes are the clubs and two clubs are connected if they have mutual members.
  - However, clubs have no objective function in our setting.
- Two agents are linked if they share a club.
- The quality of a link may depend on various characteristics of the environment.

Weights

- The weight of a link between two agents $i, i' \in N$ in G is denoted by $w(i, i', G) \in [0, 1]$.
- We will have detailed specifications of weights that are derived from club congestion and individual congestion in the upcoming analysis.
- The weighted network $g =< N, E_G, W_{G,w} >$ is induced by Environment G and weighting function w if $E_G \equiv \{\{i, j\} | i \in N, j \in N, S_G(i) \cap S_G(j) \neq \emptyset\}$ and $\forall \{i, j\} \in E_G : W_{G,w}(\{i, j\}) \equiv w(i, j, G)$.

Multiple Affiliation

Multiple affiliation is a necessary condition for indirect connections. From Fershtman and Gandal (2011):



Indirect Connections

- A path between i and i' in g is a non-empty subgraph p of g where the nodes are $\{x_1, x_2, x_3, \ldots, x_{l-1}, x_l\}$ (all distinct) and the edges are $\{x_1 x_2, x_2 x_3, \ldots, x_{l-1} x_l\}$, $x_1 = i$ and $x_l = i'$.

Definition (The Weight of a Path)

The weight of path $p = \{x_1, \ldots, x_l\}$ in the induced weighted network g is $WP_g(p) = \prod_{k=1}^{l-1} W(\{x_k, x_{k+1}\})$.

Distance

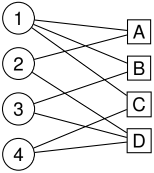
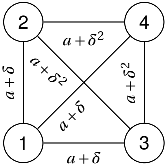
- Path p is a *shortest weighted path* between agents i and i' if there is no path p' between agents i and i' such that $WP_g(p') > WP_g(p)$.
  - The shortest path between two agents may be indirect even if they share a club in G (impossible in most network formation models).
- $d(i, i'|G, w)$ denotes the weight of a shortest path between agents i and i' in the induced weighted network g.
- $d(i, i'|G, w) = 0$ if there is no path between agents i and i'.

Preferences

- The agent benefits from her position in the network.
- Let $c > 0$ denote the homogeneous participation fees.
- The utility of Agent i from the Environment G and the weighting function w is:

$$u_i(G, w, c) = \sum_{k \in N, k \neq i} d(i, k | G, w) - s_G(i) \times c$$

Example

<table>
<tr><th>Environment</th><th>List of Clubs</th><th>Agents' Network</th><th>Utilities</th></tr>
<tr>
<td></td>
<td>Club A: 1 2
Club B: 1 3
Club C: 1 4
Club D: 2 3 4</td>
<td></td>
<td>$u_1 = 3(a+\delta) - 3c$
$\forall i \in \{2,3,4\}$:
if $\delta \geq \frac{1-a}{2}$:
$u_i = (a+\delta) + 2(a+\delta)^2 - 2c$
Otherwise:
$u_i = (a+\delta) + 2(a+\delta^2) - 2c$</td>
</tr>
</table>

The weighting function: $w(i, i', G)$ equals $a + \delta^{m-1}$ where m is the size of the smallest club i and i' share, $\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$.

Open Clubwise Stability

Open Clubwise Stability

An environment G is Open Clubwise Stable (OCS) for weighting function w and membership fees c if:

$$\forall s \in S, \forall i \in N_G(s) : u_i(G, w, c) \geq u_i(G - \{i, s\}, w, c) \qquad \text{(No Leaving)}$$

$$\forall s \in S, \forall i \notin N_G(s) : u_i(G, w, c) \geq u_i(G + \{i, s\}, w, c) \qquad \text{(No Joining)}$$

$$\forall m \subseteq N : \exists i \in m : u_i(G + m, w, c) > u_i(G, w, c) \Rightarrow \qquad \text{(No New Club)}$$

$$\exists j \in m : u_j(G + m, w, c) < u_j(G, w, c)$$

Efficiency

- An environment G is ***Pareto Efficient (PE)*** if there is no other environment G' such that $\forall i \in N : u_i(G', w, c) \geq u_i(G, w, c)$ and $\exists j \in N : u_j(G', w, c) > u_j(G, w, c)$.
- An environment G is ***Strongly Efficient (SE)*** if there is no other environment G' such that $\sum_{i \in N} u_i(G', w, c) > \sum_{i \in N} u_i(G, w, c)$.
- If Environment G is strongly efficient, it is also Pareto efficient, but the opposite is not necessarily correct.

Common Environments

- **The Empty Environment**: $G = < N, S, \emptyset >$.
- **The Grand Club**: Exactly one populated club and all the agents are affiliated with it.
- **The All Pairs Environment**: Every pair of agents shares a unique club of size two.

No Weights, No Membership Fees

- Assumptions:
  - $w(i,j,G)$ is identically 1.
  - $c = 0$.

Definition (Connected Environment)

G is a Connected Environment if its induced network is connected (any pair of agents have a path between them).

- Environment G is OCS if and only if G is connected.
  - Two disconnected agents benefit at least 1 each from forming a club together (and suffer no costs).
  - Two connected agents (directly or indirectly) cannot improve by shortening the path between them.
- The connected environments are also efficient.

No Weights, Positive Membership Fees

- $w(i, j, G)$ is identically 1.
- $c > 0$.

Definition (Minimally Connected)

$G = < N, S, A >$ is a Minimally Connected environment if

1. The induced network is connected.
2. For every affiliation $\{i, s\} \in A$, the network induced by $G - \{i, s\}$ is disconnected.

Initial Characterization

1. If G is OCS and $c < n_a - 1$ then G is a Minimally Connected environment.
2. If $c > n_a - 1$ the Empty environment is the unique OCS environment.

Intuition

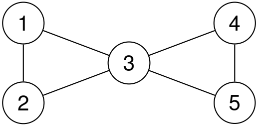
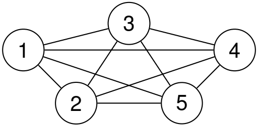
Intuition (for the case $c < n_a - 1$):

- The Grand Club environment is OCS.
- The empty environment is not OCS.
- Suppose G is OCS and disconnected (but not empty).
- There must be a component H that contains $h > 1$ agents.
- The maximal possible utility of an agent in H is $(h-1)-c$.
- Since G is OCS then $c < h-1$.
- However, any agent that is not in H can improve by $h-c > 0$ if she joins any one of H's clubs. Contradiction.
- Therefore, if G is OCS then it is connected.
- Suppose G is OCS, connected, but not minimally connected.
- There is an agent that wants to leave a club since it will not affect the network connectivity (and her benefits).

Further Classification

- Let $G =< N, S, A >$ be a Minimally Connected environment.
- Let $\{i, s\} \in A$.
- $G - \{i, s\}$ includes two components.
- Denote by $C_{-i}(G - \{i, s\})$ the component that does not include Agent i.
- Denote by $c_{-i}(G - \{i, s\})$ the number of agents in $C_{-i}(G - \{i, s\})$.
- The Class of G is $K(G) = \min_{\{i,s\} \in A} c_{-i}(G - \{i, s\})$.
- Intuition: The environment's "weakest affiliation" is an affiliation that its absence leads to the minimal loss. This loss is the environment's class.

Examples

List of Clubs	Induced Network	Class
Club A: 1 2 Club B: 2 3 Club C: 3 4 Club D: 4 5		K(G)=1 Individual 2 leaves Club A Individual 4 leaves Club D
Club A: 1 2 3 Club B: 3 4 5		K(G)=2 Individual 3 leaves Club A Individual 3 leaves Club B
Club A: 1 2 3 4 5		K(G)=4 Every individual that leaves Club A

Complete Characterization

Proposition

Suppose that for every environment G and for every pair of agents i and j that share a club in G, $w(i,j,G)=1$. Then,

1. If $n_a - 1 > c > 0$:
    1. G is OCS if and only if G is a Minimally Connected Environment of class $K(G) \geq c$.
    2. The Grand Club Environment is the unique PE and SE environment.
2. If $c > n_a - 1$, the Empty Environment is the unique OCS, PE and SE environment.

The Club Congestion Model

McPherson and Smith-Lovin (1982) p. 884

One aspect of voluntary associations which is particularly crucial for the network of informal relations is the size of a given organization. Large organizations generate more potential acquaintances. One could argue that the contacts which occur in a larger organization are more "superficial" than those in smaller organizations.

Definition (The Club Congestion Model)

A club congestion function is a non-increasing function $h:\{2,3,\dots,n_a\}\longrightarrow[0,1]$.
Given a club congestion function h, the weight of a link between two agents $i,i'\in N$ is $w_h(i,i',G)=\max_{s\in S_G(i)\cap S_G(i')} h(n_G(s))$.

No Membership Fees

Assumptions:

- $1 > h(2) > h(3)$.
- $c = 0$.

Definition (Sub Environment)

Let $G =< N, S, A >$ and $G' =< N', S', A' >$ such that $S' \subseteq S$, $N' = \cup_{s \in S'} N_G(s)$ and $A' = \{\{i, s\} | i \in N', s \in S', \{i, s\} \in A\}$ then G' is a sub environment of G and G is a super environment of G'.
If, in addition, $N' = N$ then G' is a spanning sub environment of G and G is a spanning super environment of G'.

Characterization

Club congestion with no membership fees

The only OCS environments are the spanning super environments of the All Pairs environment. These are also the only environments which are efficient (SE and PE).

Two agents that do not share a size 2 club benefit from forming such club together.

- The original value of their connection is at most $\max\{h(3), h^2(2)\}$.
- Since $1 > h(2) > h(3)$ they both strictly benefit.
- No other deviations are worthwhile (over-affiliation provides no value and no costs).

The Connections Model

- Jackson and Wolinsky (1996).
- Let g be an unweighted network.
- $u_i^{JW}(g) = \sum_{j \neq i} \delta^{d_{ij}} - n_i(g) \times c$, where
  - d_{ij} is the length of the shortest path between Agents i and j.
  - $0 \leq \delta \leq 1$ is the depreciation factor.
  - $c > 0$ is the universal direct connection costs.
  - $n_i(g)$ is the number of Agent i's direct neighbors.
- g is pairwise stable if no agent wishes to discard a link and no pair of agents wants to form a link.

The Connections Model as a Club Congestion Model

- Denote by $PS(\delta, c, n)$ the set of pairwise stable networks in the connections model and by $OCS(c, n, h)$ the set of OCS environments in the club congestion model.
- For every un-weighted network $\bar{g} =< N, \bar{E} >$ the corresponding environment $G_{\bar{g}} =< N, S, A >$ is such that for each link $\{i, j\} \in \bar{E}$ there exists a club $s_{ij} \in S$ that includes only agents i and j, and there are no other populated clubs.
- $\mathbb{G}_n$ is the set of all un-weighted networks with n agents.
- $\mathcal{G}_{\bar{\mathbb{G}}n} \subseteq \mathcal{G}_n$ is the set of all corresponding environments.

The Connections model is embedded in the Club Congestion model

Let $h(2) = \delta$ and $\forall m > 2 : h(m) = 0$.

1. $\bar{g} \in PS(\delta, c, n)$ if and only if $G_{\bar{g}} \in OCS(c, n, h)$.
2. If $G \in \mathcal{G}_n \backslash \mathcal{G}_{\bar{\mathbb{G}}n}$ then $G \notin OCS(c, n, h)$.

Two Useful Environments

Let $S_p = \{s | n_G(s) > 0\}$ denote the set of populated clubs.

Definition (m-complete)

G is an m-complete environment ($m \in \mathbb{N}$, $n_a \geq m \geq 2$) if:

$$\forall i, i' \in N : |S_G(i) \cap S_G(i')| = 1.$$

$$\forall s \in S_p : n_G(s) = m$$

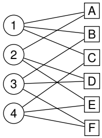
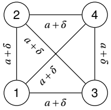
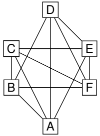
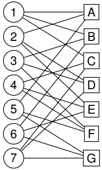
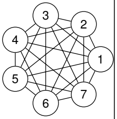
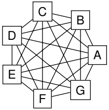
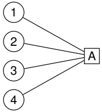
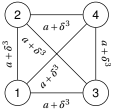
Definition (m-star)

G is an m-star environment ($m \in \mathbb{N}$, $n_a \geq m \geq 2$) if:

$$\forall s \in S_p : n_G(s) = m$$

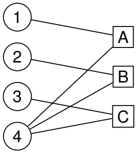
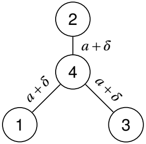
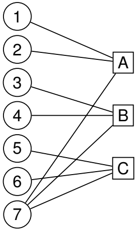
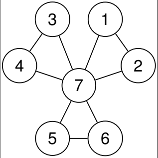
$$\forall s, s' \in S_p : N_G(s) \cap N_G(s') = \{i\}, i \in N.$$

m-complete

	Environment	List of Clubs	Network	Clubs' Network	Utilities
All Pairs or 2-Complete (n=4)		Club A: 1 2 Club B: 1 3 Club C: 1 4 Club D: 2 3 Club E: 2 4 Club F: 3 4			$\forall i \in \{1,2,3,4\}$: $u_i = 3(a+\delta) - 3c$
3-Complete (n=7)		Club A: 1 2 5 Club B: 1 3 6 Club C: 1 4 7 Club D: 2 3 7 Club E: 2 4 6 Club F: 3 4 5 Club G: 5 6 7	 The weights are all $a+\delta^2$		$\forall i \in \{1,\ldots,7\}$: $u_i = 6(a+\delta^2) - 3c$
Grand Club or 4-Complete (n=4)		Club A: 1 2 3 4		A	$\forall i \in \{1,2,3,4\}$: $u_i = 3(a+\delta^3) - c$

The weighting function: $w(i, i', G)$ equals $a + \delta^{m-1}$ where m is the size of the smallest club i and i' share, $\delta \in (0,1)$, $a \in [0,1)$ and $a + \delta \in (0,1)$.

m-star

	Environment	List of Clubs	Network	Utilities
2-Star (n=4)		Club A: 1 4 Club B: 2 4 Club C: 3 4		$\forall i \in \{1,2,3\}$: $u_i = (a+\delta)+2(a+\delta)^2 - c$ $u_4 = 3(a+\delta) - 3c$
3-Star (n=7)		Club A: 1 2 7 Club B: 3 4 7 Club C: 5 6 7	 The weights are all $a + \delta^2$	$\forall i \in \{1,\ldots,6\}$: $u_i = 2(a+\delta^2)+4(a+\delta^2)^2 - c$ $u_7 = 6(a+\delta^2) - 3c$

The weighting function: $w(i,i',G)$ equals $a+\delta^{m-1}$ where m is the size of the smallest club i and i' share, $\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$.

Efficiency

- Efficiency is hard.
- We characterize the efficient networks among those with homogeneous club size.

Definition (m-uniform environment)

G is an m**-uniform environment** ($m \in \{2, \ldots, n_a\}$) if
$\forall s \in S : n_G(s) = m \quad \text{or} \quad n_G(s) = 0$.

- Denote the set of all m-Uniform environments with n agents by $\mathcal{G}_n^m$.
- Denote the set of all Uniform environments with n agents by $\mathcal{G}_n^{all} = \cup_{k=2}^{n_a} \mathcal{G}_n^k$.

Efficiency Result

Efficienct m-Uniform environments

Let $m \in \{2, \dots, n_a\}$. For every club congestion function $h(\cdot)$ and m-Uniform Environment $G' \in \mathcal{G}_n^m$:

1. $c \in [0, (m-1)(h(m) - h^2(m)))$ and let G be an m-Complete Environment. Then, $\sum_{i=1}^{n_a} u_i(G, h, c) \geq \sum_{i=1}^{n_a} u_i(G', h, c)$.
2. $c \in ((m-1)[h(m) - h^2(m)], (m-1)h(m) + \frac{(n_a - m)(m-1)}{m} h^2(m)]$ and let G be an m-Star Environment. Then, $\sum_{i=1}^{n_a} u_i(G, h, c) \geq \sum_{i=1}^{n_a} u_i(G', h, c)$.
3. $c \geq (m-1)h(m) + \frac{(n_a - m)(m-1)}{m} h^2(m)$ and let G be the Empty Environment. Then, $\sum_{i=1}^{n_a} u_i(G, h, c) \geq \sum_{i=1}^{n_a} u_i(G', h, c)$.

This result provides motivation to study when are those architectures stable.

Club Congestion Functions

Graphs

Definition (Reciprocal Club Congestion)

$\forall m \geq 2 : h(m) = \frac{1}{m-1}$

Definition (Exponential Club Congestion)

$h(m) = a + \delta^{m-1}$ where $\delta \in (0,1)$, $a \in [0,1)$ and $a + \delta \in (0,1)$

Interpretation:

- Reciprocal: in each club, every agent divides her unit of attention uniformly between all other club members.
- Exponential: the sum of a decreasing exponent that stands for the prospects of a potential link to materialize and a constant that represents the role of the club as an institution that connects agents (the ideational component of solidarity in Moody and White (2003)).

Club Size Elasticity

Exponential with $a = 0$

- Direct Club Value (DCV): $k_h(m) = (m-1) \times h(m)$, the value of the direct links induced by a club of size m.
- Quick implication: the Empty environment is OCS if and only if $c \geq \max_{m \in \{2,\dots,n_a\}} k_h(m)$.
- Club-size elasticity of $h(m)$: $\eta_h(m) \equiv \frac{\frac{h(m+1)-h(m)}{h(m)}}{\frac{1}{m}}$ for every club size m where $h(m) > 0$ and $\eta_h(m) \equiv 0$ otherwise.
- If $\forall m \in \{2,\dots,n_a - 1\} : \eta_h(m) > -1$ ($\eta_h(m) < -1$) then $h(m)$ is said to be inelastic (elastic).

Lemma (Club Size Elasticity)

The club congestion function $h(m)$ is inelastic (elastic) if and only if $k_h(m)$ is strictly increasing (decreasing).

Stability of m-complete

Stability of m-complete

Let $m \in \mathbb{N}$, $n_a > m \geq 2$. Denote by $\hat{k}$ the club size that maximizes the DCV. The m-complete environment is OCS if and only if

$$c \in \Big[\max_{k \in \{2,\ldots,\min\{m-1,\hat{k}\}\}} (k-1)[h(k)-h(m)], (m-1)[h(m)-h^2(m)] \Big]$$

If $m = n_a$ the m-complete environment (the Grand Club) is OCS if and only if

$$c \in \Big[\max_{k \in \{2,\ldots,\min\{n_a-1,\hat{k}\}\}} (k-1)[h(k)-h(n_a)], (n_a-1)h(n_a) \Big].$$

The existence of such membership fees is not guaranteed.

Low Membership Fees

- Recall: Spanning super environments of the All Pairs environment are the only OCS when $c = 0$.
- Stability of All Pairs: the only sensible deviation is replacing a club membership with an indirect connection.
- The All Pairs environment is OCS if and only if $h(2) - h^2(2) \geq c > 0$.
- Uniqueness: environments where the smallest club between some pair of agents is of size greater than 2 are not OCS if forming a new club is worthwhile.
- The All Pairs environment is the unique OCS if and only if

$$\min\left\{h(2) - h^2(2), h(2) - h(3)\right\} \geq c > 0$$

- Efficiency: The All Pairs is the unique OCS if and only if it is PE and SE.

Grand Club

- Recall: The Grand Club environment is OCS and the unique efficient environment when there is no congestion.
- Two types of possible deviations:
  - A single agent leaves the club and gets 0.
  - A subset of agents form a smaller new club.
- The existence of membership fees where the Grand Club is OCS is not guaranteed.
  - Inelastic club congestion function: such fees exist.
  - Reciprocal club congestion function: $c \in [1 - \frac{1}{n_a - 1}, 1]$.
  - Exponential club congestion function where $\delta \in (0, \frac{1}{2})$: if $a = 0$ such fees never exist but if $a > 0$, there exists an $\bar{n}_a$ such that $\forall n_a : n_a > \bar{n}_a$, such fees exist.

Non Monotonicity

Claim

Let $h(\cdot)$ be an exponential club congestion function where $\delta \in (0, \frac{1}{2})$ and $a > 0$. There exist two integers $\bar{m} \leq \tilde{m}$ such that $\forall m : n_a > m > \bar{m}$ there exists a range of membership fees in which the m-complete environment is OCS. Moreover, there exists a range of membership fees in which every m-complete environment where $n > m > \tilde{m}$ is OCS.

- Thus, for such club congestion functions, m-complete environments with either small or big clubs are OCS for some membership fees, while no such fees exist for similar environments with intermediate size clubs.
- Example: let $a = \frac{1}{32}$ and $\delta = \frac{1}{4}$.
  - All pairs is OCS in $[0, \frac{3}{16}]$.
  - For $m \in \{3, \ldots, 9\}$ the m-complete is never OCS.
  - For $m \geq 10$ every m-complete is OCS in [0.25,0.27].

Stability of the m-star

Proposition Incentives Numeric

- The general conditions for an m-star to be OCS are cumbersome.
- Results mainly from the multiple options for coalitional deviation.
- The attractive deviation involves only peripheral agents and it depends on the severity of congestion versus the severity of the indirect connections:
  - If congestion is the main issue, then a new small club is attractive.
  - If indirect connections are the main issue, then a new big club is attractive.
  - In any case diversification is crucial.

The 2-star

- Agents in the 2-Star Environment suffer no congestion.
- One relevant deviation is a new club formed by peripheral agents.
- The other relevant deviation is dropping affiliation by the central agent.

Stability of the 2-star Environment

Denote $l_h = \min\{k \in \mathbb{Z} | h(k) \leq h^2(2)\}$.

1. The 2-Star Environment is OCS if and only if

$$h(2) \geq c \geq \max_{k \in \{2,\dots,\min\{l_h - 1, n_a - 1\}\}} (k-1)(h(k) - h^2(2))$$

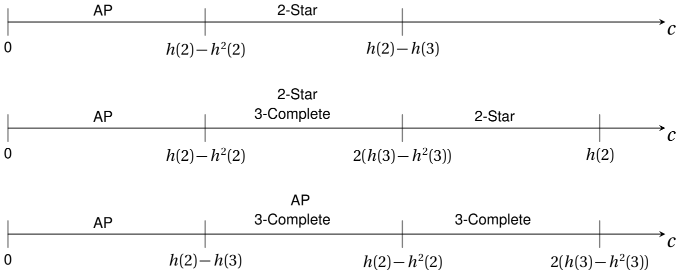
2. Let $h(\cdot)$ be an elastic club congestion function. The 2-Star Environment is OCS if and only if $h(2) \geq c \geq h(2) - h^2(2)$.

Congestion vs. Depreciation

Congestion is the stronger friction - $h(3) < h^2(2)$:

Depreciation and congestion are equal - $h(3) = h^2(2)$:

Depreciation is the stronger friction - $h(3) > h^2(2)$:



Insight

- The standard strategic theory of network formation restricts clubs to be of size 2.
- Therefore, the main trade-off is costly direct connections versus depreciated indirect connections.
- In particular, cliques can occur only if the linking costs are low.
- We endogenize the club size.
- We argue that this trade-off includes another force - congestion.
- For example, cliques can occur in high costs environment if congestion is less of a friction compared to indirect connection depreciation.

3-star

- These environments are hybrids.

3-star

Let $n_a \geq 9$ and let $h(\cdot)$ be the exponential club congestion function with $a = 0$. The 3-Star Environment is OCS if and only if $c \in [\delta + \delta^3 - 2\delta^4, 2\delta^2]$. This range exists if and only if $\delta \geq \frac{1}{2}$.

- When the 3-star is never OCS its because either:
  - The membership fees are too high, the central agent would leave the congested clubs.
  - Or, the membership fees are too low, two peripheral agents that do not share a club gain by forming a size 2 club:
    - Shorter path to the other peripheral agent.
    - Solve both congestion and indirect connections

The Individual Congestion Model

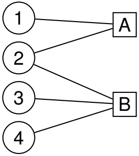
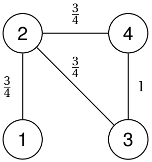
- The observation: Club affiliations require attention and time.
- Agents with a thin portfolio of affiliations are able to pay attention to each of their memberships and to form high quality connections with other members.
- Agents who are members of many clubs, possess many weak direct relations since they devote little attention to each of their memberships.

Definition (The Individual Congestion Model)

An individual congestion function is a non-increasing function $b : \{2, 3, \ldots, n_s\} \rightarrow [0, 1]$.
Given an individual congestion function b, the weight of a link between two agents $i, i' \in N$ in Environment G is,

$$w_b(i, i', G) = b(s_G(i)) \times b(s_G(i'))$$

Example

Environment	List of Clubs	Frictions	Weighted Network	Utilities
	Club A: 1 2 Club B: 2 3 4	$b(1) = 1$ $b(2) = \frac{3}{4}$ $b(3) = \frac{2}{3}$ $c = \frac{1}{4}$		$u_1 = \frac{3}{4} + 2 \times \frac{9}{16} - \frac{1}{4} = 1\frac{5}{8}$ $u_2 = 3 \times \frac{3}{4} - \frac{1}{2} = 1\frac{3}{4}$ $u_3 = \frac{3}{4} + 1 + \frac{9}{16} - \frac{1}{4} = 2\frac{1}{16}$ $u_4 = \frac{3}{4} + 1 + \frac{9}{16} - \frac{1}{4} = 2\frac{1}{16}$

Main Result

Stability and Efficiency

In the Individual Congestion model where $b(1) > b(2) > 0$:

1. Suppose $c \in [0, (n_a - 1)b^2(1))$:
    1. The Grand Club environment is the unique SE and PE environment.
    2. The Grand Club environment is OCS.
2. Let G be a non Grand Club environment.
    1. If G is OCS and non-empty then the Grand Club environment is OCS.
    2. For every $c \in [0, (n_a - 1)b^2(1))$ there exists an individual congestion function such that G is not OCS while the Grand Club environment is OCS.
3. Suppose $c > (n_a - 1)b^2(1)$:
    1. The Empty environment is the unique SE and PE environment.
    2. The Empty environment is the unique OCS.

Huge Multiplicity

Calculation

- The general intuition: joining existing or new clubs are unattractive deviations (individual congestion and membership costs).
- Suppose $c=0$ and $b(1)>b(2)$.
- The Grand Club environment is OCS.
- Consider the individual congestion function $b(k)$ such that $b(1)=1$, $b(2)=\frac{3}{4}$ and $\forall k>2: b(k)=0$.
- The circle with $n\geq 4$ agents (n clubs of 2 agents each) is also OCS.
- Intuition 1: Due to individual congestion, joining an additional (existing or new) club harms existing connections.
- Intuition 2: Due to depreciation, reducing individual congestion does not compensate for the loss of a short path.

m-complete

Stability of m-complete

Let $m \in \mathbb{N}$, $n_a > m \geq 2$ and denote $\gamma \equiv \frac{n_a-1}{m-1}$. An m-complete environment is OCS if and only if

$$c \in \left[0, (m-1)b^2(\gamma)[1-b(\gamma-1)b(\gamma)] - (n_a-m)b(\gamma)[b(\gamma-1)-b(\gamma)]\right]$$

- Key observation: A necessary condition for a formation of a new club to be an attractive deviation, is that it provides each member with at least one new direct neighbor.
- Therefore, in m-Complete environments the formation of new clubs is never an attractive deviation.
- For similar reasons no agent wishes to join an existing club.
- Thus, an m-Complete environment is OCS if the indirect connections friction induced by leaving a club is not compensated by the improved quality of the direct connections and the reduced membership fees.

The Hybrid Model

Definition (The Hybrid Model)

Given a club congestion function $h(\cdot)$ and an individual congestion function $b(\cdot)$, the weight of a link between two agents $i, i' \in N$ in the weighted network g induced by Environment $G =< N, S, A >$ is

$$w_{hb}(i,i',G) = b(s_G(i)) \times b(s_G(i')) \times \max_{s \in S_G(i) \cap S_G(i')} h(n_G(s))$$

The Co-Authors Model

- Jackson and Wolinsky (1996).
- Each agent distributes one unit of attention equally between her direct relations.
- The value of each relation depends only on the attention devoted to the link by the two end agents.
- $n_i(g)$ is the number of Agent i's direct neighbors.
- $u_i^{CA}(g) = \sum_{j \neq i : \{i,j\} \in g} [\frac{1}{n_i(g)} + \frac{1}{n_j(g)} + \frac{1}{n_i(g) n_j(g)}]$.
- g is pairwise stable if no agent wishes to discard a link and no pair of agents wants to form a link.
- Denote by $CA(n)$ the set of pairwise stable networks with n agents.

The Co-Authors Model as a Hybrid Model

- A D-Truncated model: A value from a path $p = \{x_1, \dots, x_l\}$ is the multiplication of the weights on its links if $l-1 \leq D$ and zero otherwise.
- $OCS(c, n, h, b, D)$ is the set of OCS environments with n agents where the club congestion function is h, the individual congestion function is b, the truncation parameter is D and the membership fees are c.
- We use the previous definitions for the comparison of un-weighted networks and environments.

The Co-Authors Model is embedded in the 1-Truncated Hybrid model

Let $h(2) = 1$, $\forall m > 2 : h(m) = 0$, $b(k) = \frac{1}{2}[1 + \frac{1}{k}]$.

1. $\bar{g} \in CA(n)$ if and only if $G_{\bar{g}} \in OCS(\frac{1}{4}, n, h, b, 1)$.
2. If $G \in \mathscr{G}_n \setminus \mathscr{G}_{\bar{\mathbb{G}}n}$ then $G \notin OCS(\frac{1}{4}, n, h, b, 1)$.

Individual Congestion vs. Club Congestion

- We use the Grand Club to demonstrate the interplay between the various frictions.
- GC is OCS if and only if

$$(n_a - 1) \times b^2(1) \times h(n_a) \geq c \geq$$
$$\max\left\{0, \max_{k \in \{2,\ldots,n_a-1\}} \{(k-1) \times [h(k) \times b^2(2) - h(n_a) \times b^2(1)] - (n_a - k) \times h(n_a) \times b(1) \times [b(1) - b(2)]\}\right\}$$

- The deviating agents become members of two clubs.
- The quality of their links with non-deviating agents decrease (incentive to deviate to large clubs).

Alternative Club Rules

- Very specific club rules are implicit in the model.
- Open Clubwise Stability assumes that membership is open for all agents (as long as the participation fees are paid).
- Alternative club acceptance rules:
  - Membership requires approval of (a subset of) the existing club members.
  - Membership quotas.
  - Membership criteria.
  - Exclusivity rules.
- Other implicit specifications - leaving rules, rules for forming new clubs, coordination (within and between agents).
- Assumption: all clubs in a specific environment have the same, exogenously given, rules.

Closed Clubwise Stability

Closed Clubwise Stability

An Environment G is Closed Clubwise Stable if the following conditions hold:

1. No Leaving:
$\forall s \in S, \forall i \in N_G(s): u_i(G,w,c) \geq u_i(G - \{i,s\}, w, c)$.
2. No New Club Formation:
$\forall m \subseteq N: \exists i \in m: u_i(G+m,w,c) > u_i(G,w,c) \Rightarrow$
$\exists j \in m: u_j(G+m,w,c) < u_j(G,w,c)$.
3. No Joining: $\forall s \in S, \forall i \notin N_G(s)$:
$u_i(G,w,c) \geq u_i(G+\{i,s\},w,c) \quad OR$
$\exists j \in N_G(s): u_j(G,w,c) > u_j(G+\{i,s\},w,c)$.

- OCS implies CCS.
- When there is no congestion, no negative externalities on the incumbents from admitting a new member.
- Therefore, when there is no congestion OCS and CCS coincide.

Closed Clubwise Stability in the Club Congestion Model

- When there are no membership fees OCS and CCS coincide.
- G is the Almost Grand Club environment if there is exactly one populated club and all but one agent are affiliated with it.
- Assume $h(\cdot)$ be such that $k_h(n_a - 1) > k_h(n_a) > \max_{k \in \{2,\ldots,n_a-2\}} k_h(k)$
- Let $k_h(n_a) > c > \max_{k \in \{2,\ldots,n_a-2\}} k_h(k)$ and $n_a > 3$.
- No agent wants to leave the populated club.
- No subset of agents wants to form a new club.
- The Almost Grand Club environment is not OCS since the isolated agent wishes to join the populated club.
- But, it is CCS since such a deviation will strictly hurt the incumbents.
- For the incumbents one new direct connection does not compensate for the weaker direct connections to all other incumbents due to stronger club congestion.
- Implication: Different club rules may induce different dynamics.

Closed Clubwise Stability in the Individual Congestion Model

- Let G be the $\frac{n_a+1}{2}$-Star environment and denote the central individual by i and the two populated clubs by s and t.
- Suppose n_a is odd, $b(k) = \frac{1}{k+1}$ and $c = \frac{1}{18}$.
- No agent wants to leave a club (G is minimally connected and c is low enough).
- No subset of agents wants to form a new club (the damage to existing direct connections is substantial).
- The $\frac{n_a+1}{2}$-Star environment is not OCS since every agent $j \neq i$ wishes to join the other club.
- But, it is CCS since Agent i opposes such an admission (her connection with Agent j deteriorates).
- Under non-unanimous majority rule, G is not CCS.

Clustering

- Real life networks are highly clustered: The probability of two individuals who share a common neighbor to be connected is much higher than expected if connections were random.
- The literature attributes the high clustering to one of two explanations:
  - “Preference for transitivity”: Attraction is based on the “network” properties of the individuals.
  - “Homophily”: Attraction is based on “non-network” properties of the individuals.
- Recent literature attempts to provide econometric tools for estimating network formation models that incorporate these explanations. Mainly concerned with homophily on unobservables (e.g. Goldsmith-Pinkham and Imbens (2013), Mele (2017), Graham (2015, 2016))
- Another concern is that neglecting to account for self-selection into social contexts leads to an over-estimation of the importance of these factors (Rivera et al. (2010) and Miyauchi (2016)).

Our Model and Clustering

- Indeed, our setting provides a third explanation.
- In our framework, a network must exhibit high clustering since the individual's neighbors form a tightly knit group.
- we propose clubs as linking platforms rather than individuals' linking preferences as the fundamental that drives high clustering.
- Requires a dataset where the social network (e.g. friendships) and the affiliation information (e.g. social clubs memberships) were gathered independently.
- We hypothesize that the cliques identified in the social network, should be traced back to the clubs' affiliations information (indirect evidence in Kossinets and Watts (2006)).

Other Environments

- R&D partnerships, joint ventures etc. (the projects are clubs of firms).
- Interlocking directorates (e.g. Mintz and Schwartz (1981), a director is a club of firms and a board of a firm is a club of directors).
- Standardization committees (Bar and Leiponen (2014) and Leiponen (2008), a committee is a club of firms).
- Open source code development (Fershtman and Gandal (2011), a project is a club of developers and a developer is a club of projects).
- Trade (a trade agreement is a club of countries).

Take Home

- Novel framework for strategic network formation of undirected weighted networks.
- Agents choose affiliations and benefit from their position in the underlying network.
- The framework provides insights that are absent from the link formation models.
- Empirical implications: clustering.

Thanks


THANK
YOU
GRACIAS
ARIGATO
SHUKURIA
DANKSCHEEN
BİYAN
SHUKRIA
TINGKI
TASHAKKUR ATU
YAQHANYELAY
SUKSAMA
EKHMET
GRAZIE
MEHRBANI
PALDIES
MAAKE
KOMAPSUMNIDA
GOZAIMASHITA
EFCHARISTO
JUSPAXAR
BOLZÏN
MERCI

Clubwise Stability of m-star

Proposition

Let $h(\cdot)$ be a club congestion function. Denote:

- $\gamma \equiv \frac{n_a-1}{m-1}$, $\eta_k \equiv \lceil \frac{k}{\gamma} \rceil$, $l_h = \min\{k \in \mathbb{Z} | h(k) \leq h^2(m)\}$.
- $J_h(m) = (m-1)[h(m+1) - h^2(m)]$.
- $FNS_h(k,m,n_a) = (k-1)[h(k) + (m-2)h(k)h(m) - (m-1)h^2(m)]$.
- $FNI_h(k,m,n_a) =$ $(k-1)h(k) - (\eta_k - 1)h(m) + (n_a - m - (k - \eta_k))h(m)h(k) - (n_a - m)h^2(m)$.
- $FNL_h(k,m,n_a) = (k - \eta_k)(h(k) - h^2(m))$.

1. If $\gamma \geq m$ the m-star environment is OCS if and only if

$$k_h(m) \geq c \geq \max\{\max_{m \geq k \geq 2} FNS_h(k,m,n_a), \max_{\min\{l_h, n_a\} > k > m} FNL_h(k,m,n_a)\}$$

2. If $\gamma < m$ the m-star environment is OCS if and only if

$$k_h(m) \geq c \geq \max\{J_h(m), \max_{\gamma \geq k \geq 2} FNS_h(k,m,n_a), \max_{m \geq k > \gamma} FNI_h(k,m,n_a)$$

Clubwise Stability of the m-star

back

- Leaving a club:
  - Lost connections cannot be replaced.
  - The damage is less severe for the central agent.
  - Hence, the membership fees should be low enough for the central agent to keep all her affiliations.
- Joining an existing club:
  - Irrelevant for the central agent.
  - For a standard agent it replaces $m-1$ indirect connections with $m-1$ direct connections (congested and costly).
  - The membership fees should be high enough for the peripheral agents to refrain from joining existing clubs.

Clubwise Stability of the m-star

back

- Forming a new club:
  - A standard agent gains more than the central agent.
    - If the new club is smaller than the original, the central agent gains only from the direct connections while the others also gain from increasing the value of their indirect paths.
    - Otherwise, the central agent gains nothing while the others may gain from the direct links.
  - Members of other clubs are always better partners.
    - If the new club is smaller than the original size, sharing the new club with a new partner increases the weights of all the indirect paths to other members of her club.
  - The membership fees should be high enough to deter the agents from coordinating on new clubs.

m-star - Exponential congestion

back

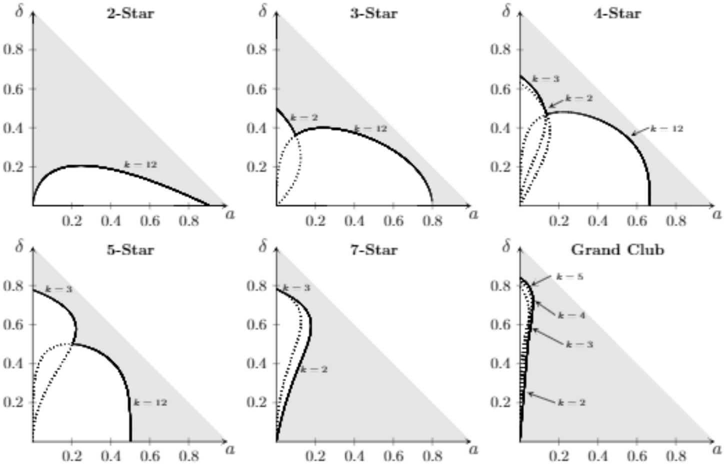


Figure 8: The existence of membership fees for which m-star environments are OCS when the club congestion function is exponential and $n_a = 13$.

m-star - Exponential congestion - continued

back

- The reason for the non-monotonicity lies in the complicated incentive structure.
- When the a is low the relative importance of congestion is high.
  - Hence, the effective lower bound is induced by a deviation of a small club, that improves on the congestion.
- When the a is high the relative importance of congestion is low.
  - Hence, the effective lower bound is induced by a deviation of a large (well diversified) club, that improves on the indirect connections.
- In the 2-star such a transition does not exist since there are no congestion issues.

Very Brief Literature Orientation

back

- Sociological Theory:
  - Simmel (1908/1955) - The Web of Group Affiliations.
  - Kadushin (1966) - Social Circles.
- Economic Theory:
  - Strategic Formation of Social Networks (surveys in Jackson (2008) and Goyal (2007)).
  - Games over Environments: Hsieh et al. (2015).
  - Contractual Stability: Caulier et al. (2013b, 2015) and Caulier et al. (2013a).
  - General Preferences: Mauleon et al. (2015).
  - Multi-Links: So et al. (2014) and Jun and Kim (2009).

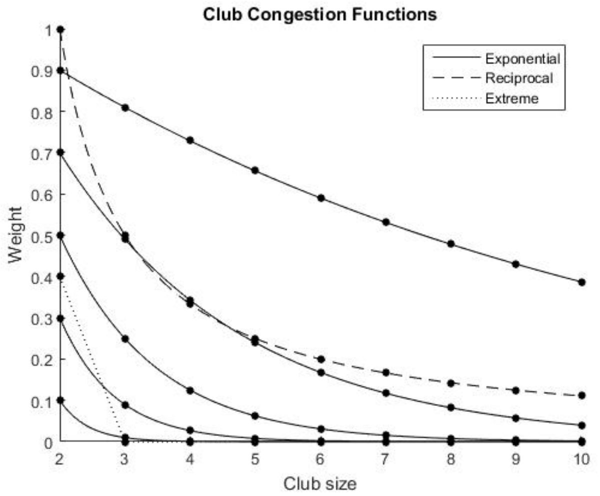
Individual Congestion Only

Back

- Consider a circle with 4 agents (4 clubs of 2 agents each) when $b(1) = 1$, $b(2) = \frac{3}{4}$ and $\forall k > 2 : b(k) = 0$.
- The weight of each link is $(\frac{3}{4})^2$.
- Therefore, each agent's utility is $2 \times (\frac{3}{4})^2 + (\frac{3}{4})^4 > 1.44$.
- The utility of an agent that leaves one of her clubs (and becomes the end of a line) is $\frac{3}{4} + (\frac{3}{4})^3 + (\frac{3}{4})^5 < 1.41$.
- Also, no agent wants to join an existing club or form a new club since $b(3) = 0$.
- Thus, the circle is OCS.

Club Congestion Functions

Back



Elasticity of Exponential Club Congestion Functions ($a = 0$)

Back