Social Clubs and Social Networks*

Chaim Fershtman† and Dotan Persitz‡

April 10, 2018

Abstract

We present a strategic network formation model which is based on membership in clubs. Agents choose a set of clubs with which they wish to be affiliated. The set of all club memberships (an environment) induces a weighted network in which two agents are directly connected if they are members of the same club. Two agents may also be indirectly connected using the multiple memberships of third parties. Agents gain from their position in the induced network and pay membership fees. Thus, both clubs and the network are formed simultaneously. Using two specifications of the weighting function we introduce two models based upon congestion - one, the club congestion model wherein the weight of each link depends upon the size of the smallest shared club and the other, the individual congestion model wherein each link's weight depends on the number of affiliations maintained by the two agents. In the club congestion model we focus on the trade-off between the size of the club, depreciation due to indirect connections and membership fees. In the individual congestion model the Grand Club environment is the unique efficient environment. However, a coordination failure arises due to the wide externalities incurred by the formation of new affiliations. We believe that this framework may serve as a basis for an empirical examination of the role of linking platforms in shaping real-life social networks.
Keywords: Strategic Formation of Clubs, Strategic Formation of Weighted Undirected Networks.
JEL Classification: D71, D85, Z13.

1 Introduction

Most of the initial social interactions between individuals occur within social circles, social groups or social clubs.[1] Clearly, some social connections can be formed randomly - like meeting someone on the street - but most friendships and acquaintances are formed within a social context like a family, school class, alumni organization, church, fraternity, academic

*We benefited from conversations with Francis Bloch, Andrea Galeotti, Li Hao, Nils Röhl, Konrad Stahl and the participants of the 15^{th} and 19^{th} CTN Workshops, the workshop in honor of Ken Judd on the occasion of his 60th birthday and of seminars at the University of California Berkeley, University of British Columbia, University of Toronto, Brown University, Paris School of Economics, the University of Hamburg, Erasmus University Rotterdam, London Business School, University of Pittsburgh, IBM research labs, the Technion, Ben Gurion University and Tel Aviv University. We thank Zvika Messing, Amit Dekel and Omri Puny for excellent research assistance. Dotan Persitz acknowledges the financial support of the Israel Science Foundation (grant number 1390/14) and the Henry Crown Institute of Business Research in Israel.

†The Eitan Berglas School of Economics, Tel Aviv University, Tel Aviv, 69978, ISRAEL and CEPR. Email: fersht@post.tau.ac.il.

‡Coller School of Management, Tel Aviv University, Tel Aviv, 69978, ISRAEL. Email: persitzd@post.tau.ac.il.

[1] Sociologists refer to social contexts as social foci. Feld (1981) introduces a "focus theory" where he defines social foci as "Social, psychological, legal, or physical entities around which joint activities are organized." For the sociological literature see, Simmel (1908/1955), Young and Larson (1965a,b), Kadushin (1966, 2011), Feld (1981), Granovetter (1983), Blau and Schwartz (1984) and the survey on non-geographical proximity by Rivera et al. (2010). See also the discussion on sub-neighborhoods in Jackson et al. (2012).

department, research group, workplace, boy scouts, youth extracurricular activities, gym or even at a bar that the individual regularly attends.[2] That is, social links are typically formed within social contexts rather than between individuals who do not share any common social foci. Thus, when considering the formation of social networks, the social environment, particularly, the number and size of the different social clubs and the type of affiliations that people maintain within these clubs also needs to be scrutinized.

Most sociologists view social clubs as preceding the formation of social networks, as stated in 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." In some social clubs, membership is automatic (e.g. family), but in most cases affiliation is by choice. People choose their gym, their university, their place of worship, as well as other social clubs that they wish to belong to, taking into account their existing club affiliations and the structure of their social network. Our work focuses on the strategic choice of an affiliation portfolio along with the simultaneous formation of club environments and social networks.[3]

Membership in a social club provides the benefit of being directly connected to other individuals in the club. Multiple club affiliations facilitate indirect connections between individuals who have no clubs in common. So, for example, an individual may have direct connections to her high school class mates in addition to having an indirect connection to an individual that attends a reading club together with one of her high school class mates.

Interaction in a small club is different than that in a large club. The "quality" of connection between two individuals generated in a large club, tends to be lower than that generated by a small club. In a small group, members are well acquainted and the flow of information is more reliable. Intuitively, the probability of any pair of members interacting and realizing the potential benefit from their mutual affiliation decreases with the size of the club. McPherson and Smith-Lovin (1982) show that the number of clubs with which an individual is affiliated is not gender dependent. However, as men tend to belong to much larger clubs than do women (see also Maccoby (1998)) gender differences in various aspects of social life can be attributed either to the larger number of direct contacts formed by men or to the higher quality of direct contacts cultivated by women.[4]

---

[2]Clearly direct benefits accrue from belonging to a club, like the positive health effects of training in a gym or the religious public goods provided by institutions of worship. These benefits are the focus of the well-established literature on club theory (see Tiebout (1956) and Buchanan (1965) for seminal contributions and Sandler and Tschirhart (1997) and Scotchmer (2002) for detailed surveys). In this paper we abstract from these benefits and focus on the role of clubs as platforms for the formation of social contacts.

[3]For sociological work that advocate for the simultaneous evolution of social networks and social foci see Feld (1981) and McPherson et al. (2001). Snijders et al. (2006) and Chandrasekhar and Jackson (2017) introduce stochastic non-strategic models of network formation that admit exogenously given clubs as platforms for link formation.

[4]There is some concrete evidence for the role of participation in social clubs in future economic outcomes. For example, Persico et al. (2004) find that participation in athletics clubs, social clubs and social activities in adolescence accounts for about half of the teen height wage premium (see also Moody (2001)). However, they are unable to point out "what precisely is acquired" in these clubs (p. 1050).

Since Granovetter (1973) the concept of "weak ties" has become central to the applied literature on social networks.[5] There are two possible interpretations of "weak ties." Two individuals may be connected directly via a large club that is subject to heavy congestion or through an interconnected sequence of small clubs. In real-life both types of weak ties are observed and their relative importance depends upon the context. This paper highlights the trade-off between the two types of weak links - indirect connections composed of "high quality" links and direct "low quality" links.

We present a model where agents choose affiliations in social clubs. Two individuals who share a club are linked in the induced social network. Each link is assigned a weight which is a non-increasing function of the size of the smallest club shared by the two agents. The weight of an indirect connection is the product of the weights associated with the links along the path. We define the shortest path between two agents as the highest quality connection between them. The benefits to an agent are the sum of the shortest paths to all other agents net of the total club membership fees. A social environment is Open Clubwise Stable (OCS) if no agent wants to leave or join a club and there is no subset of agents that are better off by forming a new club together.[6]

Open Clubwise Stability can be viewed as an extension of the pairwise stability solution concept posited by Jackson and Wolinsky (1996) to the club formation setup. Indeed, we show that the connections model with the pairwise stability solution concept, is a special case of our framework with open clubwise stability. That is, the connections model is equivalent to a club formation model with the restriction that clubs consist of exactly two members. Therefore, the setup of social clubs can be viewed as a generalization of the setup of social networks where instead of a link that connects an individual to one other individual, affiliation in a social club provides an individual with links to a group of individuals.

[5] Weak ties appear in two branches of the literature on networks in Economics. In the literature on the role of networks in labor markets, weak ties are typically viewed as cheap, infrequently used direct links that may relay useful job information (e.g. Boorman (1975), Montgomery (1992, 1994), Boxman and Flap (2000), Calvó-Armengol and Zenou (2003) and Kramarz and Skans (2014)). Calvó-Armengol (2004) studies job information transmission through indirect links but do not refer to those channels as weak ties. In the literature on the formation of weighted networks (which is frequently motivated by Granovetter (1973)) the weight is determined endogenously as some function of investments made by both end agents. A general model is introduced in Bloch and Dutta (2009) and studied further by Deroïan (2009), So (2016), Salonen (2015, 2016) and Baumann (2017). A similar approach is taken by Goyal (2005) and Goyal et al. (2008) to study R&D cooperation, by Brueckner (2006) to develop a network formation model of friendship (see also Currarini et al. (2009)) and by Rogers (2006) to explore giving and asking over networks. Another approach is to model resource allocation as a subsequent stage to the formation of the network (e.g. Ballester et al. (2006) and Cabrales et al. (2011)). Altogether, this literature also interprets weak ties as direct links (with low weights) and does not refer to indirect connections as weak ties. For a survey of the recent sociological literature on weak ties see Aral (2016).

[6] While we highlight clubs as platforms on which links form, other works concentrate on the role of the coalition as a binding agreement that constrains player activities. Myerson (1980), Slikker and Van den Nouweland (2001) and Arnold and Wooders (2005) study cooperative games with an exogenously given collection of communication subsets of players ("conferences"). In Caulier et al. (2013a,b, 2015) a coalitional network is a pair of an unweighted network and a partition. They introduce a solution concept where deviations require the consent of the original coalitions with which the deviators are affiliated.

When membership costs are sufficiently low, the environment wherein every pair of individuals share a single club of size two is the unique stable environment as the complete network is induced with high quality links. When membership costs are higher, this environment is no longer sustainable and "weak ties" appear. We show that when congestion friction is higher than depreciation friction a stable environment is one based on "weak ties" of the indirect connections type. But, contrary to most of the existing literature on strategic network formation, when depreciation is more significant than congestion, a stable environment is one based on "weak ties" generated in large clubs. In particular, our model predicts that complete networks can survive high membership costs if congestion is weaker than endogenous depreciation.

The trade-off between congestion and indirect connections is further demonstrated by considering the following two special environments: m-complete and m-star. In m-complete environments every pair of agents shares exactly one club that includes m members, and therefore every pair of individuals is directly connected via a congested link (unless $m = 2$). In m-star environments, one individual (the "star") is affiliated with all the populated clubs, the other agents (the "peripherals") are members of a single club and all the populated clubs are of size m. Therefore, in m-star environments, every peripheral agent is directly connected to $m-1$ agents and indirectly connected to all the rest. We show that when membership fees are low the efficient environment among the environments where all populated clubs are of size m, is the m-complete environment. The m-star environment is efficient for intermediate affiliation fees while the empty environment is efficient for sufficiently high membership costs.

We demonstrate that the stability of the various m-complete and m-star environments can be characterized as a function of the elasticity of congestion relative to club size. There is, however, non-monotonicity in the relationship between congestion and the size of clubs in stable environments. For a substantial set of congestion functions, m-complete environments with intermediate size clubs are never stable while m-complete environments with either small clubs (wherein each individual maintains many high quality affiliations) or large clubs (wherein each individual maintains few low quality affiliations) are open clubwise stable.

A different, yet important, type of congestion is individual congestion. When individuals belong to several clubs, time constraints or limited attention may reduce the effectiveness of each affiliation. That is, the larger the set of affiliations an individual maintains, the lower the quality of the links generated by these affiliations. Upon joining a club individuals pay participation fees but it also decreases the attention that can be devoted to other affiliations. To capture this effect we introduce a non-increasing function that assigns a value to affiliation portfolios of every size. The weight of each link is the product of the values assigned to the portfolio size of the end-agents. Clearly, in such a setting the Grand Club environment, wherein all agents are members of the same club is the unique efficient environment. However, there are many inefficient stable environments that emerge due to a coordination failure.

In real-life, both individual and club congestion exist. We briefly study such a model. In particular, we show that the co-authors model with the pairwise stability solution concept (Jackson and Wolinsky (1996)), is a special case of our richer model with open clubwise stability.

In real-life a wide range of rules regarding the formation, the joining or the leaving of social clubs are observed. For example, clubs may introduce entry barriers wherein acceptance by incumbent members is required in order to join the club.[7] Each set of rules induces a different set of possible deviations and therefore corresponds to a different stability concept. This clearly affects individual choices of clubs and consequently the stable environments. The solution concept of open clubwise stability represents an open environment wherein individuals are free to join or leave any club they wish (one at a time), and to form new clubs, as long as they pay a fixed membership fee. To demonstrate the importance of club rules we introduce the Closed Clubwise Stability (CCS) solution concept which is a strictly weaker concept wherein joining a club requires the unanimous approval of all existing club members. We demonstrate that these two concepts may lead to dramatically different stable environments.

Establishing social connections via clubs yields different types of social networks than those formed in the regular framework of network formation. In particular, the club setting provides an alternative explanation for the extensive clustering that characterizes real-life social networks. In most real-life networks the probability of two individuals being connected if they are linked with a common individual is much higher than if the connections were formed randomly (see Goyal (2007) and Jackson (2008)). Social science literature frequently attributes the high clustering in social networks to one of two explanations: First, individuals may have a preference for connections with individuals with whom they share a neighbor (preference for transitivity). Second, individuals may prefer to link to individuals with whom they share social traits (homophily). We argue that simultaneous formation of clubs and networks provides a third explanation for the high clustering observed in real-life networks. This explanation is based on the mechanics of link formation rather than on specific assumptions regarding preferences over links.

2 The Model

An *environment* is a group of agents and a set of clubs such that each agent is affiliated with a subset of clubs. Formally, $N = \{1, ..., n_a\}$ $(n_a > 2)$ is a finite set of agents and $S = \{1, ..., n_s\}$ is a finite set of clubs. The pair $\{i, s\}$ denotes the affiliation of Agent i with Club s and $A^c \equiv \big\{\{i, s\} : i \in N, s \in S\big\}$ is the set of all possible affiliations. An environment is the triplet $G \equiv < N, S, A >$ where $A \subseteq A^c$ is a set of affiliations. We denote the set of

[7] Also, there can be some interdependence between choices of clubs, in particular when belonging to one club may restrict entrance to other clubs (e.g. membership in a local antisemitic club probably restricts membership in the neighborhood's synagogue and vice versa).

all the environments with n agents by $\mathcal{G}_n$.[8] We denote by $S_G(i) \equiv \{s \in S | \{i, s\} \in A\}$ the set of clubs that Agent i is affiliated with in Environment G, and $s_G(i) \equiv |S_G(i)|$ denotes its cardinality. In addition, we denote by $N_G(s) \equiv \{i \in N | \{i, s\} \in A\}$ the set of agents that are affiliated with Club s in Environment G, and $n_G(s) \equiv |N_G(s)|$ denotes its cardinality.

The environment that results from adding (severing) $\{i, s\}$ to (from) Environment G is denoted by $G + \{i, s\} \equiv < N, S, A \cup \big\{\{i, s\}\big\} >$ (similarly, $G - \{i, s\} \equiv < N, S, A \backslash \big\{\{i, s\}\big\} >$). Let $s \in S$ be a vacant club (we assume that such a club always exists) and let $m \subseteq N$. Then, $G + m \equiv < N, S, A \cup \bigcup_{i \in m} \big\{\{i, s\}\big\} >$ is the environment that emerges from Environment G when the set m of agents populates the vacant club s.

Consider two environments $G =< N, S, A >$ and $G' =< N', S', A' >$. If $S' \subseteq S$, $N' = \cup_{s \in S'} N_G(s)$ 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'.

Every Environment G induces an undirected network g whose nodes represent agents and two agents are linked in g if they belong to the same club. We denote the weight of a link between two agents $i, i' \in N$ in G by $w(i, i', G) \in [0, 1]$. In this general setting the weight can be a function of the whole environment. We detail the specifications of the weights derived from club congestion and individual congestion in the upcoming analysis. Formally, a weighted network is a triplet $< N, E, W >$ wherein N is a set of agents, E a set of links and $W : E \to [0, 1]$ the set of weights. 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 $\forall \{i, j\} \in E_G : W_{G,w}(\{i, j\}) \equiv w(i, j, G)$. Note that each environment has another induced undirected network whose nodes represent clubs and two clubs are linked if there is an agent that affiliates with both. The club network should be the focus of an analysis of a setting wherein the clubs are strategic.[9]

We assume that agents benefit from being connected, either directly or indirectly, to other agents. Multiple affiliations facilitate indirect connections between individuals who have no clubs in common. Indirect connection between a pair of agents occurs whenever a third party shares a club with each of the two agents (see, for example, Faust (1997)). Formally, a *path* of length $l-1$ between Agent i and Agent i' in the induced network g is a sequence of agents $\{x_1, x_2, \ldots, x_{l-1}, x_l\}$ such that $x_1 = i$ and $x_l = i'$ and every consecutive pair of agents, x_k and x_{k+1}, shares at least one club in G. Two agents who share at least one club are directly

[8] A graph $G =< V, E >$ is called *bipartite* or *two mode network* if V admits a partition into two classes $(U, V \backslash U)$ such that $\forall (v_1, v_2) \in E : v_1 \in U, v_2 \in V \backslash U$. An environment can be described as a bipartite graph wherein one set of nodes is the set of agents and the other is the set of clubs. This representation is often referred to as an *affiliation network* (e.g. Ch. 8 in Wasserman and Faust (1994), Newman (2001), Bonacich et al. (2004), Latapy et al. (2008), Borgatti and Everett (2013) and Renoust et al. (2014)). An additional way to represent an environment is by a hypergraph. A *hypergraph* is a pair $H =< U, ME >$ wherein the elements of ME are subsets of U. This representation is mainly used in Mathematical Graph Theory (see Berge (1989)).

[9] Fershtman and Gandal (2011) take advantage of this duality to study the open source environment. For a methodological sociological analysis see Bonacich (1972, 1978) and Breiger (1974).

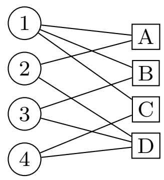
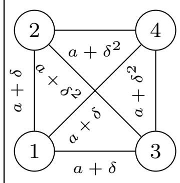
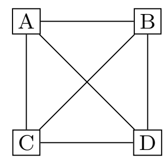
Environment	**List of Clubs**	**Induced Network**	**Club Network**	**Utilities**
	Club A: 1 2 Club B: 1 3 Club C: 1 4 Club D: 2 3 4			$u_1 = 3(a+\delta) - 3c$ $\forall i \in \{2,3,4\}$**:** **if** $\delta \geq \frac{1-a}{2}$**:** $u_i = (a+\delta) + 2(a+\delta)^2 - 2c$ **Otherwise:** $u_i = (a+\delta) + 2(a+\delta^2) - 2c$

Figure 1: An environment, its induced weighted agent network, the induced club network and the agents' utilities. The weighting function assigns a weight of $a+\delta$ to links that originate from a club of size two and $a+\delta^2$ to links that originate from a club of size three ($\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$).

connected and two agents who do not share a club in Environment G are indirectly connected if there is a path between them in g. If every pair of agents is connected (either directly or indirectly) then g is connected; otherwise, it is disconnected. We say that Environment G is connected if its induced network g is connected. The sub environment $G' =< N', S', A' >$ of $G =< N, S, A >$ is a component of G if its induced network g' is connected and there is no pair of agents $i \in N'$ and $k \in N \backslash N'$ who share a club in G. We denote the size of the component G' by $n(G') = |N'|$.

The *weight of a path* is the product of the weights on the links that constitute this path. That is, let $g =< N, E, W >$ be a weighted network. The weight of the path $p = \{x_1, \ldots, x_l\}$ is $WP_g(p) = \prod_{k=1}^{l-1} W(\{x_k, x_{k+1}\})$.[10] Path p is a *shortest weighted path* between agents i and i' if and only if there is no path p' between agents i and i' such that $WP_g(p') > WP_g(p)$.

The *distance* between agents i and i' in G using weighting function w, denoted $d(i, i'|G, w)$, is the weight of the shortest weighted path between them in the induced network g. If there is no such path, $d(i, i'|G, w) = 0$.

Agents benefit from short distances to other agents. The utility of Agent i is given by $u_i(G, w, c) = \sum_{k \in N, k \neq i} d(i, k|G, w) - s_G(i) \times c$ where c is the fixed membership fees.
Figure 1 provides a simple example. The two leftmost cells describe an environment containing four agents and four populated clubs wherein Agent 1 shares a club of size two with each of the other three agents who, among themselves share an additional club of size three. Therefore, the induced weighted network is the complete network depicted in the next cell. The weights on the links are such that a club of size two provides a link of strength $a+\delta$ while a club of size three provides a weaker link of strength $a+\delta^2$ ($\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$). Finally, the utilities of the agents are documented in the rightmost column. As this example demonstrates, the distance measure used in this model is different from the geodesic distance (the length of the path with the minimal number of links) that is used in

[10] This definition resembles the definition of the reliability of a path in Bloch and Dutta (2009) (see also Brueckner (2006)).

most models of unweighted network formation since the shortest distance is not necessarily the path that includes the least number of links.

Stability: Environment G is ***Open Clubwise Stable*** (henceforth, OCS) if no individual strictly gains from leaving a club, no individual strictly gains from joining a club and there is no subset of individuals who are all better off by forming a new club together. Formally, the conditions for OCS are:

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

(ii) No New Club Formation: $\forall m \subseteq N$:
$\exists i \in m : u_i(G + m, w, c) > u_i(G, w, c) \Rightarrow \exists j \in m : u_j(G + m, w, c) < u_j(G, w, c)$.

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

Efficiency: Environment G is ***Pareto Efficient*** (henceforth, 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 $\exists j \in N : u_j(G', w, c) > u_j(G, w, c)$. Environment G is ***Strongly Efficient*** (henceforth, SE) if there is no other environment G' such that $\sum_{i \in N} u_i(G', w, c) > \sum_{i \in N} u_i(G, w, c)$. Obviously, if Environment G is strongly efficient, it is also Pareto efficient, but the opposite is not necessarily true.

The following environments are instrumental in characterizing stable and efficient environments in this setting.

(i) **The Empty Environment**: $G = < N, S, \emptyset >$ is the Empty environment.
(ii) **The Grand Club**: G is the Grand Club environment if there is exactly one populated club and all the agents are affiliated with it.
(iii) **The All Paired Environment**: Environment G is the All Paired environment if every pair of agents shares a unique club of size two.

3 Baseline Model: No Congestion

We start by considering the simple setting in which all weights are set to 1 ($w(i, j, G)$ is identically 1), implying that the distance between two agents is 1 if they are (either directly or indirectly) connected and 0 otherwise.

When there are no membership fees, agents wish to maximize the number of other agents with which they are connected, either directly or indirectly. In this case it is hardly surprising that every connected environment is both stable and efficient.[11]

[11]For every connected environment G, $\forall i \in N : u_i = n_a - 1$ which is the maximal attainable utility. Therefore, G is OCS, SE and PE. On the other hand, if environment H is not connected, there is at least one pair of individuals who have no path between them. By forming a club together, the benefit of both increases and therefore, H is not OCS. In addition, since any other agent is not worse off by the formation of such club, H is neither PE nor SE.

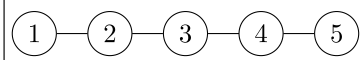
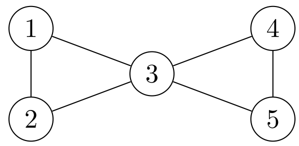
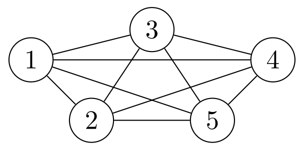
Environment	Induced Network	Class ("Weakest Affiliation")
Club A: 1 2 Club B: 2 3 Club C: 3 4 Club D: 4 5		K(G)=1 Individual 2 leaves Club A Individual 4 leaves Club D
Club A: 1 2 3 Club B: 3 4 5		K(G)=2 Individual 3 leaves Club A Individual 3 leaves Club B
Club A: 1 2 3 4 5		K(G)=4 Every individual that leaves Club A

Figure 2: Three Minimally Connected environments of 5 agents, their induced networks and their classes.

For the case of positive membership fees, we say that environment G is **Minimally Connected** if it is connected and for every affiliation $\{i,s\} \in A$, the network induced by $G - \{i,s\}$ is disconnected.

Lemma 1. *Let G be a Minimally Connected environment where Agent i is affiliated with Club s. Then, $G - \{i,s\}$ is a disconnected environment that contains exactly two components.*

All the proofs are relegated to Appendix A. Lemma 1 shows that when a single affiliation is removed from Minimally Connected environment G two components emerge - one that contains Agent i, denoted $C_i(G - \{i,s\})$, and one that does not contain Agent i, denoted $C_{-i}(G - \{i,s\})$. In the setting with no congestion, the size of $C_{-i}(G - \{i,s\})$ is the loss incurred by Agent i upon canceling the affiliation with Club s. We say that the "weakest affiliation" in Environment G is the one whose absence leads to the smallest $C_{-i}(G - \{i,s\})$. We classify the minimally connected environments by their "weakest affiliation," $K(G) = \min_{\{i,s\} \in A} n(C_{-i}(G - \{i,s\}))$. Figure 2 demonstrates this classification on some minimally connected environments that contain 5 agents.

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

1. *If $n_a - 1 > c > 0$:*
   - (a) *G is a Minimally Connected environment of class $K(G) \geq c$ if and only if G is OCS.*
   - (b) *The Grand Club 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 intuition behind Proposition 1.1a is: First note that for $n_a - 1 > c > 0$ the Grand Club Environment is OCS while the Empty Environment is not. Hence, if G is OCS and disconnected there must be a component H that contains $n_a > h > 1$ agents. Since the maximal possible utility of an agent in H is $(h-1)-c$ and since G is OCS then $c < h-1$. But then it is beneficial for every agent who is not included in H to join any one of H's clubs. Therefore, if G is OCS then it is connected. But, if it is not minimally connected there is an agent who may want to leave a club since leaving will not affect network connectivity (i.e. the agent's benefits). Finally, if the membership costs are higher than the value of the "Weakest Affiliation" there will be agents who may wish to cancel one of their affiliations.[12]

4 The Club Congestion Model

The quality of the connections generated within a club may depend on the size of the club. If club membership is high, the "quality" of the connection between any two members is probably lower than the "quality" of the connection between any two members of a small club. For example, consider the difference between belonging to a club of five individuals who attended the same college together versus being a member of the club of the class of '87 at a high school with over two hundreds members (see also Feld (1981), McPherson and Smith-Lovin (1982) and Rivera et al. (2010)). In this model we capture "quality" by assuming that links are weighted and that the weight of each link depends upon the size of the club shared by the agents. Furthermore, we assume that when agents share more than one club, the weight of the link between them is determined by the congestion in the smallest club that they share.[13] Specifically,

The club congestion function is a non-increasing function $h : \{2, 3, \ldots, n_a\} \to [0, 1]$.

Given club congestion function h, the weight of a link between two agents $i, i' \in N$ is $w_h(i, i', G) = \max_{s \in S_G(i) \cap S_G(j)} h(n_G(s))$.

Even without congestion the affiliations of one agent may affect the social network of other agents. Unilateral actions such as leaving a club or joining a club may benefit or harm other agents by creating new links (either direct or indirect) or by "breaking" some of the shortest paths. Incorporating congestion into the club formation setting introduces a new type of externality whereby these unilateral actions may also affect the quality of some links. For example, if Agent j joins a club with which Agent i is also affiliated, the quality of some links that Agent i maintains may change - either by making some paths shorter or by reducing the weight of some links due to stronger congestion. While this externality does not affect an agent's decision either to join or leave a club (Agent j in the example) it clearly affects the social desirability of the new environment.

[12]Bar (2005) also considers a model of strategic formation that includes club structure. However, she ignores any type of congestion and therefore her results are comparable to our Proposition 1 with some minor differences. Similar approaches where taken by Jun and Kim (2009) and So et al. (2015). A different approach, inspired by two-way flow model of Bala and Goyal (2000), was taken by Borgs et al. (2011) where agents organize gatherings to which they set the invitees and for which they bear the costs.

[13]See, Breiger (1974) for a discussion on networks in which the weights depend on the number of clubs that the agents share (see also Christakis et al. (2010)).

While unilateral actions may have positive or negative externalities, the formation of a new club can never hurt uninvolved agents. Hence, Lemma 2 implies that every environment such that a subset of agents can improve by forming a new club, cannot be efficient and, hence, there is a well-defined subset of the non-OCS environments that can never be efficient.

Lemma 2. *If Environment G is Pareto Efficient then it satisfies the condition of "No New Club Formation".*

Our analysis of club congestion begins by examining the simple case in which there are no membership fees. If $1 > h(2) > h(3)$, the only OCS environments are the spanning super environments of the All Paired Environment. These are also the only efficient (SE and PE) environments.[14] That is, in any efficient OCS environment every pair of agents must share a club of size two. Therefore, the induced network is complete and the weights on the links are the highest possible since club congestion is at its minimum.

We now turn to consider the club congestion model with positive membership fees. An agent holds at most $n_a - 1$ links in the induced network and this network includes at most $\frac{n_a(n_a-1)}{2}$ links. Since each link is determined by a single club - the smallest club the two agents share - we can establish an upper bound to the number of affiliations per individual and the number of populated clubs in an OCS environment assuming affiliations are costly.

Lemma 3. *Suppose $c > 0$. If Environment $G =< N, S, A >$ is OCS then:*

1. $\forall i \in N : s_G(i) \leq n_a - 1$.
2. $|\{s \in S | n_G(s) > 0\}| \leq \frac{n_a(n_a-1)}{2}$.

Note that the number of possible clubs in an environment with n_a agents is $2^{n_a} - (n_a + 1)$. Therefore, Lemma 3.2 implies that OCS environments in the club congestion model include relatively few populated clubs (e.g. for 10 agents there are 1013 possible clubs, but an OCS environment includes at most 45 populated clubs).[15]

The set of OCS environments in the club congestion model with positive membership fees crucially depends on the properties of the congestion function. We therefore introduce two forms of club congestion, Reciprocal Club Congestion and Exponential Club Congestion. These will be useful in demonstrating some of the results in the upcoming analysis.

Reciprocal Club Congestion: Environment G is characterized by Reciprocal Club Congestion if $\forall m \geq 2 : h(m) = \frac{1}{m-1}$.

[14]Suppose Environment G is a spanning super environment of the All Paired Environment. Then $\forall i \in N : u_i = (n_a - 1) \times h(2)$. This is the maximal utility that can be attained in this model. Therefore, G is OCS, SE and PE. Next, suppose that Environment G is not a spanning super environment of the All Paired Environment. Then there is at least one pair of agents (Agent i and Agent i') who share no club of size 2. By forming a club of size 2, their benefits will increase in at least $h(2) - \max\{h(3), h^2(2)\}$ which is positive since $1 > h(2) > h(3)$. Therefore, G is not OCS. By Lemma 2, G is not PE and therefore it is also not SE.

[15]Lemma 3 is true for every model in our framework where the weight of the link is determined by a single club and the solution concept includes the condition of "No Leaving". We use Lemma 3 in the Matlab code package that accompanies this work (see Section 8).

Exponential Club Congestion: Environment G is characterized by Exponential Club Congestion if $h(m) = a + \delta^{m-1}$ where $\delta \in (0, 1)$, $a \in [0, 1)$ and $a + \delta \in (0, 1)$.

The Reciprocal Club Congestion function can be interpreted as one unit of attention that agents in a club uniformly lavish upon the other club members. The Exponential Club Congestion function is the sum of two components: The first, representing the role of the club as an institution that connects agents is a constant denoted by a, and therefore depends only on agents' mutual affiliation[16] and the second, that can be interpreted as the prospects of a potential link materializing, is an exponential function that decreases with the size of the club, δ^{m-1} ($\delta \in (0, 1)$).

When agents are affiliated with a club of size m they enjoy $m - 1$ direct links to other club members. We define $k_h(m)$ as the Direct Club Value (henceforth, DCV) such that $k_h(m) = (m-1) \times h(m)$. The size of club, m, has two effects on $k_h(m)$. While a bigger club generates more direct connections, these links are of lower quality due to club congestion.

The reciprocal club congestion function is a special case of the two effects of club size on the DCV cancelling each other out as $k_h(m)$ is equal to 1 independently of m. Intuitively, for a club member, the direct value of a club is exactly the unit of attention collected from other members. The DCV of the exponential congestion function depends on a and δ. When $a = 0$, it can be shown that when $\delta < \frac{1}{2}$ the congestion effect is dominant and the DCV is maximized when the club is small ($m = 2$), but when $\delta > \frac{1}{2}$ a higher value of δ implies that the DCV is maximized by a larger value of m. When $a > 0$, the effect of the number of links is reinforced since the aggregate benefit of a increases linearly with m.[17]

To demonstrate the role of the DCV consider the Empty Environment. The Empty Environment always satisfies both the conditions of "No Leaving" and "No Joining". Therefore, the Empty Environment is OCS if and only if the condition of "No New Club Formation" holds.[18] Notice that the benefit of an agent from participating in the formation of a new club of size m is exactly the DCV of this club, $k_h(m)$. Therefore, the Empty Environment with n_a agents is OCS if and only if $c \geq \max\limits_{m \in \{2,\dots,n_a\}} k_h(m)$. Proposition B.1.1 in Appendix B.2 shows that if the congestion function is reciprocal, the Empty Environment is OCS if and only if $c \geq 1$. Proposition B.1.2 characterizes the conditions for the Empty Environment to be OCS when the congestion function is exponential.

Lemma 4 below connects the strategic aspects captured by the DCV to the properties of the club congestion function. We define the club-size elasticity of the club congestion function h as $\eta_h(m) \equiv \frac{\frac{h(m+1)-h(m)}{h(m)}}{\frac{1}{m}}$ for every club size m where $h(m) > 0$ and $\eta_h(m) \equiv 0$ otherwise. $h(m)$ is non-negative and non-increasing and, therefore, $\eta_h(m) \leq 0$. We say that $h(m)$ is

[16]In the sociological context Moody and White (2003) refer to it as the ideational component of solidarity. See also the discussion on trustworthiness in closed structures in Coleman (1988).

[17]Lemma B.1 in Appendix B.1 characterizes the club size that maximizes the DCV for various sets of parameters of the exponential club congestion function.

[18]Hence by Lemma 2 if the Empty Environment is not OCS it is also inefficient.

inelastic (elastic) if $\forall m \in \{2, \ldots, n_a - 1\} : \eta_h(m) > -1$ (respectively, $\eta_h(m) < -1$).

Lemma 4. *The club congestion function $h(m)$ is inelastic (elastic) if and only if $k_h(m)$ is strictly increasing (decreasing).*

4.1 The Relationship with the Connections Model

Jackson and Wolinsky (1996) introduce the connections model in which the utility of Agent i in the unweighted network g is

$$u_i^{JW}(g) = \sum_{j \neq i} \delta^{d_{ij}} - n_i(g) \times c$$

where d_{ij} is the geodesic distance between agents i and j, $\delta \in (0,1)$ the depreciation factor, $c > 0$ the direct connection cost and $n_i(g)$ the number of Agent i's direct neighbors. Network g is pairwise stable if no single agent gains by severing any of their links and no pair of unlinked agents would wish to establish a link between themselves.

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

Proposition 2. *The Connections model is a special case of the Club Congestion model. Specifically, let the congestion function be $h(2) = \delta$ and $\forall m > 2 : h(m) = 0$.*

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

The concept of pairwise stability is closely related to OCS. Both solution concepts imply that leaving a club of size two destroys the club and the formation of a new club of size two is an acceptable deviation. However, OCS also allows for the formation of bigger clubs, for leaving bigger clubs without destroying them and for deviations in which an individual can join an existing club. Naturally, when the discussion is limited to clubs of size two, the two concepts coincide. Letting $h(2) = \delta$ and $h(m) = 0$ for every $m > 2$ implies that there is no OCS environment with clubs of size larger than 2.

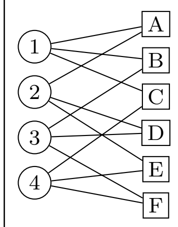
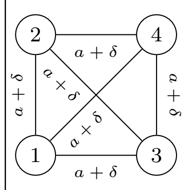
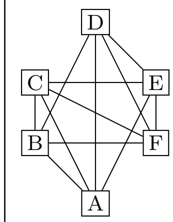
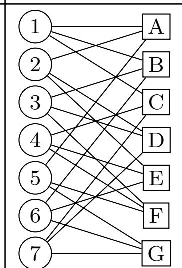
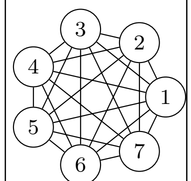
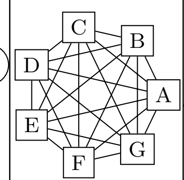
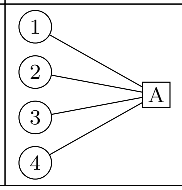
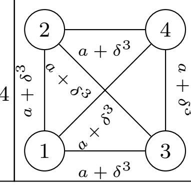
	Environment	List of Clubs	Induced Network	Club Network	Utilities
All Paired or 2-Complete (n=4)		Club A: 1 2 Club B: 1 3 Club C: 1 4 Club D: 2 3 Club E: 2 4 Club F: 3 4			$\forall i \in \{1,2,3,4\}$: $u_i = 3(a+\delta) - 3c$
3-Complete (n=7)		Club A: 1 2 5 Club B: 1 3 6 Club C: 1 4 7 Club D: 2 3 7 Club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

Figure 3: Three m-Complete environments, their induced weighted networks (weighted by the exponential club congestion function), induced club networks and the agents' utilities.

4.2 Two Useful Clubs Architectures

4.2.1 The m-Complete Environment

In m-Complete environments every pair of agents shares exactly one club and all the populated clubs are of the same size, m.[19] Formally,

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

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

$$\forall s \in S : n_G(s) = m \quad \text{or} \quad n_G(s) = 0.$$

Figure 3 provides three examples of m-complete environments. First note that every m-complete environment induces a complete weighted network. The first example in Figure 3 exhibits the All Paired Environment. Compared to other m-complete environments, the links are stronger but an agent needs to join more clubs in order to be connected to all the other agents. To demonstrate this trade-off consider the case of the 3-Complete Environment

[19]Given n_a and m, a necessary condition for the existence of an m-Complete environment is that $\frac{n_a-1}{m-1}$ and $\frac{n_a(n_a-1)}{m(m-1)}$ are integers. Here we discuss only cases in which m-Complete environments exist (See Arnold and Wooders (2005) and Page and Wooders (2007) for the important role of "leftovers" when agents are farsighted). As a combinatorial object an m-Complete Environment with n_a agents is the "Steiner System" $S(t, m, n_a)$ where $t = 2$.

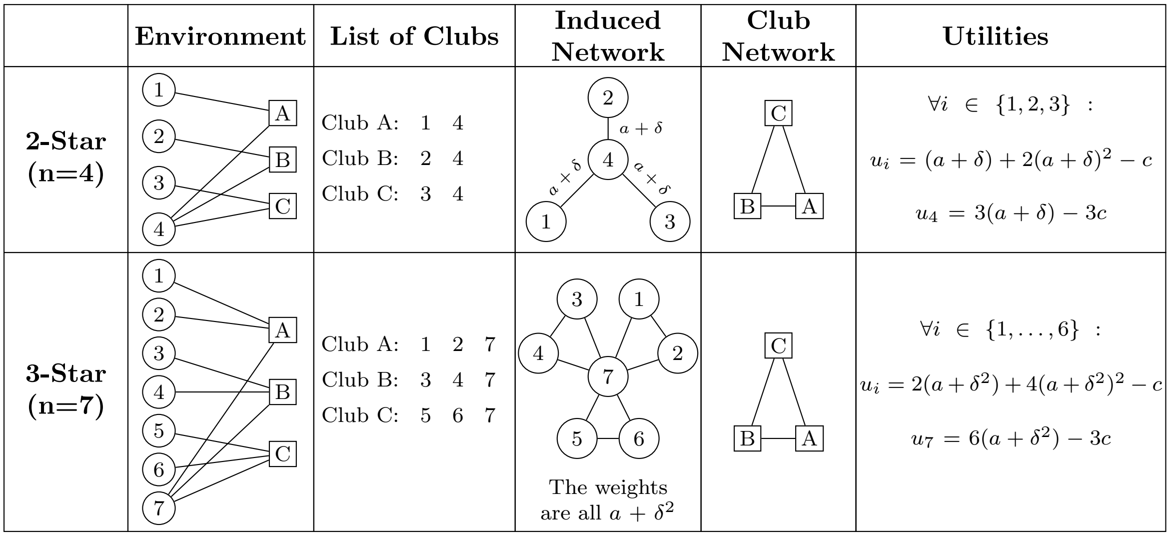


Figure 4: Two m-Star environments, their induced weighted networks (weighted by the exponential club congestion function), the induced club network and the agents' utilities.

with seven agents. In this environment each agent is a member of three clubs while in the All Paired Environment with $n_a = 7$ each agent pays for six memberships. At the same time, the links in the network induced by the All Paired Environment are stronger than those in the network induced by the 3-Complete Environment. The third example demonstrates that the Grand Club Environment is an m-complete environment where $m = n_a$.

Another important observation is that in m-complete environments indirect connections are never the shortest paths since every pair of agents is connected by a direct link and all links are identically weighted. While most of the other environments contain frictions due both to congestion and indirect connections, m-complete environments are free of the friction caused by indirect connections.

4.2.2 The m-Star Environment

In the literature on the strategic formation of social networks the star network often emerges as both stable and efficient for medium levels of linking costs. The star structure has one agent who has connections with all the other agents while these agents have no additional connections. We generalize this topology by defining the m-Star Environment, where the size of the clubs is $m \geq 2$, one agent is a member of all clubs and all other agents are members of a single club.[20] Formally,

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

$$\forall s \in S : n_G(s) = m \quad \text{or} \quad n_G(s) = 0.$$

[20] Given n_a and m, a necessary and sufficient condition for the existence of an m-star environment is that $\frac{n_a-1}{m-1}$ is an integer. We only discuss cases in which m-Star environments exist. In the graph theory literature the unweighted networks induced by m-Star Environments are called Windmill Graphs. In addition, these networks are specific cases of the $m-1$-quilts introduced by Jackson et al. (2012).

$$\exists i \in N \quad \text{such that} \quad \forall s', s'' \in \{s | n_G(s) > 0\} : N_G(s') \cap N_G(s'') = \{i\}.$$
$$\forall j \in N \backslash \{i\} : s_G(j) = 1.$$

Two m-Star environments are demonstrated in Figure 4. In 2-Star environments there is one agent who is a member of $n_a - 1$ clubs of size two with all the other agents and therefore provides all the connectivity in the induced network. Each club, in this example, induces a weight of $a + \delta$ and the distance between each pair of these $n_a - 1$ agents is $(a + \delta)^2$. Note that 2-Star environments are congestion free.

In the 3-Star Environment, all clubs are of size three and all include one special agent. Compared to the 2-Star Environment the central agent in the 3-Star Environment pays lower membership fees but suffers greater congestion. The larger m, the more direct connections peripheral agents have but the lower the quality of these connections, both direct and indirect. Thus, while agents in m-Complete environments suffer only from congestion and agents in 2-Star environments suffer only from depreciation caused by their indirect connections, agents in m-Star environments generally suffer from both types of friction. Note that the Grand Club Environment is an m-Star environment wherein $m = n_a$ and therefore contains no depreciation friction.

4.3 Efficiency

In many standard homogeneous models of strategic network formation (e.g. the connections model of Jackson and Wolinsky (1996)), strongly efficient topologies are the complete network for low linking costs, the star network for medium linking costs and the empty network for high linking costs. These results reflect the benefits of direct linking and the role of short indirect connections as a substitute for direct connections when linking costs are substantial.

Proposition 3 demonstrates that a similar intuition pertains in the Club Congestion Model with respect to constant levels of congestion. In order to control for the level of congestion friction we consider the set of m-Uniform environments in which all populated clubs are of size m. That is, 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 $\mathcal{G}_n^m$ and denote the set of all uniform environments with n agents by $\mathcal{G}_n^{all} = \cup_{k=2}^{n_a} \mathcal{G}_n^k$. Proposition 3 implies that among all uniform environments the strongly efficient ones are either m-Complete, m-Star, or empty.

Proposition 3. *Let $m \in \{2, \ldots, n_a\}$. For every club congestion function $h(\cdot)$ and m-Uniform Environment $G' \in \mathcal{G}_n^m$:*

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

Proposition 3 considers the set of environments in which all clubs are of size m. When membership fees are low, the m-Complete environments are efficient. Since m-Complete environments are symmetric across agents, the upper bound is independent of n_a and represents the agents' preference for costly direct links ($(m-1)h(m)-c$) over free indirect links ($(m-1)h^2(m)$). When membership fees increase, the importance of short indirect connections relative to costly direct connections and the low quality of long indirect connections emerges. The architecture of m-Star environments implements these preferences since the direct connections of the central agent keep the environment connected while making all other connections as short as possible. While the lower bound is independent of n_a, the upper bound increases with n_a since the larger the environment, the larger the return for membership for everyone except the central agent.

The proof is inspired by the proof of Proposition 1 in Jackson and Wolinsky (1996). We first show that when $c \leq (m-1)(h(m)-h^2(m))$, the m-Complete Environment achieves maximal total utility among all connected m-Uniform environments with no more than $\frac{n_a(n_a-1)}{m(m-1)}$ clubs due to the high quality of the direct connections. This result also holds when the m-Complete Environment is compared to large connected m-Uniform environments (since additional clubs are redundant) and to disconnected m-Uniform environments (since the total utility of the m-Complete Environment is convex in n_a). A result on hypergraphs from Berge (1989) is adopted to show that m-Star environments minimize the number of clubs required for an m-Uniform Environment to be connected. When $c \geq (m-1)(h(m)-h^2(m))$, the m-Star Environment achieves the maximal total utility among all connected m-Uniform environments due to the high quality of the indirect connections and the low total participation fees. This result also holds when the m-Star Environment is compared to non-empty disconnected m-Uniform environments since the union of two stars has a higher total utility than the sum of the totals of the two stars (due to additional indirect connections). The third part results from the fact that when $c > (m-1)h(m) + \frac{(n_a-m)(m-1)}{m}h^2(m)$, the total utility of the m-Star Environment is negative.

The comparison of efficient uniform environments across club sizes depends on the specific congestion function. However, two implications can be drawn from Proposition 3. One is that when membership fees are below $\min_{m\in\{2,\ldots,n_a\}}(m-1)(h(m)-h^2(m))$, the environment that achieves the maximal total utility among all uniform environments must be an m-Complete Environment. The other is that since the maximal DCV across club sizes is greater than $\max_{m\in\{2,\ldots,n_a\}}(m-1)(h(m)-h^2(m))$, among all uniform environments only an m-Star Environment or the Empty Environment may achieve maximal total utility when the membership fees are higher than the maximal DCV across club sizes.

4.4 The Stability of m-Complete Environments

The m-complete environment is OCS only when membership fees are neither too high nor too low. Low membership fees are an incentive to form new small clubs (if $m > 2$) while high membership fees are an incentive to leave one of the clubs, replacing direct connections with indirect ones. Note that in the m-Complete environment, the "No Joining" condition is irrelevant since if an agent joins an existing club, that agent pays membership fees but creates no new (or better) connections.

Proposition 4. *Denote by $\hat{k}$ the club size that maximizes the DCV.*
Let $m \in \mathbb{N}$, $n_a > m \geq 2$. An m-Complete Environment is OCS if and only if

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

Let $m = n_a$. The m-Complete Environment (the Grand Club Environment) is OCS if and only if

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

In m-Complete environments, forming new clubs of size smaller than m may reduce congestion. Hence, the lower bound of both parts of Proposition 4 implies that in order for an m-Complete Environment to be OCS, membership fees should be high enough to preclude new clubs from being formed. The benefit of a coalitional deviation to a club of size $k < m$ is its DCV ($(k-1)h(k)$) net the value of these links in the original environment ($(k-1)h(m)$). The DCV applies here since indirect connections never constitute the shortest path for new club formation deviations in m-Complete environments. Note that the larger the new club, the larger the number of original links whose value has been lost. Hence, deviation to clubs with more than $\hat{k}$ members is less attractive than deviation to clubs of size $\hat{k}$ because of the lower DCV and the greater loss of original value.

An agent with multiple memberships in an m-Complete environment ($n_a > m$) may consider leaving a club to trade-off reduced membership payments with replacing some direct connections with indirect ones. The first case of Proposition 4 guarantees that membership fees will not be high enough to make such trade-offs worthwhile. The second case relates to the Grand Club Environment wherein an agent who has left the club will not be compensated by indirect connections to the other agents.

The existence of a stable m-Complete environment is not guaranteed. It is possible that the lower bound may be higher than the upper bound. Claim 1 uses the exponential club congestion function with strong congestion (low δ) to demonstrate that even then m-Complete environments with large clubs may be OCS.

Claim 1. *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 an 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.*

When congestion is strong, but there exists a non-congested part to the club congestion function, if the clubs are large enough ($m > \bar{m}$) the existence of membership fees for which an m-Complete environment is OCS is guaranteed. The second part of Claim 1 shows that there even exists a range of membership fees for which multiple m-Complete environments are OCS (all those with $m > \tilde{m}$). This result implies non-monotonicity in the relationship between congestion and the size of clubs in stable environments: m-complete environments with intermediate size clubs are unstable while m-complete environments with either small clubs (wherein each individual maintains many high quality affiliations) or large clubs (wherein each individual maintains few low quality affiliations) are open clubwise stable.[21]

Two extreme cases are of special interest - the All Paired and the Grand Club environments. The stability of these environments depends on the relative importance of the two frictions - club congestion and affiliation fees - that obtain in m-complete environments.

4.4.1 The Stability of the All Paired Environment

In the All Paired Environment the agents suffer no congestion and no depreciation. When membership costs are introduced the strict super environments of the All Paired are no longer OCS since redundancy is costly. In the All Paired Environment, joining an existing club or forming a new one never constitutes a beneficial deviation since the additional affiliations are costly and agents already share small clubs with every other agent. Therefore, only incentives to leave a club and use an indirect connection instead need to be examined. As long as the gain from staying in the club ($h(2) - c$) is greater than an indirect connection ($h^2(2)$), the All Paired environment is OCS.

A similar argument guarantees that no sub-environment of the All Paired Environment is OCS when $h(2) - c > h^2(2)$. This argument, however, does not rule out the stability of environments wherein the smallest club shared by some pair of agents is larger than size 2. Such environments are not OCS if the costs of forming a new club are lower than the benefit derived from eliminating the club congestion suffered by this pair, that is, when $h(2)-h(3) > c$. Therefore, the uniqueness of the All Paired Environment is guaranteed when membership costs are small enough to allow agents to form new two-agent clubs in order to resolve the friction created by indirect connections and the friction of club congestion. Formally, the All Paired Environment is the unique OCS environment when $c \in (0, \min\{h(2) - h^2(2), h(2) - h(3)\})$. In fact, in this range of affiliation fees, the All Paired Environment is also the unique PE and SE.[22]

4.4.2 The Stability of the Grand Club Environment

In the Grand Club Environment agents suffer severe club congestion but no depreciation friction and minimal membership fees. Proposition 1 states that when there is no club congestion and $n_a - 1 > c > 0$ the Grand Club Environment is both OCS and the unique efficient environment. But as Proposition 4 implies, when club congestion is introduced, the Grand Club en-

[21] Consider the case where $h(m) = \frac{1}{32} + (\frac{1}{4})^{m-1}$. The All Paired Environment is OCS in $[0, \frac{3}{16}]$, for $m \in \{3, \ldots, 9\}$ the m-complete is never OCS and for $m \geq 10$ every m-complete is OCS in [0.25,0.27].

[22] When $c \in (0, \min\{h(2) - h^2(2), h(2) - h(3)\})$ every environment that is not a strictly spanning super environment of the All Paired Environment is Pareto dominated by an environment that is a strictly spanning super environment of the All Paired Environment (by adding the missing clubs of size 2). In addition, since $c > 0$, every strictly spanning super environment of the All Paired Environment is Pareto dominated by the All Paired Environment due to redundancies.

vironment is OCS if and only if $c \in \big[\max_{k\in\{2,\dots,\min\{n_a-1,\hat{k}\}\}}(k-1)[h(k)-h(n_a)], (n_a-1)h(n_a)\big]$.

Since the DCV of the reciprocal club congestion function is unity, the most attractive deviation from the Grand Club Environment is to a club of size 2 (wherein the loss of original value is minimal). Therefore, when club congestion is reciprocal, the Grand Club Environment is OCS if and only if $c \in [1 - \frac{1}{n_a-1}, 1]$. Hence, when the club congestion function is reciprocal a range of membership fees in which the Grand Club environment is OCS always exists.

Generally, as mentioned above, Proposition 4 does not guarantee that such a range exists. Nevertheless, Lemma 4 is used to demonstrates that when club congestion is not too sensitive to club size, the Grand Club Environment is OCS for some range of membership fees.

Claim 2. *If the club congestion function is inelastic, a range of membership fees in which the Grand Club Environment is OCS exists.*

When the club congestion function is exponential, Claim 2 guarantees that the Grand Club Environment is OCS for some range of membership fees in some cases (e.g. when $a = 0$ and $\delta > 1 - \frac{1}{n_a}$). The case of $\delta \in (0, \frac{1}{2})$ wherein the congestion component of the club congestion function is substantial is analyzed in Claim 3 (note that the Grand Club Environment is not covered by Claim 1).

Claim 3. *Let $n_a \geq 4$ and $h(\cdot)$ be an exponential club congestion function where $\delta \in (0, \frac{1}{2})$. For $a = 0$ there is no range of membership fees in which the Grand Club environment is OCS. But if $a > 0$, there exists an $\bar{n}_a$ such that $\forall n_a : n_a > \bar{n}_a$, a range of membership fees in which the Grand Club environment is OCS exists.*

The first part demonstrates that when there is no non-congested component, the congestion is too strong for a Grand Club environment to be OCS. However, as long as there is a non-congested component, the Grand Club environment can be OCS for some range of membership fees, as long as the set of agents is large enough to make the non-congested part important (see also Appendices B.1 and C.2). A sociological interpretation of this result may imply that social solidarity (which does not depend on club size) may be useful in maintaining big clubs even when club congestion is strong.

4.5 The Stability of m-Star Environments

An m-Star environment is OCS if no agent wishes to join or leave an existing club and no subset of agents benefits from forming a new club.

The central agent prefers to leave a club when membership costs are higher than the benefit derived from direct links to the agents in the club. A peripheral agent wishes to leave a club when participation fees are higher than the benefit of direct links to other club members and indirect connections to all other peripheral agents. Therefore, the peripheral agents' incentives to leave a club are weaker than those of the central agent. Thus, the upper bound on the range of membership fees in which an m-Star environment is OCS depends on the central agent's incentives. Since by leaving a club, the central agent disconnects from the other members of the club, the upper bound is higher than in m-Complete environments (when $m < n_a$) where direct links that have been lost can be replaced by indirect connections.

Joining an existing populated club is not a relevant consideration for the central agent since this agent is already a member of all populated clubs. When joining an existing club, a peripheral agent replaces $m - 1$ indirect connections with direct connections that are both congested (relative to existing direct connections) and costly. Thus, the lower bound on the range of membership fees in which an m-Star environment is OCS should be high enough to make the existing indirect connections to be more attractive than new direct connections for a peripheral agent.

The third consideration is the formation of a new club. In forming a new club, a peripheral agent always gains more than the central agent. If the new club is smaller than m, the central agent only gains from improved direct connections while the peripheral agents also gain from better indirect connections. When the new club is weakly larger than m, the central agent gains nothing while the peripheral agents may gain from the new direct links created. Therefore, the lower bound on the range of membership fees for which an m-Star environment is OCS should be high enough to deter the formation of clubs that include only peripheral agents.

A peripheral agent always prefers to form a new club with members of other clubs. If the new club is of size $k < m$, sharing it with an agent who is also affiliated with the original club yields a single improved direct connection $(h(k) - h(m))$ while forming the new club with an agent who belongs to a different original club turns an indirect connection into a direct one $(h(k) - h^2(m))$ (and it may also have a positive effect upon indirect connections). If the new club is not smaller than m, then, sharing this new club with an agent who is also affiliated with the original club yields nothing while forming a new club with an agent who belongs to a different original club may improve one indirect connection. Moreover, when the new club is small, its attractiveness increases with the number of original clubs that are represented in the new club. Hence, the lower bound on the range of membership fees in which an m-Star environment is OCS should be high enough to deter peripheral agents from coordinating the formation of a new club that includes a diverse collection of members relative to the original clubs. Proposition 5 summarizes these incentives.

Proposition 5. *Let $n_a > m \geq 2$ and let $h(\cdot)$ be the club congestion function. Denote $\gamma \equiv \frac{n_a - 1}{m - 1}$, $\eta_k \equiv \lceil \frac{k}{\gamma} \rceil$ and $l_h = \min\{k \in \mathbb{Z} | h(k) \leq h^2(m)\}$.*

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

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

where

$$FNS_h(k, m) = (k - 1)[h(k) + (m - 2)h(k)h(m) - (m - 1)h^2(m)]$$
$$FNL_h(k, m, n_a) = (k - \eta_k)(h(k) - h^2(m))$$

2. *If $\gamma < m$ the m-Star environment is OCS if and only if*[23]

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

where

$$FNI_h(k,m,n_a) = (k-1)h(k) - (\eta_k - 1)h(m) + (n_a - m - (k - \eta_k))h(m)h(k) - (n_a - m)h^2(m)$$

$$J_h(m) = (m-1)[h(m+1) - h^2(m)]$$

4.5.1 The Stability of a 2-Star Environment

Proposition 5 implies that a 2-Star Environment is OCS if the membership fees are high enough to preclude any subset of peripheral agents from founding a new club that does not include the central agent and low enough that it is beneficial for the central agent to maintain each affiliation. Claim 4 summarizes the conditions for stability for a 2-Star Environment for various characteristics of the club congestion function:

Claim 4. *Denote $l_h = \min\{k \in \mathbb{Z} | h(k) \leq h^2(2)\}$.*

1. *Let $h(\cdot)$ be the club congestion function. The 2-Star Environment is OCS if and only if $h(2) \geq c \geq \max_{k \in \{2,\ldots,\min\{l_h-1,n_a-1\}\}} (k-1)(h(k) - h^2(2))$.*
2. *Let $h(\cdot)$ be an elastic club congestion function. The 2-Star Environment is OCS if and only if $h(2) \geq c \geq h(2) - h^2(2)$.*
3. *Let $h(\cdot)$ be the reciprocal club congestion function. The 2-Star Environment is OCS if and only if $c \in [0,1]$.*
4. *Let $h(\cdot)$ be the exponential club congestion function. The 2-Star Environment is OCS if and only if*

$$a + \delta \geq c \geq \max_{k \in \{2,\ldots,\min\{l_h-1,n_a-1\}\}} (k-1)((a + \delta^{k-1}) - (a+\delta)^2)$$

In particular, if $a = 0$, then the 2-Star Environment is OCS if and only if $c \in [\delta - \delta^2, \delta]$.

Agents in a 2-Star Environment suffer no congestion. Therefore, indirect paths in the 2-Star Environment can only be improved by direct links. By forming a new club of size k that does not include the central agent, peripheral agents only gain from direct links to other deviators, $(k-1)(h(k) - h^2(2))$. The second part shows that when the club congestion function is elastic

[23] If $\frac{m(m-1)}{n_a - 1} < 2$ then $FNI_h(m,m,n_a) \geq J_h(m)$. Thus, in this case $J_h(m)$ is not the maximizing element of the lower bound.

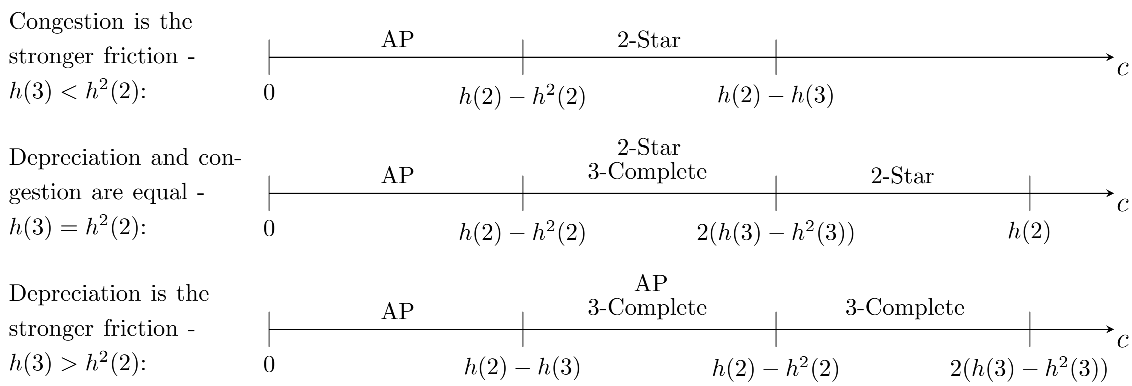


Figure 5: A graphical representation of Corollary 1.

there is always a range of membership fees wherein the 2-Star Environment is OCS.[24] The third and the fourth parts characterize the stability of the 2-Star Environment in cases of reciprocal and exponential club congestion functions, respectively.

4.5.2 The Emergence of Weak Links

We are now ready to demonstrate the effects of club congestion and depreciation in indirect connections on the structure of stable environments. As previously discussed, as long as membership fees are low enough, each agent is able to avoid both congestion and depreciation by forming intimate clubs with all other agents. However, once membership fees become too high for agents to maintain so many clubs, weak links emerge as low quality substitutes.

Corollary 1. *Let* $n_a > 3$ *and let* $1 > h(2) > h(3) \geq 0.15$.[25]

1. *If* $h(3) < h^2(2)$ *then*
   (a) *The All-Paired Environment is the unique OCS Environment if and only if* $c \in (0, h(2) - h^2(2))$.
   (b) *For* $c \in (h(2) - h^2(2), h(2) - h(3))$ *the 2-Star Environment is OCS while the All-Paired and the 3-Complete Environments are not.*
2. *If* $h(3) = h^2(2)$ *then*
   (a) *The All-Paired Environment is the unique OCS environment if and only if* $c \in (0, h(2) - h^2(2))$.

[24]Generally, the existence of such a range is not guaranteed. Consider, for example, the case where $h(2) = 0.3$, $h(3) = 0.25$ and $n_a \geq 4$. In this case, the central agent would abort her affiliations for every membership fee above 0.3. However, a triad of peripheral agents will form a new club if the membership fees are lower than 0.32. Thus, if $h(2) = 0.3$, $h(3) = 0.25$ and $n_a \geq 4$, the 2-star Environment is never OCS.

[25]The lower bound on $h(3)$ is necessary because for very small values of $h(3)$ the 3-Complete Environment may not be OCS even if $h(3) \geq h^2(2)$. However, the lower bound on $h(3)$ stated in Corollary 1 is not tight. The exact condition required for the third part of Corollary 1 is $h(3) > \max\{h^2(2), \frac{3}{4}[1 - \sqrt{1 - \frac{8h(2)}{9}}]\}$.

(b) *For* $c \in (h(2) - h^2(2), 2(h(3) - h^2(3)))$ *the* 2*-Star and the* 3*-Complete environments are OCS while the All-Paired Environment is not.*

(c) *For* $c \in (2(h(3) - h^2(3)), h(2))$ *the* 2*-Star Environment is OCS while the* 3*-Complete and the All-Paired Environments are not.*

3. *If* $h(3) > h^2(2)$ *then*

   (a) *The All-Paired Environment is the unique OCS environment if and only if* $c \in (0, h(2) - h(3))$.

   (b) *For* $c \in [h(2)-h(3), h(2)-h^2(2))$ *the All-Paired and the* 3*-Complete Environments are OCS while the* 2*-Star Environment is not.*

   (c) *For* $c \in (h(2) - h^2(2), 2(h(3) - h^2(3))]$ *the* 3*-Complete Environment is OCS while the All-Paired Environment is not.*

Two interpretations of weak links are suggested by Corollary 1. First, when congestion is the stronger friction ($h(3) < h^2(2)$), once the All-Paired Environment ceases to be stable, the 2-Star Environment becomes OCS while the 3-Complete Environment is not OCS. Indeed, as in the standard literature on strategic network formation, costless indirect links emerge as low quality substitutes for costly direct connections (e.g. Jackson and Wolinsky (1996)). Hence, when congestion is more substantial than depreciation, weak ties are indirect (not congested) links. However, when depreciation is the stronger friction ($h(3) > h^2(2)$), and membership fees become high, the 3-Complete Environment, rather than the 2-Star Environment emerges as OCS (for some membership fees both the All-Paired and the 3-Complete environment are OCS). Thus, Corollary 1 suggests a new insight - larger clubs, that induce low quality direct connections at low cost (per link) turn out to be substitutes for costly intimate connections. Hence, when congestion is less substantial than depreciation, weak ties are direct congested links.

One important implication of this concerns the architecture of the induced social network. Most of the literature predicts that the complete network cannot survive non-negligible maintenance costs. In these cases, according to the literature, the frugal star architecture emerges as an equilibrium that efficiently maintains connectivity at much lower costs. By incorporating club formation into the setup of strategic network formation, we show that this prediction holds only if congestion is the predominant friction in the formation process (as implicitly assumed by most existing models). When depreciation is stronger we show that complete networks may survive high maintenance costs by reducing the quality of the links.

The second part of Corollary 1 shows that when the frictions are of the same magnitude ($h(3) = h^2(2)$), both 3-Complete and 2-Star environments are OCS once membership fees are too high for the ideal All-Paired Environment to be OCS. Note, however, that the 2-Star Environment is a Minimally Connected Environment while the 3-Complete Environment is not. As a result, the marginal utility of each affiliation in the 2-Star Environment is higher than that of the 3-Complete Environment. Therefore, the range of costs wherein the 2-Star

Environment is OCS is larger than that wherein the 3-Complete Environment is OCS.

As was previously discussed, the All-Paired Environment is the SE environment for $c \in (0, \min\{h(2) - h^2(2), h(2) - h(3)\})$. When congestion is the stronger friction, once the membership fees become too high ($c > h(2) - h^2(2)$) the 2-Star Environment is socially preferred to the All-Paired and the 3-Complete environments.[26] In fact, this is still the case when depreciation is the stronger friction ($h(3) > h^2(2)$) but congestion is still considerable (e.g. when $h(2) - h(3) > h(3) - h^2(2)$ and n_a is large). Only when depreciation is the stronger friction and congestion is less of an issue - does the 3-Complete Environment achieve higher total utility than both the All-Paired and the 2-Star environments for some range of membership fees. It should be noted that generally the social attractiveness of the 2-Star Environment compared to that of the 3-Complete Environment grows with the number of agents since the number of clubs (and therefore membership payments) grows linearly in the former while in the latter it grows quadratically.

4.5.3 Other m-Star Environments

m-Star environments where $m > 2$ are hybrids of the architectures previously discussed. Unlike in m-Complete environments, agents in m-Star environments do suffer depreciation. Unlike the 2-Star Environment, m-Star environments where $m > 2$ include some congestion friction.

Claim 5. *Let $n_a \geq 9$.*

1. *Let $h(\cdot)$ be the reciprocal club congestion function. The 3-Star Environment is OCS if and only if $c = 1$.*
2. *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 $\delta \geq \frac{1}{2}$.*

Claim 5 shows that when the club congestion function is reciprocal then the 3-Star Environment is OCS only when $c = 1$. Note that the reciprocal club congestion function at this level of membership fees exhibits extensive multiplicity of equilibria including the Empty Environment (Proposition B.1.1), the Grand Club Environment (see the discussion preceding Claim 2) and the 2-Star Environment (Claim 4.3).

When the club congestion function is exponential with $a = 0$, we prove that the most attractive new club is one formed by two peripheral agents that do not share a club in the 3-Star Environment. This club provides these agents with a direct link between themselves and an improved indirect link to non-central agent affiliated with their partner in the original

[26] The total utility in the All-Paired Environment is $n_a(n_a - 1)(h(2) - c)$. The total utility in the 3-Complete Environment is $\frac{n_a(n_a-1)}{2}(2h(3) - c)$. The total utility in the 2-Star Environment is $(n_a - 1)(h(2) - c) + (n_a - 1)(h(2) + (n_a - 2)h^2(2) - c)$. The All-paired Environment dominates the 2-Star Environment if and only if $c < h(2) - h^2(2)$ and the 3-Complete Environment if and only if $c < 2(h(2) - h(3))$. The 3-Complete Environment dominates the 2-Star Environment if and only if $c < \frac{2n_a-4}{n_a-4}(h(3) - h^2(2)) - \frac{4}{n_a-4}(h(2) - h(3))$ that approaches $2(h(3) - h^2(2))$ from above when n_a grows larger.

environment. If congestion friction is strong then the 3-Star Environment is never OCS. On the one hand, strong congestion leads to small benefits accruing to the central agent from each affiliation. On the other hand, due to congestion the benefit of forming a new small club is relatively high. When congestion is weakened (δ increases) the incentive for the central agent to leave a club weakens since membership in the club becomes more profitable. In addition, peripheral agents refrain from coalitional deviation since the links induced by the original clubs are satisfactory.

The case wherein the club congestion function is exponential with $a = 0$ corresponds to the second part of Corollary 1 (assuming δ is high enough). Thus, the All-Paired Environment is the unique OCS environment when membership fees are very low. Then, when membership fees increase, the 2-Star and the 3-Complete environments become OCS. But, when $c > 2(\delta^2 - \delta^4)$, the 3-Complete Environment ceases to be OCS since aborting existing affiliations becomes worthwhile as indirect connections are now an attractive alternative. However, even though congestion in the 3-Star Environment is similar to that of the 3-Complete Environment, leaving an existing club is not compensated by an indirect connection. Indeed, by Claim 5, if $\delta \geq \frac{1}{2}$ then for slightly higher costs ($\delta + \delta^3 - 2\delta^4 > 2(\delta^2 - \delta^4)$), the 3-Star Environment becomes OCS. In fact, when $c \in (\delta, 2\delta^2]$, the 2-Star Environment is not OCS, while the 3-Star Environment still is due to the higher value of a single affiliation to the central agent. See Appendix C for further analysis of the stability of m-Star environments including numerical results for the case of an exponential congestion function with $a > 0$.

4.6 The Stability-Efficiency Gap

Proposition 4 shows that when $n_a > m$ the upper bound on membership fees for which an m-Complete Environment is OCS is $(m-1)(h(m) - h^2(m))$. Therefore, Proposition 3 and Lemma 2 imply that there is never a case wherein an m-Complete Environment ($n_a > m$) is strongly efficient and not OCS.[27]

This, however, is not true for m-Star environments. Recall that by Proposition 3, m-Star environments are efficient relative to m-Uniform environments for some range of costs such that $c = k_h(m)$ is always strictly included within this range. By Proposition 5, m-Star environments are never OCS when $c > k_h(m)$, meaning that a range of costs always exists where m-Star environments are not OCS although they are efficient relative to all m-Uniform environments. In fact, Figure 6 demonstrates multiple cases where m-Star environments are strongly efficient but not OCS.

We used our Matlab code package (see Section 8) to calculate the strongly efficient 5-agents environment for various exponential congestion functions and membership fees.[28] Each shape

[27] The opposite, however, is possible. Consider the case where $h(2) = 0.5$ and $h(3) = 0.3$. By Proposition 4, the 3-Complete Environment is OCS for membership fees between 0.2 and 0.42. However, when $n_a = 9$, the total utility of the 2-Star Environment is $22 - 16c$ while the total utility of the 3-Complete Environment is $21.6 - 36c$. Thus, the 3-Complete Environment is OCS but not SE.

[28] $a \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ and $\delta \in \{0, 0.05, 0.1, \ldots, 0.95 - a\}$ for the exponential congestion function and $c \in \{0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.4, 2.7, 3\}$ for the membership fees.

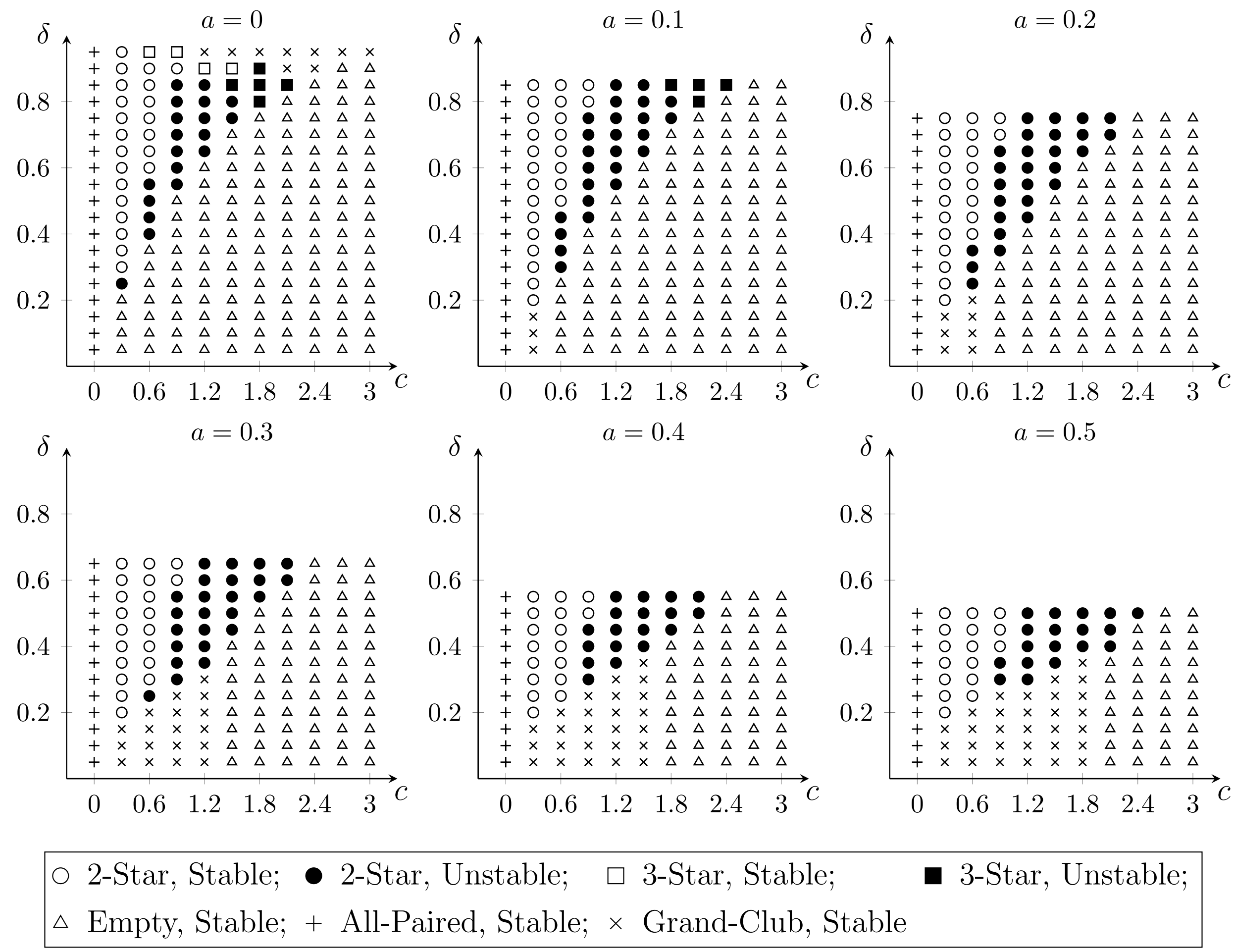


Figure 6: Strong efficiency analysis for 5 agents and the exponential club congestion function.

in the graphs in Figure 6 represents the type of the strongly efficient environment and whether it is OCS. First, note that all the strongly efficient environments are either m-Complete (All-Paired or Grand-Club), m-Star (2-Star or 3-Star) or empty. Second, the unstable strongly efficient environments are all m-Stars (2-Star or 3-Star).

5 The Individual Congestion Model

The Club Congestion model focuses on the "quality" of links within a club. But one can think about other types of congestion. The Individual Congestion model deals with club affiliations that require attention and time. Agents with a thin portfolio of affiliations are able to pay more attention (or time) to each one of their club memberships and thus form connections of high quality with other members. On the other hand, agents who are members of many clubs possess many weak direct relations since they cannot devote much attention to each one of their memberships. We now introduce individual congestion by relating the weight of each link in the induced network to a non-increasing function of the number of affiliations possessed by the agents involved.

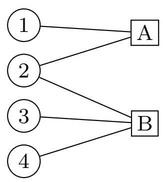
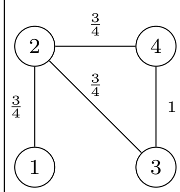
Environment	List of Clubs	Friction	Induced Network	Utilities
	Club A: 1 2 Club B: 2 3 4	$b(1) = 1$ $b(2) = \frac{3}{4}$ $b(3) = \frac{2}{3}$ $c = \frac{1}{4}$		$u_1 = \frac{3}{4} + 2 \times \frac{9}{16} - \frac{1}{4} = 1\frac{5}{8}$ $u_2 = 3 \times \frac{3}{4} - \frac{1}{2} = 1\frac{3}{4}$ $u_3 = \frac{3}{4}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

Figure 7: An environment, its induced weighted network and the agents' utilities.

An **individual congestion function** is a non-increasing function $b : \{1, 2, \ldots, n_s\} \to [0, 1]$ such that the weight of a link between two agents $i, i' \in N$ who share the same club in Environment G is,

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

The weight of the link between agents i and i' is determined by the product of the congestion attributed to each of their affiliation portfolios. In so doing, we implicitly assume that the specifics of the clubs with which these agents are affiliated are inconsequential with regard to the attention paid by agents to their affiliations. Also, as in the Club Congestion model, we assume that the number of clubs that two agents share does not affect the strength of the link between them.

Figure 7 provides an example of the Individual Congestion model. Agent 2 shares Club A with Agent 1 and Club B with Agents 3 and 4. The individual congestion function is such that there is no congestion when a single affiliation is maintained. However, maintaining two affiliations leads to an individual congestion of $\frac{3}{4}$. Thus, in the induced weighted network, Agent 2 is connected to all other agents by a congested weighted link. Agent 3 is linked to Agent 4 by a non-congested connection, since both maintain a single affiliation. The rightmost column demonstrates the costs and benefits in this example. The shortest paths between agents provide the benefits. Since the induced network is connected but not complete, some of the shortest paths are indirect connections. Consider, for example, the path between Agent 1 and Agent 3 through Agent 2. The weight of this path is the product of the weight of the link between Agents 1 and 2 and the weight of the link between Agents 2 and 3.[29] Notice that both weights include the individual congestion of Agent 2. We consider this double counting a requirement since Agent 2 maintains these two links independently. Finally, the membership fees constitute the cost. While Agents 1, 3 and 4 pay membership fees for only one club affiliation, Agent 2 pays for two memberships.

[29] In this example there is an additional shortest path between Agents 1 and 3, which goes through agents 2 and 4. This longer path has the same weight since the link between Agents 3 and 4 is not congested.

5.1 Stability and Efficiency

The Grand Club Environment is the efficient environment in the Individual Congestion model (when membership fees are not too high) since both individual congestion and membership fees drive agents away from multiple affiliations.

Proposition 6. *In the Individual Congestion model where $b(1) > 0$:*

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

The stability and efficiency of the Grand Club Environment reflect the fact that agents minimize individual congestion friction by sharing the same club. Similar incentives exist when (not too high) membership fees are the sole friction introduced into the model (the case studied in Section 3). Indeed, in both cases we observe that agents may fail to coordinate on the Grand Club Environment and thus find themselves in some other, inefficient, OCS environment.

Consider, for example, a Minimally Connected environment other than the Grand Club Environment. A necessary condition for such an environment being OCS is that the loss of connection to some club members outweighs the gain of leaving a club. When membership fees are introduced alone, Proposition 1 suggests that indeed if membership fees (which are the gain of leaving a club) are sufficiently low this environment is OCS. Minimally Connected environments may also be OCS when the only friction introduced is individual congestion, but for a different reason. In this case, leaving a club enables the agents to devote more attention to their other affiliations thereby improving all remaining direct connections. Thus, Minimally Connected environments may be OCS in the Individual Congestion model if the congestion is not very strong. In this case improvement of the deviating agent's other direct connections is not large enough to cover the loss of connections to club members who have been discarded.[30]

[30] For example, let G be the 3-Star Environment with 5 agents (the second example in Figure 2). By Proposition 1, in a model with no congestion, if $c \leq 2$ then G is OCS. In the Individual Congestion model with no membership fees, if $b(1) = 0.75$ and $b(2) = 0.38$ then G is OCS.

While Proposition 1 guarantees that Minimally Connected environments are the only OCS environments in a setup with membership fees and no congestion friction, when individual congestion is introduced there may be many other OCS environments that are either connected (but not minimally)[31] or disconnected.[32] This extensive multiplicity is due to the unattractiveness of deviations that enlarge the portfolio of affiliations. Such deviations damage the value of all existing direct connections (See, for example, Lemma 9 and Lemma 10 in Appendix A).

Part 3a shows that every combination of an individual congestion function and membership costs that makes a non-empty environment OCS, must also make the Grand Club Environment OCS. Moreover, Part 3b points out that while the efficient architecture is not sensitive to the details of individual congestion friction, the type of mis-coordination is sensitive to its specifics. Thus, by Proposition 6, the Individual Congestion model demonstrates that when club formation is mainly driven by attention constraints and membership fees, no gap between stability and efficiency emerges but there is a high potential for lack of coordination. One extreme example is given by Proposition 7 that characterizes the stability of m-Complete environments in the Individual Congestion model.

Proposition 7. *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*

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

m-Complete environments induce considerable congestion, at least for small clubs, since agents are members of many clubs and avoid using indirect connections. Therefore, it might be surprising that for a large set of parameters these environments are OCS. The proof uses the fact that in the Individual Congestion model a necessary condition for the attractiveness of a new club is that it provides each member with at least one new direct neighbor (Lemma 9). Hence, in m-Complete environments the formation of new clubs is never attractive. For similar reasons no agent wishes to join an existing club (Lemma 10). Thus, since for low membership fees agents are reluctant to terminate affiliations, the m-Complete environment is OCS. Incentive to terminate an affiliation increases with the number of agents who do not belong to the club $(n_a - m)$ and with the relief in congestion produced by leaving a club $(b(\gamma - 1) - b(\gamma))$. Hence, m-Complete environments are OCS when membership fees are low if the number of agents is not too high and if the individual congestion function is not too steep.

[31]The simplest example of a connected, but not minimal, OCS environment is the circle with 4 agents ($S = \{\{1,2\},\{2,3\},\{3,4\},\{1,4\}\}$) when $b(1) = 1$, $b(2) = \frac{3}{4}$ and $b(3) = 0$.

[32]For example, suppose $c = 0$, let P_{n_a} be the Partitioned environment with n_a agents and two populated clubs wherein half the agents are affiliated with one club and the other half with the other club (assume $n_a > 2$ is even). It turns out that P_{n_a} is OCS for every even n_a if and only if $b(1) \geq 3b(2)$. To see that, note that for every Agent i affiliated with Club s_1, $u_i(P_{n_a}) = (\frac{n_a}{2} - 1)b^2(1)$. If Agent i joins Club s_2, the utility is $u_i(P_{n_a} + \{i, s_2\}) = (n_a - 1)b(1)b(2)$. $u_i(P_{n_a}) \geq u_i(P_{n_a} + \{i, s_2\})$ for any n_a if and only if $b(1) \geq 3b(2)$. $b(1) \geq 3b(2)$ also prevents deviations to a new club (because formation of a new club is less attractive than joining the another already populated club). Finally, leaving a club is not a profitable deviation since $c = 0$.

5.2 A Model with Club Congestion and Individual Congestion

Consider the model containing both club congestion and individual congestion. Given club congestion function $h(\cdot)$ and individual congestion function $b(\cdot)$, the weight of a link between two agents $i, i' \in N$ that share a club in 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 weight of each link in the induced network is the product of a non-increasing function of the size of the smallest club shared by the end agents and a non-increasing function of the number of affiliations of those same agents.[33]

5.2.1 The Coauthors Model (Jackson and Wolinsky (1996))

In the Coauthors model agents equally distribute one unit of attention between their direct relations. The value of each relation only depends on the attention devoted to the link by the two end agents. Specifically, the utility of Agent i from Network g is

$$u_i^{CA}(g) = \sum_{j \neq i: \{i,j\} \in g} [\frac{1}{n_i(g)} + \frac{1}{n_j(g)} + \frac{1}{n_i(g) n_j(g)}]$$

where $n_k(g)$ is the number of Agent k's direct neighbors. Denote by $CA(n)$ the set of pairwise stable networks with n agents.

Agents' preferences are **truncated at geodesic distance** D if the weight of path $p = \{x_1, \ldots, x_l\}$ is the product of the weights on its links if $l - 1 \leq D$ and zero otherwise. For example, when $D = 1$ Agent i benefits from the path to Agent j only if both agents are directly connected in the induced network. Denote by $OCS(c, n, h, b, D)$ the set of OCS environments with n agents wherein the club congestion function is h, the individual congestion function is b, the membership fees are c and the agents' preferences are truncated at geodesic distance D.

Proposition 8 uses the definitions and notations introduced in Section 4.1 for a comparison of unweighted networks and environments to show that the Coauthors Model is a truncated

[33] Zhou et al. (2007) study the loss of information when a bipartite network is projected onto a directed one-mode network. They propose the following weighting scheme: Each node is endowed with some resources that are distributed equally among affiliations. Then, the total amount of resources each club accumulates is equally distributed among all its members. The weight of the direct link from Node i to Node j is the sum of the resources transferred from Node i to Node j. While Zhou et al. (2007) is a technical work on projection and not a model of strategic network formation, its weighting scheme bears some relation the model presented in this section. Specifically, the distribution of the node's resources and the distribution of the club's accumulated resources resemble the Individual Congestion function and the Club Congestion function, respectively.

version of our model containing the two types of congestion.[34]

Proposition 8. *The Coauthors Model is a special case of the truncated model with club congestion and individual congestion. Specifically, let the club congestion function be* $h(2) = 1$, $\forall m > 2 : h(m) = 0$ *and let the club congestion function be* $b(k) = \frac{1}{2}[1 + \frac{1}{k}]$*:*

1. $g \in CA(n)$ *if and only if* $G_g \in OCS(\frac{1}{4}, n, h, b, 1)$.
2. *If* $G \in \mathcal{G}_n \backslash \mathcal{G}_{\mathbb{G}n}$ *then* $G \notin OCS(\frac{1}{4}, n, h, b, 1)$.

6 Club Rules: Closed Clubwise Stability

There are many possible rules regarding the forming, joining or leaving of social clubs. Each set of rules induces a different set of possible deviations and therefore corresponds to a different stability concept. So far we have only considered Open Clubwise Stability that implements an open environment in which the joining, leaving and formation of clubs is done freely as long as membership fees are paid. But there are environments in which clubs have more restrictive rules. For example, clubs in which the acceptance of new members requires the agreement of incumbent club members, are very common. Such a requirement is observed in various social groups (e.g. academic departments, fraternities and sororities and kibbutzim) as well as in international organizations (e.g. the European Union). We demonstrate the application of this club rule by introducing the Closed Clubwise Stability solution concept which considers stability under the requirement that incumbents must unanimously approve every new member.[35]

An Environment G is ***Closed Clubwise Stable*** (henceforth, CCS) if the following conditions obtain:

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

(ii) No New Club Formation: $\forall m \subseteq N$:
$\exists i \in m : u_i(G + m, w, c) > u_i(G, w, c) \Rightarrow \exists j \in m : u_j(G + m, w, c) < u_j(G, w, c)$.

(iii) No Joining: $\forall s \in S, \forall i \notin N_G(s)$:
$u_i(G, w, c) \geq u_i(G + \{i, s\}, w, c)$ *OR* $\exists j \in N_G(s) : u_j(G, w, c) > u_j(G + \{i, s\}, w, c)$.

[34] The nontransferable social value model in Hout et al. (2013) is also a special case of the truncated model with club congestion and individual congestion wherein $h(2) = 1$, $\forall m > 2 : h(m) = 0$, $b(k) = \frac{\sqrt{V^s}}{k^\rho}$, $c = 0$ and $D = 1$. In addition, this version of our model can also accommodate other approaches to the individual's limited capacity for maintaining direct connections (e.g. Currarini et al. (2016) and Moody (2001)).

[35] The literature on the stability of coalition partitions and jurisdictions also explores various admission rules. The basic rule is free mobility (Tiebout (1956)) wherein each group of agents can freely move from one coalition to another. Well studied restrictions of free mobility in this context are exclusion rules and capacity thresholds (e.g. Greenberg (1979), Drèze and Greenberg (1980), Jehiel and Scotchmer (2001), Bogomolnaia and Jackson (2002), Scotchmer (2002) and Watts (2007)).

For an environment to be CCS, the unanimous agreement of incumbent club members is required in order to join a club. Since this requirement makes the joining deviation harder to execute, Open Clubwise Stability is a refinement of Closed Clubwise Stability.

Generally, admission of new members into the club induces both positive and negative externalities upon incumbent members. Positive externalities stem from new (or shorter) paths provided by the new member. Negative externalities stem from the effects of congestion. Clearly when there is no congestion, incumbents receive only positive externalities from admitting new members, and therefore they would never object it. This implies that in the baseline case of no congestion (Section 3) the OCS and CCS solution concepts coincide.

6.1 Closed Clubwise Stability: The Club Congestion Model

When there are no membership fees the set of CCS environments is the set of spanning super environments of the All Paired Environment.[36] Hence, the difference between Open Clubwise Stability and Closed Clubwise Stability exists only when there are positive membership fees and strict club congestion.

In order to demonstrate the difference between OCS and CCS we define the **Almost Grand Club Environment** as an environment in which there is exactly one populated club and all agents except for one are affiliated with it.[37]

For a given number of agents $n_a > 3$, let $h(\cdot)$ be a club congestion function such that $k_h(n_a - 1) > k_h(n_a) > \max_{k\in\{2,\ldots,n_a-2\}} k_h(k)$.[38] Consider the Almost Grand Club Environment when $k_h(n_a) > c > \max_{k\in\{2,\ldots,n_a-2\}} k_h(k)$. In this case, no agent would want to leave the populated club since $k_h(n_a - 1) > c$. No subset of agents would want to form a new club since $c > \max_{k\in\{2,\ldots,n_a-2\}} k_h(k)$.[39] Since $k_h(n_a) > c$, the isolated agent wishes to join the populated club and the Almost Grand Club Environment is not OCS. But, since $k_h(n_a - 1) > k_h(n_a)$ such a deviation will strictly harm incumbents of the populated club who lose out because the positive externalities (one new direct connection) are lower than the negative externalities (weaker direct connections to all other incumbents due to stronger club congestion).

[36] If Environment G is a spanning super environment of the All Paired Environment it is OCS and therefore also CCS. If Environment G is not a spanning super environment of the All Paired Environment, then there are two agents that are better off forming a new club of size 2 and G is not CCS.

[37] Formally, $\exists s \in S$ such that $n_G(s) = n_a - 1$, $\forall s' \in S\backslash\{s\} : n_G(s') = 0$ and $A = \cup_{i=1}^{n_a-1}\{i, s\}$.

[38] One example of a club congestion function that satisfies these properties is the exponential congestion function with $a = 0$ and $\delta = \frac{\sqrt{4n_a^2-16n_a+14}}{2(n_a-1)}$ (to see this, note that $\delta \in (1 - \frac{1}{n_a-2}, 1 - \frac{1}{n_a-1})$ and recall the logic used in the proof of Lemma 4).

[39] For every size k of the new club, the benefits for deviating agents are bounded from above by $k_h(k)$. Therefore $c > \max_{k\in\{2,\ldots,n_a-2\}} k_h(k)$ guarantees no deviations to clubs of size $n_a - 2$ or smaller. The maximal gain for an agent affiliated with the populated club from being involved in the formation of a new club of size $n_a - 1$ or n_a is bounded by $h(n_a - 1)$. Since $n_a > 3$, $c > k_h(n_a - 2) = (n_a - 3)h(n_a - 2) \geq h(n_a - 1)$. This means that $c > \max_{k\in\{2,\ldots,n_a-2\}} k_h(k)$ guarantees no deviations to form new clubs.

Hence, the Almost Grand Club Environment is CCS. Although we do not pursue a dynamic analysis of our setting wherein agents join the environment sequentially, it is intuitive that while OCS encourages integration (in this example, one big club), CCS may drive the environment toward segregation (uniform partition).

6.2 Closed Clubwise Stability: The Individual Congestion Model

In the Individual Congestion model, the externalities induced by admission of new members into a club depend upon the connections between incumbents and the agents desirous of admission prior to the entrance. Positive externalities arise from new and shorter paths provided to incumbents by new members. These externalities are high when prior to admission, new members are distant in the induced network. Upon admitting these new members, incumbents get closer to them and their locality. Negative externalities arise from the damage caused to existing connections due to the deterioration in the quality of the direct links maintained by the new members. These externalities are high when prior to admission new members are involved in many of the shortest paths originating from incumbents.

To demonstrate both the positive and negative externalities in the admission of new members and their effect upon stability in the Individual Congestion model, suppose n_a is odd and consider the $\frac{n_a+1}{2}$-Star Environment (denoted by G). This environment includes two populated clubs of equal size ($m = \frac{n_a+1}{2}$), Club s and Club t that have exactly one member, Agent i, in common ($N_G(s) \cap N_G(t) = \{i\}$). When Agent $j \neq i$ from Club s wishes to join Club t, each agent in Club t who is not a member of Club s would approve since their connection with Agent j improves from being indirect of value $b^2(1)b^2(2)$ to being direct of value $b(1)b(2)$. However, Agent i would oppose such an admission since the connection with Agent j deteriorates from $b(1)b(2)$ to $b^2(2)$. Thus, the $\frac{n_a+1}{2}$-Star Environment is not OCS but it is CCS.[40] This example, just like the previous one, demonstrates that while OCS leads to integration, the CCS internalizes the negative externalities of new member admissions and identify some segregated environments as stable in addition to the OCS environments characterized in Sections 4 and 5.

[40]For example, let $b(k) = \frac{1}{k+1}$ be the individual congestion function and let $c = \frac{1}{18}$. Agent i gains $\frac{1}{6}(n_a - 1)$ in G and only $\frac{1}{4}\frac{n_a-1}{2}$ in $G - \{i, s\}$ or $G - \{i, t\}$ so this agent would not want to abort any of her affiliations. Any other agent $j \neq i$ gains at least $\frac{1}{6} - \frac{1}{18}$ in G (from the link with Agent i) and zero if isolated. Thus, no agent would want to leave a club. Note that Agent i already has a direct link to all other agents, so joining a new club would produce a loss. Next, consider the most profitable new club for Agent $j \neq i$ that is affiliated with Club s. Such a club includes Agent j and all the members of Club t excluding Agent i. It is easy to see that the benefits of so doing sum to $\frac{1}{36}$ while the costs are $\frac{1}{18}$. Therefore even this attractive deviation is not worthwhile. Hence, no subset of agents would want to form a new club. It is also easy to show that by joining Club t the utility of Agent j would increase by $\frac{n_a-2}{36}$ (which is positive since $n_a > 2$). Hence, the $\frac{n_a+1}{2}$-Star Environment is not OCS. However, Agent i (who is a member of Club t) would oppose this admission since it decreases that agent's utility by $\frac{1}{18}$. Hence, the $\frac{n_a+1}{2}$-Star Environment is CCS.

7 Clustering

A well-known real-life phenomenon in social networks is that they are characterized by high clustering. That is, in most real-life networks the probability of two individuals who share a common neighbor to be connected is much higher than would be expected if the connections had formed randomly (see Goyal (2007) and Jackson (2008)). High clustering affects the spread of information and therefore access to jobs, ideas and other resources. As a result, clustering has become a central topic of interest for social networks researchers.

Social sciences literature (see Rivera et al. (2010) for a recent survey) frequently attributes high clustering in social networks to one of two explanations. First, individual preference for connections with individuals with whom a shared connection already exists. Termed "preference for transitivity," it can be based on various motives, such as reduced uncertainty, improved monitoring, conflict mitigation and minimization of opportunism (see Heider (1946), Cartwright and Harary (1956), Coleman (1988) and Hummon and Doreian (2003)). Another explanation based on homophily describes an environment wherein individuals prefer to link to individuals with whom they share common social traits such as race, gender, country of origin, etc (see McPherson et al. (2001)).

A relatively recent body of literature attempts to provide econometric tools for estimating various network formation models that incorporate homophily, preference for transitivity and state dependence in links.[41] A growing concern in this literature is that neglect of self-selection into social contexts leads to over-estimation of the importance of homophily and preference for transitivity in the process of network formation (see Rivera et al. (2010), Currarini et al. (2010) and Miyauchi (2016)).

Indeed, we believe that our setting provides a third explanation for the high clustering observed in real-life networks. Since every pair of agents who share a club is connected in the induced network, the affiliation portfolios chosen by individuals induce a social network composed of a collection of cliques. Therefore, in our framework, a network induced by non-trivial clubs (e.g. of size greater than 2) must exhibit high local clustering since an individual's neighbors form a tightly knit group (see also Bar (2005)).[42] Hence, we propose considering clubs as linking platforms rather than individuals' linking preferences as the

[41] One of the main challenges of this literature is the treatment of homophily on unobservables. Goldsmith-Pinkham and Imbens (2013) introduce homophily on unobservables by assuming that the relevant unobservables are binary and distributed independently of all observables. In Mele (2017) individuals are partitioned exogenously to unobserved communities and they exhibit preference for transitivity only within these communities. Graham (2015, 2016) proposes to exploit the fact that homophily is independent of network structure. See the discussion in Jackson (2014).

[42] Jackson et al. (2012) point out that such environments are characterized also by high support. That is, the fraction of links which are part of a triad is high. Latapy et al. (2008) suggest clustering indices for affiliation networks based on intersections of affiliation portfolios.

fundamental that drives the high clustering observed in real-life networks.[43]

8 Concluding comments

The focus of our paper is on the formation of social networks based on the endogenous formation of social clubs. Our analysis relies on two important assumptions: clubs (beside their size) and agents are homogeneous. A more complete picture of the social architecture may include the endogenous formation of a variety of clubs that may differ in membership costs, quality of induced links and rules of entry, exit and formation. For example, in some clubs the interaction among members may be more intense than in others and as a result they may differ in their congestion functions and membership fees (e.g. Young and Larson (1965a)). We also assume that individuals are homogeneous. In a model of heterogeneous agents the weight of each link may depend on the identity of the agents (e.g. Bruggeman (2016)) and may also be asymmetric. The authors' benefit derived from being connected to Leo Messi is probably different from the benefit he derives from being connected to us. These benefits affect the attractiveness of different clubs and their composition. As a result they also affect the stable social environments that may emerge from our endogenous affiliation setting.

The concept of clubs as platforms for link formation can be applied also to non-social contexts. For example, consider the environment of open source software development wherein individuals work on different R&D projects. A project can be viewed as a social club and a developer can be viewed as a club of projects (See Fershtman and Gandal (2011)). A similar view can be taken of countries that participate in multi-lateral trade agreements, interlocking corporate directorates (e.g. Mizruchi (1996)), standardization committees (Bar and Leiponen (2008) and Leiponen (2008)), etc.

Finally, a Matlab package (including a detailed user manual) that provides tools for exploring stability and efficiency in the Club Congestion model, the Individual Congestion model and the model that includes both types of congestion, accompanies this paper. The package can be found on the second author's website.

[43] Obviously, in some contexts these explanations coexist. One example is the discussion in Currarini et al. (2009) on social clubs as the platform on which matching biases (as homophily) evolve. Another example is Kossinets and Watts (2006) who track emails of students, faculty, and staff at a large research university over an academic year. They find that among students who do not share a common class, having a mutual contact increase the probability of communication by 140 times. However, if students do share a class they were only 3 times more likely to begin communicating if they shared a common correspondent. Datasets that contain both club affiliations and the social network, as in the one studied by Kossinets and Watts (2006) (or by Young and Larson (1965a)), may enable researchers to disentangle these three explanations (see also the discussion in Feld (1981)). The method suggested in Chandrasekhar and Jackson (2017) for the estimation of network formation models, can be interpreted as an econometric analysis of network data where club affiliations are assumed but not observed (see Section 3 therein).

References

Aral, Sinan (2016) "The future of weak ties," *American Journal of Sociology*, Vol. 121, No. 6, pp. 1931–1939.

Arnold, Tone and Myrna Wooders (2005) "Dynamic Club Formation With Coordination," Working Paper 05-W22.

Bala, Venkatesh and Sanjeev Goyal (2000) "A Noncooperative Model of Network Formation," *Econometrica*, Vol. 68, pp. 1181–1229.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou (2006) "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, Vol. 74, No. 5, pp. 1403–1417.

Bar, Talia (2005) "Coalition Networks and Induced Graphs," *Cornell Department of Economics Working Paper*, Vol. 0, p. 0.

Bar, Talia and Aija Leiponen (2008) "Committees and Networking in Standard Setting," Working Paper.

Baumann, Leonie (2017) "A Model of Weighted Network Formation."

Berge, Claude (1989) *Hypergraphs: Combinatorics of Finite Sets*, Vol. 45 of North-Holland Mathematical Library: Elsevier Science Publishers B. V.

Blau, Peter M and Joseph E Schwartz (1984) "Crosscutting social circles."

Bloch, Francis and Bhaskar Dutta (2009) "Communication Networks with Endogenous Link Strength," *Games and Economic Behavior*, Vol. 66, pp. 39–56.

Bogomolnaia, Anna and Matthew O Jackson (2002) "The stability of hedonic coalition structures," *Games and Economic Behavior*, Vol. 38, No. 2, pp. 201–230.

Bonacich, Phillip (1972) "Technique for Analyzing Overlapping Memberships," *Sociological Methodology*, Vol. 4, pp. 176–185.

——— (1978) "Using Boolean Algebra to Analyze Overlapping Memberships," *Sociological Methodology*, Vol. 9, pp. 101–115.

Bonacich, Phillip, Annie Cody Holdren, and Michael Johnston (2004) "Hyper-edges and multidimensional centrality," *Social networks*, Vol. 26, No. 3, pp. 189–203.

Boorman, Scott A (1975) "A combinatiorial optimization model for transmission of job information through contact networks," *The bell journal of economics*, pp. 216–249.

Borgatti, Stephen P. and Martin G. Everett (2013) "The Dual-Projection Approach for Two-Mode Networks," *Social Networks*, Vol. 35, pp. 204–210.

Borgs, Christian, Jennifer Chayes, Jian Ding, and Brendan Lucier (2011) "The hitchhiker's guide to affiliation networks: A game-theoretic approach," in Bernard Chazelle ed. *Innovations in Computer Science - ICS 2011, Tsinghua University, Beijing, China*: Tsinghua University Press, pp. 389–400.

Boxman, Ed and Henk Flap (2000) "Getting started: the influence of social capital on the start of the occupational career," in Nan Lin, Karen S. Cook, and Ronald S. Burt eds. *Social capital: Theory and Research.*: New York: Aldine de Gruyter, pp. 159–181.

Breiger, Ronald L. (1974) "The Duality of Persons and Groups," *Social Forces*, Vol. 53, No. 2, pp. 181–190.

Brueckner, Jan K. (2006) "Friendship Networks," *Journal of Regional Science*, Vol. 46, No. 5, pp. 847–865.

Bruggeman, Jeroen (2016) "The strength of varying tie strength: Comment on Aral and Van Alstyne," *American Journal of Sociology*, Vol. 121, No. 6, pp. 1919–1930.

Buchanan, James M. (1965) "An Economic Theory of Clubs," *Economica, New Series*, Vol. 32, No. 125, pp. 1–14.

Cabrales, Antonio, Antoni Calvó-Armengol, and Yves Zenou (2011) "Social interactions and spillovers," *Games and Economic Behavior*, Vol. 72, No. 2, pp. 339–360.

Calvó-Armengol, Antoni (2004) "Job contact networks," *Journal of economic Theory*, Vol. 115, No. 1, pp. 191–206.

Calvó-Armengol, Antoni and Yves Zenou (2003) "Does crime affect unemployment? The role of social networks," *Annales d'Economie et de Statistique*, pp. 173–188.

Cartwright, Dorwin and Frank Harary (1956) "Structural balance: a generalization of Heider's theory.," *Psychological review*, Vol. 63, No. 5, p. 277.

Caulier, Jean-François, Ana Mauleon, José J. Sempere-Monerris, and Vincent Vannetelbosch (2013a) "Stable and Efficient Coalitional Networks," *Review of Economic Design*, Vol. 17, pp. 249–271.

Caulier, Jean-François, Ana Mauleon, and Vincent Vannetelbosch (2013b) "Contractually Stable Networks," *International Journal of Game Theory*, Vol. 42, No. 2, pp. 483–499.

——— (2015) "Allocation rules for coalitional network games," *Mathematical Social Sciences*, Vol. 78, pp. 80–88.

Chandrasekhar, A. and M. O. Jackson (2017) "A Network Formation Model Based on Subgraphs," *ArXiv e-prints*.

Christakis, Nicholas A, James H Fowler, Guido W Imbens, and Karthik Kalyanaraman (2010) "An empirical model for strategic network formation."

Coleman, James S (1988) “Social Capital in the Creation of Human Capital,” *American Journal of Sociology*, pp. S95–S120.

Currarini, Sergio, Matthew O Jackson, and Paolo Pin (2009) “An economic model of friendship: Homophily, minorities, and segregation,” *Econometrica*, Vol. 77, No. 4, pp. 1003–1045.

——— (2010) “Identifying the roles of race-based choice and chance in high school friendship network formation,” *Proceedings of the National Academy of Sciences*, Vol. 107, No. 11, pp. 4857–4861.

Currarini, Sergio, Jesse Matheson, and Fernando Vega-Redondo (2016) “A simple model of homophily in social networks,” *European Economic Review*, Vol. 90, pp. 18–39.

Deroïan, Frédéric (2009) “Endogenous link strength in directed communication networks,” *Mathematical Social Sciences*, Vol. 57, No. 1, pp. 110–116.

Drèze, Jacques H. and Joseph Greenberg (1980) “Hedonic Coalitions: Optimality and Stability,” *Econometrica*, Vol. 48, No. 4, pp. 987–1003.

Faust, Katherine (1997) “Centrality in affiliation networks,” *Social networks*, Vol. 19, No. 2, pp. 157–191.

Feld, Scott L (1981) “The focused organization of social ties,” *American journal of sociology*, Vol. 86, No. 5, pp. 1015–1035.

Fershtman, Chaim and Neil Gandal (2011) “Direct and Indirect Knowledge Spillovers: the ‘Social Network‘ of Open-Source Projects,” *RAND Journal of Economics*, Vol. 42, No. 1, pp. 70–91.

Goldsmith-Pinkham, Paul and Guido W Imbens (2013) “Social networks and the identification of peer effects,” *Journal of Business & Economic Statistics*, Vol. 31, No. 3, pp. 253–264.

Goyal, Sanjeev (2005) “Strong and Weak Links,” *Journal of the European Economic Association*, Vol. 3, No. 2-3, pp. 608–616.

——— (2007) *Connections: An Introduction to the Economics of Networks*: Princeton University Press.

Goyal, Sanjeev, Aleander Konovalov, and José Luis Moraga-González (2008) “Hybrid R&D,” *Journal of the European Economic Association*, Vol. 6, No. 6, pp. 1309–1338.

Graham, Bryan S (2015) “Methods of identification in social networks,” *Annu. Rev. Econ.*, Vol. 7, No. 1, pp. 465–485.

——— (2016) “Homophily and Transitivity in Dynamic Network Formation,” Working Paper.

Granovetter, Mark (1973) "The Strength of Weak Ties," *The American Journal of Sociology*, Vol. 78, No. 6, pp. 1360–1380.

——— (1983) "The Strength of Weak Ties: A Network Theory Revisited," *Sociological Theory*, Vol. 1, pp. 201–233.

Greenberg, Joseph (1979) "Stability when mobility is restricted by the existing coalition structure," *Journal of Economic Theory*, Vol. 21, No. 2, pp. 213–221.

Heider, Fritz (1946) "Attitudes and cognitive organization," *The Journal of psychology*, Vol. 21, No. 1, pp. 107–112.

Harmsen-van Hout, Marjolein JW, P Jean-Jacques Herings, and Benedict GC Dellaert (2013) "Communication network formation with link specificity and value transferability," *European Journal of Operational Research*, Vol. 229, No. 1, pp. 199–211.

Hummon, Norman P and Patrick Doreian (2003) "Some dynamics of social balance processes: bringing Heider back into balance theory," *Social Networks*, Vol. 25, No. 1, pp. 17–49.

Jackson, Matthew O. (2008) *Social and Economic Networks*: Princeton University Press.

Jackson, Matthew O (2014) "Networks in the understanding of economic behaviors," *The Journal of Economic Perspectives*, Vol. 28, No. 4, pp. 3–22.

Jackson, Matthew O, Tomas Rodriguez-Barraquer, and Xu Tan (2012) "Social capital and social quilts: Network patterns of favor exchange," *American Economic Review*, Vol. 102, No. 5, pp. 1857–97.

Jackson, Matthew O and Asher Wolinsky (1996) "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, Vol. 71, No. 1, pp. 44–74.

Jehiel, Philippe and Suzanne Scotchmer (2001) "Constitutional Rules of Exclusion in Jurisdiction Formation," *The Review of Economic Studies*, Vol. 68, No. 2, pp. 393–413.

Jun, Tackseung and Jeong-Yoo Kim (2009) "Hypergraph formation game," *Hitotsubashi Journal of Economics*, pp. 107–122.

Kadushin, Charles (1966) "The Friends and Supporters of Psychotherapy: on Social Circles in Urban Life," *American Sociological Review*, Vol. 31, No. 6, pp. 786–802.

——— (2011) *Understanding Social Networks: Theories, Concepts, and Findings*: Oxford University Press.

Kossinets, Gueorgi and Duncan J Watts (2006) "Empirical analysis of an evolving social network," *science*, Vol. 311, No. 5757, pp. 88–90.

Kramarz, Francis and Oskar Nordström Skans (2014) "When strong ties are strong: Networks and youth labour market entry," *Review of Economic Studies*, Vol. 81, No. 3, pp. 1164–1200.

Latapy, Matthieu, Clémence Magnien, and Nathalie Del Vecchio (2008) "Basic Notions for the Analysis of Large Two-Mode Networks," *Social Networks*, Vol. 30, pp. 31–48.

Leiponen, Aija Elina (2008) "Competing Through Cooperation: The Organization of Standard Setting in Wireless Telecommunications," *Management Science*, Vol. 54, No. 11, pp. 1904–1919.

Maccoby, Eleanor E (1998) *The two sexes: Growing up apart, coming together*, Vol. 4: Harvard University Press.

McPherson, J Miller and Lynn Smith-Lovin (1982) "Women and weak ties: Differences by sex in the size of voluntary organizations," *American Journal of Sociology*, Vol. 87, No. 4, pp. 883–904.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook (2001) "Birds of a feather: Homophily in social networks," *Annual review of sociology*, pp. 415–444.

Mele, Angelo (2017) "A Structural Model of Homophily and Clustering in Social Networks."

Miyauchi, Yuhei (2016) "Structural estimation of pairwise stable networks with nonnegative externality," *Journal of Econometrics*, Vol. 195, No. 2, pp. 224–235.

Mizruchi, Mark S (1996) "What Do Interlocks Do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates," *Annual Review of Sociology*, Vol. 22, No. 1996, pp. 271–298.

Montgomery, James D (1992) "Job search and network composition: Implications of the strength-of-weak-ties hypothesis," *American Sociological Review*, pp. 586–596.

——— (1994) "Weak ties, employment, and inequality: An equilibrium analysis," *American Journal of Sociology*, pp. 1212–1236.

Moody, James (2001) "Race, school integration, and friendship segregation in America," *American journal of Sociology*, Vol. 107, No. 3, pp. 679–716.

Moody, James and Douglas R. White (2003) "Structural Cohesion and Embeddedness: A Hierarchical Concept of Social Groups," *American Sociological Review*, Vol. 68, No. 1, pp. 103–127.

Myerson, Roger B. (1980) "Conference Structures and Fair Allocation Rules," *International Journal of Game Theory*, Vol. 9, No. 3, pp. 169–182.

Newman, Mark EJ (2001) "Scientific collaboration networks. I. Network construction and fundamental results," *Physical review E*, Vol. 64, No. 1, p. 016131.

Page, Frank H. and Myrna Wooders (2007) "Networks and Clubs," *Journal of Economic Behavior and Organization*, Vol. 64, pp. 406–425.

Persico, Nicola, Andrew Postlewaite, and Dan Silverman (2004) “The effect of adolescent experience on labor market outcomes: The case of height,” *Journal of Political Economy*, Vol. 112, No. 5, pp. 1019–1053.

Renoust, Benjamin, Guy Melançon, and Marie-Luce Viaud (2014) “Entanglement in multiplex networks: understanding group cohesion in homophily networks,” in Missaoui R. and Sarr I. eds. *Social Network Analysis-Community Detection and Evolution*: Springer, pp. 89–117.

Rivera, Mark T, Sara B Soderstrom, and Brian Uzzi (2010) “Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms,” *annual Review of Sociology*, Vol. 36, pp. 91–115.

Rogers, Brian W. (2006) “A Strategic Theory of Network Status,” Working paper.

Salonen, Hannu (2015) “Reciprocal equilibria in link formation games,” *AUCO Czech Economic Review*, Vol. 9, No. 3, p. 169.

——— (2016) “Equilibria and centrality in link formation games,” *International Journal of Game Theory*, Vol. 45, No. 4, pp. 1133–1151.

Sandler, Todd and John T. Tschirhart (1997) “Club Theory: Thirty Years Later,” *Public Choice*, Vol. 93, pp. 335–355.

Scotchmer, Suzanne (2002) “Local Public Goods and Clubs,” *Handbook of public economics*, Vol. 4, pp. 1997–2042.

Simmel, Georg (1908/1955) “The Web of Group-Affiliations,” in *Conflict and the Web of Group-Affiliations*: The Free Press. (Translated by Reinhard Bendix in 1955. Original appeared in “Soziologie” published by Duncker and Humblot, Berlin, 1908).

Slikker, Marco and Anne Van den Nouweland (2001) *Social and Economic Networks in Cooperative Game Theory*: Kluwer Academic Publishers.

Snijders, Tom AB, Philippa E Pattison, Garry L Robins, and Mark S Handcock (2006) “New specifications for exponential random graph models,” *Sociological methodology*, Vol. 36, No. 1, pp. 99–153.

So, Chiu Ki (2016) “Network Formation with Endogenous Link Strength and Decreasing Returns to Investment,” *Games*, Vol. 7, No. 4, p. 40.

So, Chiu Ki, Vai-lam Mui, and Birendra Rai (2015) “Event Subscription and Non-cooperative Network Formation,” Working paper.

Tiebout, Charles M. (1956) “A Pure Thoery of Local Expenditures,” *The Journal of Political Economy*, Vol. 64, No. 5, pp. 416–424.

Wasserman, Stanley and Katherine Faust (1994) *Social Network Analysis*: Cambridge University Press.

Watts, Alison (2007) "Formation of segregated and integrated groups," *International Journal of Game Theory*, Vol. 35, No. 4, pp. 505–519.

Young, Ruth C and Olaf F Larson (1965a) "The contribution of voluntary organizations to community structure," *American Journal of Sociology*, Vol. 71, No. 2, pp. 178–186.

Young, Ruth C. and Olaf F. Larson (1965b) "A New Approach to Community Structure," *American Sociological Review*, Vol. 30, No. 6, pp. 926–934.

Zhou, Tao, Jie Ren, Matúš Medo, and Yi-Cheng Zhang (2007) "Bipartite network projection and personal recommendation," *Physical Review E*, Vol. 76, No. 4, p. 046115.

Appendix

A Proofs

A.1 Lemma 1

Proof. Obviously, $G - \{i, s\}$ contains one component that includes Agent i. Since G is minimally connected, this environment contains at least one additional component.
Suppose C and C' are two components of $G - \{i, s\}$ that do not include Agent i. Let Agent j be a member of C and Agent j' be a member of C'. Since G is connected, there is a path between j and j' in G and since they reside in different components in $G - \{i, s\}$, there is no path between them in $G - \{i, s\}$.
Hence, the path between j and j' in G must include the affiliation of Agent i in Club s. Therefore, this path must be of the form $\{j, \ldots, k, i, k', \ldots, j'\}$ where either k is affiliated with Club s or k' is affiliated with Club s but not both.
With no loss of generality, assume that k is affiliated with Club s while k' is not affiliated with Club s. Therefore, agents i and k' share a club in G which is not s. Hence, i and k' share a club in $G - \{i, s\}$. But, then there is a path between Agent i and Agent j' in $G - \{i, s\}$ in contradiction to C' being a component of $G - \{i, s\}$ that does not include Agent i. Thus, $G - \{i, s\}$ contains exactly two components. $\square$

A.2 Proposition 1

Proof. Suppose $n_a - 1 > c > 0$. First, consider the case where G is a Minimally Connected environment of class $K(G) \geq c$. Since G is connected no agent can benefit by joining a club or by forming a new club. Also, by leaving a club, every agent loses connection to at least $K(G)$ agents while gaining the membership fees. Since $K(G) \geq c$, the agent can not gain by leaving a club. Therefore, G is OCS.
Next, consider the case where G is not a Minimally Connected environment of class $K(G) \geq c$. If G is the Empty environment then consider the deviation where all the agents form a new club. The benefit for each individual is $n_a - 1 - c > 0$. Hence, the Empty environment

is not OCS.
If G is a non-empty disconnected environment then there must exist a component H that contains $h > 1$ agents. The maximal possible utility of an agent in H is $(h-1)-c$. If $c > h-1$ then every member of this component would like to leave any of her clubs. If $c \leq h-1$ then any agent that is not in H can improve if she joins one of H's clubs since she gets $h-c > 0$. Hence, no disconnected environment is OCS.
If G is connected, but not minimally connected, there is an affiliation that can be removed while leaving the induced network connected. Denote this affiliation by $\{i,s\}$. Then, Agent i, by leaving club s can improve his net utility by c. Hence, no connected, but not minimally connected, environment is OCS.
Next, suppose that G is a minimally connected network of class $K(G) < c$. Consider an affiliation $\{i^\star, s^\star\} \in \arg \min_{\{i,s\}\in A} n(C_{-i}(G-\{i,s\}))$. Agent $i^\star$ wishes to leave club $s^\star$ since while losing the connection to $K(G)$ agents she gains c, and $K(G) < c$. Hence, if G is not a Minimally Connected environment of class $K(G) \geq c$ then G is not OCS. This completes the proof of Part 1a.
Since $n_a - 1 > c$, the maximal utility an agent can obtain is $u_i(G) = n_a - 1 - c$. In the Grand Club Environment every agent achieves the maximal utility. Therefore, the Grand Club Environment is PE. Moreover, any other environment is either disconnected or it contains at least one agent that maintains multiple affiliations. In both cases there is at least one individual with a utility lower than $n_a - 1 - c$. Therefore, the Grand Club Pareto dominates any other environment. Hence, the Grand Club Environment is the unique PE environment. Therefore, the Grand Club Environment is also the unique SE environment.
Last, suppose that $c > n_a - 1$. Every agent that maintains memberships in a non-empty environment, wishes to leave any of her clubs. Therefore, every non-empty environment G is not OCS. Moreover, none of the agents in those environments have positive utility and there is at least one with negative utility. However, in the Empty environment, no coalition of agents is better off by establishing a new club. Therefore, the Empty Environment is OCS. Since all the agents have zero utility, it is also PE and SE. Hence, the Empty environment is the unique OCS, PE and SE environment. □

A.3 Lemma 2

Proof. Suppose that Environment G is Pareto Efficient but does not satisfy the "No New Club Formation" condition. Thus, there exists a coalition $m \subseteq N$ such that for every member i of m, $u_i(G+m,w,c) \geq u_i(G,w,c)$ and for at least one of them the inequality is strict. Note that the weighted network induced by $G+m$ (henceforth, g_{G+m}) differs from the weighted network induced by G (henceforth, g_G) only in links where both end agents belong to m. Moreover, if $i,j \in m$ share a smaller club in G then the link between them and its weight are the same in g_{G+m} and in g_G. Thus, the differences between g_{G+m} and g_G are only in links between members of m that share bigger clubs in G or share no clubs at all in G. Therefore, g_{G+m} may have links that do not exists in g_G or links with higher weights then in G. This implies that the direct links of every agent $k \in N \backslash m$ are the same in g_{G+m} and in g_G. Moreover, her indirect links can only get shorter in g_{G+m} compared to g_G (either due to new links or to higher weights). Since the costs of the agents in $N \backslash m$ are identical in G and in $G+m$, $\forall k \in N \backslash m : u_k(G+m,w,c) \geq u_k(G,w,c)$. Therefore, for every $i \in N$, $u_i(G+m,w,c) \geq u_i(G,w,c)$ and for at least one of them the inequality is strict. Hence, G is not Pareto Efficient, contradiction. Therefore, if Environment G is Pareto Efficient it satisfies the "No New Club Formation" condition. □

A.4 Lemma 3

Proof. By the definition of the club congestion function, the weight of each link is determined by a single club - the smallest club the two end agents share. Since the benefit part of the agent's preferences depends only on the weights in the induced network, there may be at most $n_a - 1$ affiliations that determine the individual's benefits. Similarly, at most $\frac{n_a(n_a-1)}{2}$ clubs contribute to the benefits of any agent in the environment. However, all populated clubs contribute to the costs part of the preferences, since each affiliation is costly ($c > 0$). Hence, every G that includes an individual that maintains more than $n_a - 1$ affiliations is not OCS. Moreover, if G includes more than $\frac{n_a(n_a-1)}{2}$ populated clubs, there is at least one club that does not contribute to the benefits of any of its members. Therefore, each one of its members would wish to cancel this affiliation and G is not OCS. □

A.5 Lemma 4

Proof. By definition, $h(m)$ is inelastic if and only if $\forall m \in \{2,\ldots,n_a-1\} : \eta_h(m) > -1$. $\eta_h(m) > -1$ if and only if $\frac{m \times h(m+1)}{h(m)} - m > -1$. Therefore $h(m)$ is inelastic if and only if $k_h(m+1) = m \times h(m+1) > (m-1) \times h(m) = k_h(m)$. Hence, $h(m)$ is inelastic if and only if $k_h(m)$ is strictly increasing. Similar argument shows that $h(m)$ is elastic if and only if $k_h(m)$ is strictly decreasing. □

A.6 Proposition 2

A.6.1 Lemma 5

Lemma 5. *If $c > 0$ and the Club Congestion function is $h(2) = \delta$ and $\forall m > 2 : h(m) = 0$ then $\forall g \in \mathbb{G}_n, \forall i \in N : u_i(G_g) = u_i^{JW}(g)$.*

Proof. Note that for every un-weighted network $g =< N, E >$, the induced network of G_g denoted by $\bar{g} =< N, \bar{E}, W >$ is such that $\bar{E} = E$ and, by the choice of $h(\cdot)$, each link has a weight of δ since the clubs are all of size two.
Since all the weights in $\bar{g}$ are the same, the length of the shortest weighted path between agents i and j in $\bar{g}$ is the same as the length of the shortest path between them in g. Therefore, the distance between agents i and j in $\bar{g}$ equals $\delta^{d_{ij}}$ where d_{ij} is the geodesic distance between agents i and j in $\bar{g}$ and therefore also in g. As a consequence, the benefits of the agents in G_g equal their benefits in g.
Moreover, by construction, the number of direct links each agent maintains in g equals the number of her affiliations in $G_{\bar{g}}$. Therefore, the costs of the agents in G_g equal their costs in g. Hence, $\forall i \in N : u_i(G_g) = u_i^{JW}(g)$. □

A.6.2 Lemma 6

Lemma 6. *Let* $g =< N, E >$ *be an un-weighted network and let* $G_g =< N, S, A >$ *be the corresponding environment.* $\forall i, j \in N$ *such that* $\exists s \in S : \{\{i, s\}, \{j, s\}\} \subseteq A$ *then* $u_i(G_{g-\{i,j\}}) = u_i(G_g - \{i, s\})$.

Proof. By construction, the Environment $G_{g-\{i,j\}}$ includes the same clubs as G_g excluding Club s. Therefore its induced weighted network $\bar{g}_{-s}$ is identical to g excluding the link between agents i and j. Denote the benefits of Agent i in $\bar{g}_{-s}$ by B. Then $u_i(G_{g-\{i,j\}}) = B - (s_{G_g}(i) - 1) \times c$.
Environment $G_g - \{i, s\}$ includes the same clubs as G_g, but the affiliation of Agent i in Club s is dropped. Since Club s is a singleton in $G_g - \{i, s\}$, it induces no links. Therefore, $\bar{g}_{is}$, the weighted network induced by $G_g - \{i, s\}$ is identical to $\bar{g}_{-s}$. Hence, the benefits of Agent i in $\bar{g}_{is}$ are B and $u_i(G_g - \{i, s\}) = B - (s_{G_g}(i) - 1) \times c$.[44] Thus, we get $u_i(G_{g-\{i,j\}}) = u_i(G_g - \{i, s\})$. □

A.6.3 The Proof

Proof. We suppose that $g \in PS(\delta, c, n)$ and show that $G_g \in OCS(c, n, h)$. The "No Joining" condition holds since the utility from a club of size 3 is zero while the participation fees are positive. For the same reason, no coalition of size greater than two wishes to form a new club.
Next, consider two agents, i and j, that do not share a club in G_g. Then, by construction, Agent i and Agent j are not linked in g. Since g is pairwise stable, if $u_i^{JW}(g) < u_i^{JW}(g + \{i, j\})$ then $u_j^{JW}(g) > u_j^{JW}(g + \{i, j\})$. By Lemma 5, if $u_i(G_g) < u_i(G_{g+\{i,j\}})$ then $u_j(G_g) > u_j(G_{g+\{i,j\}})$. Denote by m_{ij} the coalition that includes only agents i and j. Then, note that $G_{g+\{i,j\}}$ is identical to $G_g + m_{ij}$ since both denote the addition of Club s that includes agents i and j to Environment G_g. Hence, if $u_i(G_g) < u_i(G_g + m_{ij})$ then $u_j(G_g) > u_j(G_g + m_{ij})$. Therefore, no coalition of size two wishes to form a new club and the "No New Club Formation" condition holds.
For the "No Leaving" condition, consider Agent i that participates, together with Agent j, in Club s in G_g. Then, by construction, Agent i and Agent j are linked in g. Since g is pairwise stable $u_i^{JW}(g) \geq u_i^{JW}(g - \{i, j\})$. By Lemma 5, $u_i(G_g) \geq u_i(G_{g-\{i,j\}})$. By Lemma 6, $u_i(G_g) \geq u_i(G_g - \{i, s\})$, meaning that this condition also holds. Therefore, $G_g \in OCS(c, n, h)$.
For the other direction, we suppose that $G_g \in OCS(c, n, h)$ and show that $g \in PS(\delta, c, n)$. First, consider Agent i that is linked with Agent j in g. By construction Agent i participates, together with Agent j, in Club s in G_g. Since G_g is OCS, $u_i(G_g) \geq u_i(G_g - \{i, s\})$. By Lemma 6, $u_i(G_g) \geq u_i(G_{g-\{i,j\}})$. By Lemma 5, $u_i^{JW}(g) \geq u_i^{JW}(g - \{i, j\})$, meaning that no agent wishes to discard an existing link. Next, consider two agents, i and j, that are not linked in g. By construction agents i and j do not share a club in G_g. Since G_g is OCS, if $u_i(G_g) < u_i(G_g + m_{ij})$ then $u_j(G_g) > u_j(G_g + m_{ij})$. But, as mentioned above, $G_{g+\{i,j\}}$ is identical to $G_g + m_{ij}$. Therefore, if $u_i(G_g) < u_i(G_{g+\{i,j\}})$ then $u_j(G_g) > u_j(G_{g+\{i,j\}})$. By Lemma 5, if $u_i^{JW}(g) < u_i^{JW}(g + \{i, j\})$ then $u_j^{JW}(g) > u_j^{JW}(g + \{i, j\})$, meaning that no pair

[44] The single difference in agents' utilities between Environment $G_{\bar{g}-\{i,j\}}$ and Environment $G_{\bar{g}} - \{i, s\}$ is that in the latter the costs of Agent j are $n_{G_{\bar{g}}}(j) \times c$ while in the former her costs reduce to $(n_{G_{\bar{g}}}(j) - 1) \times c$.

of agents wishes to form a new link. Therefore, $g \in PS(\delta, c, n)$.
For the second part note that since we assume that $\mathcal{G}_n$ includes only environments with distinct clubs, every environment $G \in \mathcal{G}_n \backslash \mathcal{G}_{\bar{\mathbb{G}}n}$ includes at least one populated club of size greater than two. However, every agent that participates in a club of size greater than two wishes to leave the club since its benefits are zero (all induced links of such club are of weight zero) while the membership fees are positive. Therefore, $G \notin OCS(c, n, h)$. □

A.7 Proposition 3

Proof. Throughout the proof we assume that $\frac{n_a-1}{m-1}$ and $\frac{n_a(n_a-1)}{m(m-1)}$ are integers. Consider first the maximal sum of utilities of a connected environment $G \in \mathcal{G}_n^m$ with at most $\frac{n_a(n_a-1)}{m(m-1)}$ clubs. By Proposition 2 in Berge (1989), the minimal number of clubs in a connected m-Uniform Environment is $\frac{n_a-1}{m-1}$. Denote the number of clubs by $\frac{n_a(n_a-1)}{m(m-1)} \geq k \geq \frac{n_a-1}{m-1}$. The maximal total number of direct connections across all agents in the environment is $km(m-1)$ and their value is $km(m-1)h(m)$. Since the induced network is connected, the number of indirect connections is $n_a(n_a-1)-km(m-1)$ and their maximal value is $[n_a(n_a-1)-km(m-1)]h^2(m)$. The total cost of participation in k clubs of size m is kmc. Thus, the maximal sum of utilities of a connected m-Uniform Environment with k clubs is

$$km(m-1)h(m) + [n_a(n_a-1) - km(m-1)]h^2(m) - kmc$$

An m-Complete Environment includes $\frac{n_a(n_a-1)}{m(m-1)}$ clubs (each club generates $\frac{m(m-1)}{2}$ links out of the $\frac{n_a(n_a-1)}{2}$ possible links and each link is generated exactly once). Therefore, the sum of utilities of an m-Complete Environment is

$$n_a(n_a-1)h(m) - n_a\frac{n_a-1}{m-1}c$$

The difference between these two expressions is

$$[km(m-1) - n_a(n_a-1)]h(m) + [n_a(n_a-1) - km(m-1)]h^2(m) - [km - n_a\frac{n_a-1}{m-1}]c$$

Or,

$$[km(m-1) - n_a(n_a-1)][h(m) - h^2(m) - \frac{c}{m-1}]$$

Obviously, there must be a k such that the maximal total sum of utilities of a connected m-Uniform Environment is greater than the total sum of utilities of an m-Complete Environment. Therefore, this difference must be non-negative for some k. If $c < (m-1)[h(m) - h^2(m)]$, $[km(m-1) - n_a(n_a-1)]$ must be non-negative, meaning that $k \geq \frac{n_a(n_a-1)}{m(m-1)}$. Since $\frac{n_a(n_a-1)}{m(m-1)} \geq k \geq \frac{n_a-1}{m-1}$ it must be that $k = \frac{n_a(n_a-1)}{m(m-1)}$ and the difference is zero. Also, if $c = (m-1)[h(m) - h^2(m)]$ the difference is zero. Therefore, the m-Complete Environment achieves the maximal sum of utilities of the set of connected m-Uniform environments with at most $\frac{n_a(n_a-1)}{m(m-1)}$ clubs when $c \leq (m-1)[h(m) - h^2(m)]$.

Next, let G' be some m-Uniform environment with $k > \frac{n_a(n_a-1)}{m(m-1)}$ populated clubs. Each agent $i \in \{1, \ldots, n_a\}$ in G' gets at most $(n_a - 1)h(m)$ (in case she is directly connected to all other agents). Therefore, the total benefits in G' are at most $n_a(n_a - 1)h(m)$ while the total membership fees are $kmc > \frac{n_a(n_a-1)}{(m-1)}c$. Hence, the total sum of utilities of an m-Complete Environment is weakly greater than the total sum of utilities of any m-Uniform environment with $k > \frac{n_a(n_a-1)}{m(m-1)}$ populated clubs. In particular, this means that the m-Complete Environment achieves the maximal sum of utilities of the set of connected m-Uniform environments when $c \leq (m-1)[h(m) - h^2(m)]$.

This result implies that an environment that maximizes the sum of utilities from the set of non-empty m-Uniform environments when $c \leq (m-1)[h(m) - h^2(m)]$ is a collection of m-Complete components and isolated agents. Note that the sum of utilities of an m-Complete component with n agents ($n > 1$) is $n(n-1)[h(m) - \frac{c}{m-1}]$. Since $[h(m) - \frac{c}{m-1}]$ is non-negative, the sum of utilities of an m-Complete component is a weakly increasing and weakly convex function of the number of agents in the component. Therefore, if $c \leq (m-1)[h(m) - h^2(m)]$ an environment that achieves the maximum of the sum of utilities from the set of non-empty m-Uniform environments is the m-Complete Environment. Also, when $c \leq (m-1)[h(m) - h^2(m)]$, the sum of utilities of the m-Complete Environment is non-negative. Thus, when $c \in [0, (m-1)(h(m) - h^2(m)]$ the m-Complete Environment achieves the maximum of the sum of utilities from the set of all m-Uniform environments.

An m-Star Environment includes $\frac{n_a-1}{m-1}$ clubs. Therefore, the sum of utilities of an m-Star Environment is

$$\frac{n_a - 1}{m-1} m(m-1)h(m) + [n_a(n_a - 1) - \frac{n_a - 1}{m-1} m(m-1)]h^2(m) - \frac{n_a - 1}{m-1} mc$$

The difference between the maximal sum of utilities of a connected m-Uniform Environment with k clubs is

$$[k - \frac{n_a - 1}{m-1}]m(m-1)h(m) - [k - \frac{n_a - 1}{m-1}]m(m-1)h^2(m) - [k - \frac{n_a - 1}{m-1}]mc$$

Or,

$$[k - \frac{n_a - 1}{m-1}]m[(m-1)h(m) - (m-1)h^2(m) - c]$$

Again, there must be a k such that the maximal total sum of utilities of a connected m-Uniform Environment is greater than the total sum of utilities of an m-Star Environment. Therefore, this difference must be non-negative for some k. If $c > (m-1)[h(m) - h^2(m)]$, $[k - \frac{n_a-1}{m-1}]m$ must be non-positive and since $k \geq \frac{n_a-1}{m-1}$ it must be that $k = \frac{n_a-1}{m-1}$ and the difference is zero. Also, if $c = (m-1)[h(m) - h^2(m)]$ the difference is zero. Therefore, the m-Star Environment achieves the maximal sum of utilities of the set of connected m-Uniform environments when $c \geq (m-1)[h(m) - h^2(m)]$.

This result implies that when $c \geq (m-1)[h(m) - h^2(m)]$ an environment that maximizes the sum of utilities from the set of non-empty m-Uniform environments is a collection of m-Star sub-environments and isolated agents. In fact, an environment that maximizes the sum of utilities from the set of m-Uniform environments is a collection of m-Star sub-environments

with non-negative sum of utilities and isolated agents.
Suppose that C_1 and C_2 are two m-Star environments with $n_1 > 1$ and $n_2 > 1$ agents, respectively. Let b_1 and b_2 be the central agents of C_1 and C_2, respectively. Consider a new environment C that includes the clubs of C_1 and the clubs of C_2 where b_2 is replaced by b_1. Thus, C is an m-Star environment with $n_1 + n_2 - 1$ agents with an additional isolated agent. The utility of the central agent in C is the sum of utilities of b_1 and b_2 in C_1 and C_2, respectively. The utility of all other agents improves due to the additional free indirect connections. Thus, uniting two m-stars into one bigger m-Star environment (and an isolate) always increase the sum of utilities. That is, when $c \geq (m-1)[h(m) - h^2(m)]$, an environment that achieves the maximal sum of utilities from the set of non-empty m-Uniform environments is an m-Star environment and some isolated agents. Thus, assuming that $n_a - 1$ is a multiple of $m-1$, if $c \geq (m-1)[h(m) - h^2(m)]$, an environment that achieves the maximal the sum of utilities from the set of non-empty m-Uniform environments is the m-Star Environment.
To complete the proof notice that the m-Star Environment achieves the maximal sum of utilities from the set of all m-Uniform environments if and only if it has a non-negative sum of utilities. If it has non-positive sum of utilities, the Empty Environment achieves the maximal the sum of utilities from the set of m-Uniform environments. The condition for the sum of utilities of the m-Star Environment with n_a agents to be non-negative is

$$(m-1)h(m) + \frac{(n_a - m)(m-1)}{m} h^2(m) \geq c$$

Thus, when $c \in [(m-1)(h(m) - h^2(m)), (m-1)h(m) + \frac{(n_a-m)(m-1)}{m} h^2(m)]$ the m-Star Environment achieves the maximal sum of utilities from the set of m-Uniform environments. When $c \geq (m-1)h(m) + \frac{(n_a-m)(m-1)}{m} h^2(m)$ the Empty Environment achieves the maximal the sum of utilities from the set of m-Uniform environments. □

A.8 Proposition 4

Proof. Let G be an m-Complete Environment. Assume, first, that $n_a > m \geq 2$ (there is more than one club in the environment). The utility of Agent i from Environment G is (denote $\gamma \equiv \frac{n_a-1}{m-1} \in \mathbb{N}$):[45] $u_i(G) = (n_a - 1)h(m) - \gamma c$.
To calculate the utility of Agent i from aborting any one of her affiliations, suppose that Agent i leaves Club s which she shares with Agent i'. In addition, suppose she shares the Club s' with Agent i''. By the definition of an m-Complete environment, $S_{G-\{i,s\}}(i) \cap S_{G-\{i,s\}}(i') = \emptyset$. Thus, the new shortest path between Agent i and Agent i' must be indirect. Again, by the definition of m-Complete environments the populated clubs in $G - \{i, s\}$ are of size m except Club s which is of size $m-1$. Therefore, the new shortest path is of length 2 and its weight must be either $h(m-1)h(m)$ or $h^2(m)$ (any path of length of more than 2 has a lower or equal weight than $h^2(m)$). Now, let us show that in $G - \{i, s\}$ there is no shortest path between Agent i and Agent i' of the weight $h(m-1)h(m)$. Suppose such a path exists. Then, there is an Agent j who shares a club with Agent i (denoted by t) and also shares Club s with Agent i'.

[45] We assume that $\frac{n_a-1}{m-1}$ and $\frac{n_a(n_a-1)}{m(m-1)}$ are integers. For further details see Footnote 19.

Thus, in G, Agent j shared s also with Agent i which implies, however, $S_G(i) \cap S_G(j) = \{s, t\}$ and the m-completeness of G is violated. However, a shortest path of weight $h^2(m)$ between Agent i and Agent i' in $G - \{i, s\}$ does exist. Recall that Agent i shares Club s' with Agent i'' and note that Agent i'' is not a member of Club s (otherwise agent i and i'' share two clubs in G) and that Agent i' is not a member of Club s' (otherwise agent i and i' share two clubs in G). Hence, by the definition of an m-complete environment, $\exists s'' \in S \backslash \{s, s'\} : \{i', i''\} \subseteq N_G(s'')$. Thus, Agent i has a link of weight $h(m)$ with Agent i'' (Club s') and Agent i'' has a link of weight $h(m)$ with Agent i' (Club s''). Therefore, there is a path of weight $h^2(m)$ between Agent i and Agent i' in Environment $G - \{i, s\}$. Thus, the utility of Agent i from Environment $G - \{i, s\}$ is $u_i(G - \{i, s\}) = (n_a - m)h(m) + (m - 1)h^2(m) - (\gamma - 1)c$ and $u_i(G - \{i, s\}) - u_i(G) = (m-1)h^2(m) - (m-1)h(m) + c$. Agent i would not wish to leave any of her clubs if and only if $u_i(G - \{i, s\}) \leq u_i(G)$, meaning that she would not wish to leave any of her clubs if and only if $(m - 1)[h(m) - h^2(m)] \geq c$. Thus, $(m - 1)[h(m) - h^2(m)] \geq c$ guarantees that the "No Leaving" condition holds.

Next, let us calculate the utility of Agent i from joining an existing Club s. Since G is m-complete, $\forall i' \in N_G(s) : |S_{G+\{i,s\}}(i) \cap S_{G+\{i,s\}}(i')| = 2$. Moreover, since $\forall i' \in N_G(s)$: $w(i, i', G) = h(m)$, $n_{G+\{i,s\}}(s) = m + 1$ and $h(m) \geq h(m + 1)$, Agent i does not improve any of her shortest paths by joining Club s. However, she pays c as participation fees. Therefore, $c \geq 0$ guarantees that the "No Joining" condition holds.

Next, let us calculate the utility of Agent i from the formation of a new club by the group K ($i \in K, K \subseteq N$) and let $|K| = k$. Note that $c \geq 0$ guarantees the "No New Club Formation" condition for the case of $k \geq m$ due to similar considerations to those used in the case of the "No Joining" condition above. For $m > k \geq 2$, the utility of Agent i from the Environment $G + K$ is $u_i(G + K) = (n_a - k)h(m) + (k - 1)h(k) - (\gamma + 1)c$. Since $u_i(G + K) - u_i(G) = (k - 1)h(k) - (k - 1)h(m) - c$ we get that if $c \geq (k - 1)[h(k) - h(m)]$ then $u_i(G + K) \leq u_i(G)$. Thus, Agent i will refuse to establish a new club as part of Group K if and only if $c \geq (k - 1)[h(k) - h(m)]$. However, in order to ensure that Agent i will refuse to establish a new club with any subset of agents it must be that this condition will hold $\forall k \in \{2, \ldots, m - 1\}$. Therefore, $c \geq \max\limits_{k\in\{2,\ldots,m-1\}} (k - 1)[h(k) - h(m)]$ guarantees that the "No New Club Formation" condition holds for $k < m$. Since $\forall k \in \{2, \ldots, m\}$: $(k - 1)[h(k) - h(m)] \geq 0$, this condition also ensures that the No Joining condition and the No New Club Formation condition for the case of $k \geq m$ hold.

Denote $k^\star = \min\{\arg\max_{k\in\{2,\ldots,m-1\}} (k - 1)[h(k) - h(m)]\}$. Note that the condition above can be rewritten as $\max_{k\in\{2,\ldots,m-1\}} k_h(k) - (k - 1)h(m)$. Let $k' \in \{\hat{k} + 1, \ldots, n_a\}$. By the definition of $\hat{k}$, we get $k_h(\hat{k}) \geq k_h(k')$. In addition, since given an m-Complete environment $h(m)$ is fixed, we get $(\hat{k} - 1)h(m) \leq (k' - 1)h(m)$. Hence, for every $k' \in \{\hat{k} + 1, \ldots, n_a\}$, we have $k_h(\hat{k}) - (\hat{k} - 1)h(m) \geq k_h(k') - (k' - 1)h(m)$. Therefore, $k^\star \leq \hat{k}$. This implies that the "No New Club Formation" and the "No Joining" conditions hold if $c \geq \max\limits_{k\in\{2,\ldots,\min\{m-1,\hat{k}\}\}} (k - 1)[h(k) - h(m)]$.

To complete the proof consider the case where $n_a = m$. In this case, G includes one club that consists of all the agents in the environment. Note that the considerations stated above for the lower bound hold also when $n_a = m$. Thus, if $c \geq \max\limits_{k\in\{2,\ldots,\min\{n_a-1,\hat{k}\}\}} (k - 1)[h(k) - h(n_a)]$ the "No Joining" and the "No New Club Formation" conditions hold. However, the "No

Leaving" condition is different since if an agent decides to leave the club her utility is zero. The utility of the agents from G is $u_i(G) = (n_a - 1)h(n_a) - c$. Therefore, an agent will not leave the club as long as $(n_a - 1)h(n_a) \geq c$. Thus, $(n_a - 1)h(n_a) \geq c$ guarantees that the "No Leaving" condition holds when $n_a = m$. □

A.9 Claim 1

A.9.1 Lemma 7

Lemma 7. *Let $h(\cdot)$ be an exponential club congestion function where $\delta \in (0, \frac{1}{2})$. Then,* $\max_{k \in \{2,\dots,m-1\}} (k-1)[h(k) - h(m)] = h(2) - h(m)$.

Proof.

$$\forall l \in \{2, \dots, m-2\}, \forall k \in \{0, \dots, m-l-1\} : \frac{h(l+k) - h(l+k+1)}{h(l) - h(l+1)} = \delta^k$$

Thus,

$$\forall l \in \{2, \dots, m-2\} : \sum_{k=0}^{m-l-1} \frac{h(l+k) - h(l+k+1)}{h(l) - h(l+1)} = \sum_{k=0}^{m-l-1} \delta^k$$

Or,

$$\forall l \in \{2, \dots, m-2\} : \frac{h(l) - h(m)}{h(l) - h(l+1)} = \sum_{k=0}^{m-l-1} \delta^k$$

Since this is a geometric series and since $\delta \in (0, \frac{1}{2})$, we get

$$\forall l \in \{2, \dots, m-2\} : \frac{h(l) - h(m)}{h(l) - h(l+1)} = \frac{1 - \delta^{m-l}}{1 - \delta} < \frac{1}{1 - \delta} < 2$$

Therefore,

$$\forall l \in \{2, \dots, m-2\} : h(l) - h(m) < 2[h(l) - h(l+1)]$$

And,

$$\forall l \in \{2, \dots, m-2\} : h(l) - h(m) < l[h(l) - h(l+1)]$$

Or,

$$\forall l \in \{2, \dots, m-2\} : l[h(l+1) - h(m)] < (l-1)[h(l) - h(m)]$$

Therefore, $\max_{k \in \{2,\dots,m-1\}} (k-1)[h(k) - h(m)] = h(2) - h(m)$. □

A.9.2 The Proof

Proof. Let $h(\cdot)$ be an exponential club congestion function where $\delta \in (0, \frac{1}{2})$ and $a > 0$. By Proposition 4 and Lemma 7, for every $n_a > m$, there exists a range of membership fees where the m-complete environment is OCS if $(m-1)[(a+\delta^{m-1}) - (a+\delta^{m-1})^2] > \delta - \delta^{m-1}$. Note that the right-hand-side of the inequality is bounded from above by $\delta < \frac{1}{2}$. In addition, note that the left-hand-side of the inequality can be written as

$$(m-1)(a-a^2) + (1-2a)(m-1)\delta^{m-1} - (m-1)\delta^{2(m-1)}$$

Then, $(1-2a)(m-1)\delta^{m-1}$ and $(m-1)\delta^{2(m-1)}$ go to zero when m goes to infinity, while, since $a \in (0,1)$, $(m-1)(a-a^2)$ goes to infinity when m goes to infinity. Thus, the left-hand-side of the inequality is not bounded. Moreover, since the left-hand-side of the inequality is monotonic from some club size (depends on δ and a) there exists $\bar{m}$ such that $\forall m : m > \bar{m}$ the inequality holds. Thus, $\forall m : n_a > m > \bar{m}$ there exists a range of membership fees in which the m-complete environment is OCS.
For similar reasons there exists an integer $\tilde{m}$ such that $\forall m : m > \tilde{m}$ the upper bound is higher than δ (since δ is greater than the right-hand-side for every $m \geq 2$, $\tilde{m} \geq \bar{m}$). Let $\bar{c} = (\tilde{m}-1)[(a+\delta^{\tilde{m}-1}) - (a+\delta^{\tilde{m}-1})^2]$. Thus, in the membership costs range $(\delta, \bar{c})$, every m-complete environment where $m \geq \tilde{m}$ is OCS. $\square$

A.10 Claim 2

Proof. By Proposition 4 the Grand Club environment is OCS if and only if

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

where $\hat{k}$ denotes the club size that maximizes the DCV.
Therefore, there exists a range of membership fees in which the Grand Club environment is OCS if and only if

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

Or, alternatively, such a range exists if and only if

$$\forall k \in \{2,\ldots,\min\{n_a-1,\hat{k}\}\} : (k-1)[h(k) - h(n_a)] \leq (n_a-1)h(n_a)$$

Using the DCV notation, such a range exists if and only if

$$\forall k \in \{2,\ldots,\min\{n_a-1,\hat{k}\}\} : k_h(k) - (k-1)h(n_a) \leq k_h(n_a)$$

Equivalently, there exists a range of membership fees in which the Grand Club environment is OCS if and only if

$$\forall k \in \{2,\ldots,\min\{n_a-1,\hat{k}\}\} : k_h(k) - \frac{k-1}{n_a-1}k_h(n_a) \leq k_h(n_a)$$

Thus, there exists a range of membership fees in which the Grand Club environment is OCS if and only if

$$\forall k \in \{2, \ldots, \min\{n_a - 1, \hat{k}\}\} : \frac{n_a - 1}{n_a + k - 2} \times k_h(k) \leq k_h(n_a)$$

Since for every $k \in \{2, \ldots, \min\{n_a - 1, \hat{k}\}\}$ we have $\frac{n_a-1}{n_a+k-2} < 1$, if the DCV is increasing then the inequality is satisfied. By Lemma 4, if the club congestion function is inelastic the DCV is strictly increasing, and therefore if the club congestion function is inelastic, a range of membership fees in which the Grand Club environment is OCS exists. □

A.11 Claim 3

Proof. By Lemma 7, the lower bound of the range of membership costs in which the Grand Club Environment is OCS becomes $h(2) - h(n_a) = \delta - \delta^{n_a-1}$.
When $a = 0$, the upper bound is $(n_a - 1)h(n_a) = (n_a - 1)\delta^{n_a-1}$. Since $n_a \geq 4$ we can write $n_a \leq 2^{n_a-2}$ or $\frac{1}{n_a} \geq \frac{1}{2^{n_a-2}}$. Using $\delta \in (0, \frac{1}{2})$ we have $\frac{1}{n_a} > \delta^{n_a-2}$ or $1 > n_a\delta^{n_a-2}$ or $\delta > n_a\delta^{n_a-1}$. Hence, $\delta - \delta^{n_a-1} > (n_a - 1)\delta^{n_a-1}$, that is the lower bound is always greater than the upper bound. Thus, for $a = 0$, $\delta \in (0, \frac{1}{2})$ and $n_a \geq 4$ the Grand Club Environment is never OCS.
A range of membership fees for which the Grand Club Environment is OCS exists if and only if $\delta - \delta^{n_a-1} \leq (n_a - 1)(a + \delta^{n_a-1})$. That is, there exists a range of membership fees for which the Grand Club Environment is OCS if and only if $(n_a - 1)a + n_a\delta^{n_a-1} \geq \delta$. Let $\bar{n}_a = \frac{\delta}{a} + 1$. Then, $(\bar{n}_a - 1)a = \delta$, and therefore, for every $n_a > \bar{n}_a$ there exists a range of membership fees for which the Grand Club Environment is OCS. □

A.12 Proposition 5

Proof. Let $G =< N, S, A >$ be an m-Star environment where $n_a > m \geq 2$ and denote the number of populated clubs in G by $\gamma \equiv \frac{n_a-1}{m-1}$ (we assume that γ is an integer). For simplicity, we refer to the central agent as Agent b and to the other agents as agents i, i', etc.
We begin with an upper bound on the range of membership fees where G is OCS. The upper bound is set by the membership fees above which agents would wish to wave any of their affiliations. The utility of the central agent from Environment G is $u_b(G, h, c) = (n_a - 1)h(m) - \gamma c$. Consider Club s. Since all non-central club members have no other affiliations, no path exists between Agent b and these agents once Agent b leaves Club s. Therefore, for every $\{b, s\} \in A$, $u_b(G - \{b, s\}, h, c) = (n_a - m)h(m) - (\gamma - 1)c$. Therefore, Agent b will have no incentive to leave any of her affiliations if and only if $(m - 1)h(m) \geq c$. The utility of a non-central agent i from Environment G is $u_i(G, h, c) = (m - 1)h(m) + (n_a - m)h^2(m) - c$. Consider Club s such that $\{i, s\} \in A$. The utility of Agent i after aborting her affiliation with Club s is zero. Therefore, Agent i will not have an incentive leave the club if and only if $(m - 1)h(m) + (n_a - m)h^2(m) \geq c$. Thus, no agent has an incentive to leave a club in G if and only if $(m - 1)h(m) \geq c$ or $k_h(m) \geq c$.
We continue with the lower bound on the range of membership fees where G is OCS. The lower bound is set by the membership fees below which agents would wish to form new

affiliations either by joining a new club or by forming a new club.

We begin by considering the benefits for a subset of agents ($K \subseteq N$, $k = |K|$) from forming a new club (r). Denote by $K_s = N_{G+K}(s) \cap N_{G+K}(r)$ the set of agents that share Club s in G and are affiliated with the new Club r and denote its magnitude by k_s. Denote the set of clubs represented in K by $Q = \{s \in S : k_s > 0\}$ and its magnitude by q.

We first consider the case where the new club is no larger than the existing clubs, $k \leq m$. Each agent in K gets a benefit from the direct connections with the other members in K. These connections replace either a link with a weight of $h(m)$ (if they share a club in G) or a path with a weight of $h^2(m)$ (if they do not share a club in G). Obviously, an improvement on an indirect connection is larger than an improvement on a direct connection ($h(k) - h^2(m) \geq h(k) - h(m)$). Therefore, non-central agents get the same benefit as the central agent on links with agents they already share a club with in G and higher benefit than the central agent on links with agents they do not share a club with in G. Thus, in the case where $b \in K$, since the utility from environment $G + K$ to Agent b is $u_b(G+K, h, c) = (k-1)h(k) + (n_a - k)h(m) - (\gamma+1)c$ we get that $c \geq \max_{m \geq k \geq 2}(k-1)(h(k) - h(m))$ prevents the formation of K since it does not benefit Agent b. For the case where $b \notin K$ we begin by considering the case where the size of the new club is not greater than the original club size and the original number of clubs ($m \geq k$ and $\gamma \geq k$). Note that on top of the improved direct connections that each agent in K gets, the partners with whom she did not share a club with in G supply her with improved indirect paths to the agents in their original clubs that do not participate in K. These paths are better than the paths supplied in G by the central agent, since the new club is small ($k \leq m$). The utility from Environment $G + K$ for Agent i such that $\{i, s\} \in A$, $b \notin K$ and $i \in K$ is

$$\begin{aligned} u_i(G+K, h, c) =& (k-1)h(k) + (m - k_s)h(m) + ((q-1)(m-1) - (k - k_s))h(k)h(m) \\ &+ (\gamma - q)(m-1)h^2(m) - 2c \end{aligned}$$

For every Agent i and every m, $h(\cdot)$, c, q and k, $u_i(G+K, h, c)$ is maximized if $k_s = 1$.[46] Thus, the utility of Agent i from K is maximized if no other member in this club shares her original club in G. The utility of Agent i ($(i, s) \in A$) from $G + K$ when K includes no other agent from Club s is

$$\begin{aligned} u_i(G+K, h, c) =& (k-1)h(k) + (m-1)h(m) + ((q-1)(m-1) - (k-1))h(k)h(m) \\ &+ (\gamma - q)(m-1)h^2(m) - 2c \end{aligned}$$

In addition, to maximize $u_i(G + K, h, c)$ given m, $h(\cdot)$, c and k, q should be as high as possible.[47] Thus, the utility of Agent i from a new club where $q = k$ (no pair of agents in

[46] $\frac{\partial u_i(G+K)}{\partial k_s} = -h(m) + h(k)h(m) = -h(m)(1 - h(k)) \leq 0$. Since q and k are held fixed, increasing k_s by 1 means that Agent j' of Club s' ($k_{s'} > 1$ since q is fixed) is replaced in K by a member j of Club s ($j \notin \{i, b\}$). The gain from this change is the improved path to Agent j ($h(k) - h(m)$) while the loss is the longer path to j' ($h(k)h(m) - h(k)$). Thus, the net benefit is $-h(m) + h(k)h(m)$.

[47] $\frac{\partial u_i(G+K)}{\partial q} = (m-1)h(k)h(m) - (m-1)h^2(m) \geq 0$ (equality is achieved if and only if $k = m$). Since k is held fixed, increasing q by 1 means that Agent j' of Club s' ($k_{s'} > 1$) is replaced in K by a member j of Club s that was not represented in K. The gain from this change is the paths to Agent j and her club members ($h(k) + (m-2)h(k)h(m) - (m-1)h^2(m)$) while the loss is the longer path to j' ($h(k)h(m) - h(k)$). Thus, the net benefit is $(m-1)(h(k)h(m) - h^2(m))$.

the new club share a club in G) is

$$u_i(G+K,h,c) = (k-1)h(k) + (m-1)h(m) + ((k-1)(m-1)-(k-1))h(k)h(m) + (\gamma-k)(m-1)h^2(m) - 2c$$

Therefore, the non-central agents have no incentive to form a new club of size $\min\{m,\gamma\} \geq k$ if

$$c \geq \max_{\min(\gamma,m)\geq k\geq 2} (k-1)[h(k)+(m-2)h(k)h(m)-(m-1)h^2(m)]$$

To complete the case of $k \leq m$, we consider the case of a new club of size $m \geq k > \gamma$ that does not include the central agent. Suppose $q < \gamma$. Then, there is a non-empty Club s in G such that $k_s = 0$. In this case, in $G+K$, $\forall i \in K$ there are some indirect paths with weight $h(k)h(m)$ and some indirect paths with weight $h^2(m)$. Alternatively, suppose $q = \gamma$. For every no non-empty club s in G, $k_s > 0$. Then, $\forall i \in K$ the direct links are the same as in the previous case, but all the indirect paths are of weight $h(k)h(m)$. Clearly, for each agent, the incentives to form a new club are weakly stronger when $q = \gamma$.
The utility from Environment $G+K$ to Agent i who participates in Club s and in Group K where $q = \gamma$ ($b \notin K$) is

$$u_i(G+K) = (k-1)h(k) + (m-k_s)h(m) + (n_a - m - (k-k_s))h(k)h(m) - 2c$$

Given m, $h(\cdot)$ and c, $u_i(G+K)$ increases when k_s decreases (see Footnote 46). Thus, the most attractive K is the one that minimizes the maximal k_s (over all $s \in S$) where $q = \gamma$. In this new optimal club $\max_{s,s'\in S} |k_s - k_{s'}| \leq 1$ and the agents that belong to the original clubs with the higher k_s have lower utility. Denote the optimal k_s by $\eta_k \equiv \lceil \frac{k}{\gamma} \rceil$. Then, the utility of $i \in K$ that belongs to the original Club $s \in \{s \in S | \forall s' \in S, k_s \geq k_{s'}\}$ is

$$u_i(G+K) = (k-1)h(k) + (m-\eta_k)h(m) + (n_a - m - (k-\eta_k))h(k)h(m) - 2c$$

Thus, the membership fees required to prevent the formation of a new club of size k where $\gamma < k \leq m$ are

$$c > \max_{m\geq k>\gamma} (k-1)h(k) - (\eta_k - 1)h(m) + (n_a - m - (k-\eta_k))h(k)h(m) - (n_a - m)h^2(m)$$

It is easy to see that the membership fees required to prevent the formation of a new club of size $k \leq m$ are higher when the central agent is not included in the group. Denote the fees required to prevent a deviation to a club of size k when $m \geq k$ and $\gamma \geq k$,

$$FNS_h(k,m) = (k-1)[h(k) + (m-2)h(k)h(m) - (m-1)h^2(m)]$$

and the fees required to prevent a deviation to a club of size k when $m \geq k$ and $k > \gamma$,

$$FNI_h(k,m,n_a) = (k-1)h(k) - (\eta_k - 1)h(m) + (n_a - m - (k-\eta_k))h(m)h(k) - (n_a - m)h^2(m)$$

Therefore, we can conclude that the minimal membership fees required to prevent the formation of a new club that is no larger than the existing clubs, $k \leq m$, depends on the relation

between m and γ.
If $m > \gamma$

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

while if $\gamma \geq m$

$$c \geq \max_{m \geq k \geq 2} FNS_h(k,m)$$

Note that when $k > m$ there are no gains to the members of the new club from shorter indirect paths. In addition, they have no gains from the members of the new club with whom they already share a club in G (therefore the central agent can never benefit from participating in clubs of size $k > m$). Moreover, if $h^2(m) \geq h(k)$ there are no gains also from the other members of the new club. Thus, no new club of size $k \geq l_h$ will be formed where $l_h = \min\{k \in \mathbb{Z} | h(k) \leq h^2(m)\}$. However, the net gains for a non-central Agent i, that belongs to Club s in Environment G, from establishing a new club of size $\min\{l_h, n_a\} > k > m$ are $(k-k_s)(h(k) - h^2(m)) - c$. Since there is at least one agent in K for which $k_s \geq \eta_k$, she would refuse to deviate if $c > (k - \eta_k)(h(k) - h^2(m))$. Denote the fees required to prevent a deviation to a club of size k when $\min\{l_h, n_a\} > k > m$ by $FNL_h(k,m,n_a) = (k-\eta_k)(h(k) - h^2(m))$. Thus, the minimal membership fees required to prevent the formation of a new club are
If $m > \gamma$

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

while if $\gamma \geq m$

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

Finally, we analyse the incentive to join a new club. Agent b is irrelevant since she is already present in all the populated clubs. A non-central Agent i who joins an existing Club s shortens her paths to the members of this club (excluding the center) while she pays the membership fees (and intensifies the congestion in Club s). The utility of Agent i from environment $G + \{i,s\}$ (where $\{i,s\} \notin A$ and $n_G(s) \geq 2$) is

$$u_i(G + \{i,s\}, h, c) = (m-1)h(m) + (m-1)h(m+1) + (n_a - 2m + 1)h^2(m) - 2c$$

Therefore, the net benefit for Agent i from joining an existing club s is $(m-1)[h(m+1) - h^2(m)] - c$. Thus, no agent wishes to join a new club in G if and only if $c \geq (m-1)[h(m+1) - h^2(m)]$. Denote $J_h(m) = (m-1)[h(m+1) - h^2(m)]$.
Note that $FNS_h(m,m) = (m-1)(h(m) - h^2(m))$ and therefore $FNS_h(m,m) \geq J_h(m)$. Thus, the lower bound on the membership are

If $m > \gamma$

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

while if $\gamma \geq m$

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

□

A.13 Claim 4

Proof. First note that since $n_a > 2$ and $m = 2$, the number of clubs is never smaller than m and therefore only the first part of Proposition 5 is relevant for the 2-Star Environment.
Let us begin with Part 1. When $m = 2$ we get $k_h(2) = h(2)$ as the upper bound. For the lower bound only $FNS_h(2,2)$ and $FNL_h(k,2,n_a)$ for $k \in \{2,\ldots,\min\{l_h - 1, n_a - 1\}\}$ are relevant. $FNS_h(2,2) = h(2) - h^2(2)$. Since $m = 2$ and $k \leq n_a - 1$, we get $\eta_k = 1$. Therefore, $FNL_h(k,2,n_a) = (k-1)(h(k) - h^2(2))$. The 2-Star Environment is therefore OCS if and only if

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

Next, by Lemma 4, since the club congestion function is elastic then $k_h(\cdot)$ is strictly decreasing. Note that $(k-1)(h(k) - h^2(2)) = k_h(k) - (k-1)h^2(2)$. Thus, the first part decreases with k while the second part increases with k, meaning that $(k-1)(h(k) - h^2(2))$ is maximized by $k = 2$. Therefore, since $h(\cdot)$ is an elastic club congestion function, the 2-Star Environment is OCS if and only if $h(2) \geq c \geq h(2) - h^2(2)$.
The reciprocal club congestion function implies that $(k-1)(h(k) - h^2(2))$ equals $1 - (k-1)h^2(2) = 1 - (k-1) = 2 - k$. Thus, using Part 1, if $h(\cdot)$ is the reciprocal club congestion function, the 2-star environment is OCS if and only if $c \in [0,1]$.
Finally, suppose that $h(\cdot)$ is an exponential club congestion function. A straight forward application of Part 1 suggests that the 2-Star Environment is OCS if and only if

$$a + \delta \geq c \geq \max_{k \in \{2,\ldots,\min\{l_h - 1, n_a - 1\}\}} (k-1)((a + \delta^{k-1}) - (a + \delta)^2)$$

It is easy to see that when $a = 0$ the upper bound becomes δ. Since $a = 0$ implies that $l_h = 3$, the lower bound becomes $\delta - \delta^2$. Therefore, if $a = 0$ then the 2-Star Environment is OCS if and only if $c \in [\delta - \delta^2, \delta]$. □

A.14 Corollary 1

Proof. $h(3) < h^2(2)$ implies that $h(2) - h^2(2) < h(2) - h(3)$. Thus, as our previous analysis indicates, the All-paired Environment is the unique OCS if and only if $c \in (0, h(2) - h^2(2))$

while when $c > (h(2) - h^2(2))$ the All-Paired Environment is not OCS. By Proposition 4 the lower bound of the range of membership fees for which the 3-Complete Environment is OCS is $h(2) - h(3)$. Therefore, the 3-Complete Environment is not OCS for $c \in (h(2) - h^2(2), h(2) - h(3))$. $h(3) < h^2(2)$ implies that $l_h = 3$. Therefore, by Claim 4 the 2-Star Environment is OCS if and only if $c \in [h(2) - h^2(2), h(2)]$. As a result, the 2-Star Environment is OCS when $c \in (h(2) - h^2(2), h(2) - h(3))$.
$h(3) = h^2(2)$ implies that $h(2) - h^2(2) = h(2) - h(3)$. First, let us show that $h(2) > 2[h(3) - h^2(3)] > h(2) - h^2(2)$. We begin with the left inequality. Since the maximal value of $f(x) = x - x^2$ is $\frac{1}{4}$, $2[h(3) - h^2(3)] \leq \frac{1}{2}$ and this inequality is correct for $h(2) > \frac{1}{2}$. In addition, since $f(x) = x - x^2$ is strictly increasing in $x \in [0, \frac{1}{2}]$, $h(2) - h^2(2) > h(3) - h^2(3)$ when $h(2) \leq \frac{1}{2}$. Thus, $h(2) - h(3) = h(2) - h^2(2) > h(3) - h^2(3) > h(3) - 2h^2(3)$. Therefore, also for $h(2) \leq \frac{1}{2}$, $h(2) > 2[h(3) - h^2(3)]$. Next we show that $2[h(3) - h^2(3)] > h(2) - h^2(2)$. Since $f(x) = x - x^2$ is strictly decreasing in $x \in [\frac{1}{2}, 1]$, when $1 > h(2) > h(3) \geq \frac{1}{2}$ we get $2[h(3) - h^2(3)] > h(3) - h^2(3) > h(2) - h^2(2)$. In addition, when $\frac{1}{2} \geq h(3) \geq 0.15$ we get $2[h(3) - h^2(3)] > \frac{1}{4} \geq h(2) - h^2(2)$ where the first inequality is since $h(3) \geq 0.15$ and $f(x) = x - x^2$ is increasing below $x = \frac{1}{2}$. The second inequality is correct since the maximal value of $f(x) = x - x^2$ is $\frac{1}{4}$. As our previous analysis indicates the All-Paired Environment is the unique OCS if and only if $c \in (0, h(2) - h^2(2))$ and it is not OCS when $c > h(2) - h^2(2)$. By Proposition 4 the upper bound of the range of membership fees for which the 3-Complete Environment is OCS is $2[h(3) - h^2(3)]$ and the lower bound is $h(2) - h(3) = h(2) - h^2(2)$. Since $l_h = 3$, by Claim 4 the 2-Star Environment is OCS if and only if $c \in [h(2) - h^2(2), h(2)]$.
$h(3) > h^2(2)$ implies that $h(2) - h^2(2) > h(2) - h(3)$. Also, recall from the previous part that $2[h(3) - h^2(3)] > h(2) - h^2(2)$ when $1 \geq h(3) \geq 0.15$. As our previous analysis indicates the All-Paired Environment is the unique OCS if and only if $c \in (0, h(2) - h(3))$, it is OCS (but not unique) when $c \in [h(2) - h(3), h(2) - h^2(2)]$ and it is not OCS when $c > h(2) - h^2(2)$. By Proposition 4 the 3-Complete Environment is OCS when $c \in [h(2) - h(3), h(2) - h^2(2)]$ and when $c \in [h(2) - h^2(2), 2(h(3) - h^2(3))]$. By Claim 4, the lower bound of the range of membership fees for which the 2-Star Environment is OCS is $\max_{k \in \{2,\ldots,\min\{l_h-1, n_a-1\}\}} (k-1)(h(k) - h^2(2))$. Since for $k = 2$, $(k-1)(h(k) - h^2(2)) = h(2) - h^2(2)$ it is guaranteed that for $c < h(2) - h^2(2)$ the 2-Star Environment is not OCS. $\square$

A.15 Claim 5

Proof. Since $n_a \geq 9$ then $\gamma \geq 4 > 3$, so that only the first part of Proposition 5 is relevant. First, let $h(m) = \frac{1}{m-1}$. Note that $l_h = 5$. Thus, the 3-Star Environment is OCS if and only if

$$2h(3) \geq c \geq \max\{h(2) + h(2)h(3) - 2h^2(3), 2[h(3) - h^2(3)], 3[h(4) - h^2(3)]\}$$

and using the functional form we get $1 \geq c \geq \max\{1, \frac{1}{2}, \frac{1}{4}\}$ and therefore the 3-Star Environment is OCS if and only if $c = 1$.
Next, let $h(m) = \delta^{m-1}$ for $\delta \in (0, 1)$. Note that again $l_h = 5$. Thus, again the 3-star environment is OCS if and only if

$$2h(3) \geq c \geq \max\{h(2) + h(2)h(3) - 2h^2(3), 2[h(3) - h^2(3)], 3[h(4) - h^2(3)]\}$$

and using the functional form the 3-Star Environment is OCS if and only if

$$2\delta^2 \geq c \geq \max\{\delta + \delta^3 - 2\delta^4, 2\delta^2 - 2\delta^4, 3\delta^3 - 3\delta^4\}$$

Note that since $\delta \in (0, 1)$ it must be that $\delta(1 - \delta)^2 > 0$. Therefore, $\delta + \delta^3 > 2\delta^2$ and $\delta + \delta^3 - 2\delta^4 > 2\delta^2 - 2\delta^4$. Meaning that the 3-Star Environment is OCS if and only if

$$2\delta^2 \geq c \geq \max\{\delta + \delta^3 - 2\delta^4, 3\delta^3 - 3\delta^4\}$$

Note that since $\delta \in (0, 1)$ it must be that $\delta^2(2 - \delta) < 1$.[48] Therefore, $2\delta^2 - \delta^3 < 1$ or $2\delta^3 - \delta^4 < \delta$ or $3\delta^3 - \delta^4 < \delta + \delta^3$ or $3\delta^3 - 3\delta^4 < \delta + \delta^3 - 2\delta^4$. Meaning that the 3-Star Environment is OCS if and only if

$$2\delta^2 \geq c \geq \delta + \delta^3 - 2\delta^4$$

Note that given that $\delta \in (0, 1)$ then $(\delta^2 + 1)(2\delta - 1) \geq 0$ if and only if $\delta \in [\frac{1}{2}, 1)$. Thus, $2\delta^3 - \delta^2 + 2\delta \geq 1$ if and only if $\delta \in [\frac{1}{2}, 1)$. And, $2\delta^4 - \delta^3 + 2\delta^2 \geq \delta$ if and only if $\delta \in [\frac{1}{2}, 1)$. Meaning that $2\delta^2 \geq \delta + \delta^3 - 2\delta^4$ if and only if $\delta \in [\frac{1}{2}, 1)$. Thus, since the 3-Star Environment is OCS if and only if $c \in [\delta + \delta^3 - 2\delta^4, 2\delta^2]$, there is a range of membership fees for which it is OCS if and only if $\delta \geq \frac{1}{2}$. □

A.16 Proposition 6

Proof. Let G be the Grand Club Environment with n_a agents. The utility of each agent in G is $u_i(G, c, b) = (n_a - 1)b^2(1) - c > 0$. Note that no agent can achieve more than $(n_a - 1)b^2(1) - c$ since the gains are maximal and the fees are minimal. Thus, the Grand Club Environment is PE. To show that it is the unique PE environment it suffices to show that every other environment includes at least one agent that obtains strictly lower utility. Obviously, the Empty Environment is not PE. If an environment includes a single populated club, and it is not the Grand Club Environment, then there is at least one agent with no affiliations, and therefore zero utility. Hence there is no PE environment, other than the Grand Club Environment, which includes at most one populated club. Next, consider a disconnected environment that includes at least two populated clubs. Then, no agent is connected to all other agents and therefore the maximal possible utility is $(n_a - 2)b^2(1) - c$ which is strictly smaller than $(n_a - 1)b^2(1) - c$. Hence, no disconnected environment is PE. Finally, consider a connected environment that includes at least two clubs. Then, there exists at least one agent that is a member of more than one club, and therefore her maximal utility is $(n_a - 1)b(2)b(1) - 2c$ which is strictly smaller than $(n_a - 1)b^2(1) - c$ since $\max\{b(1) - b(2), c\} > 0$. Hence, no environment other than the Grand Club Environment is PE. Thus, the Grand Club Environment is the unique PE and therefore also the unique SE, and the proof of Part 1a is completed.
To prove Part 1b note that no agent i wants to leave the club since $u_i(G, c, b) = (n_a - 1)b^2(1) - c > 0$ by keeping her membership and zero by leaving the club. The "No Joining" condition is vacuously satisfied since there are no other populated clubs. Finally, no subset

[48] $\delta^2(2 - \delta)$ has a local maximum at $\frac{4}{3}$, a local minimum at 0 and its value at $\delta = 1$ is 1.

of agents want to form a new club, since by being members of two clubs, their utility drops to $u_i(G+m,c,b) = (k-1)b^2(2) + (n_a-k)b(1)b(2) - 2c$ where k is the number of agents in m. Hence, the Grand Club Environment is OCS.

For any membership fees, no agent with at least one affiliation can achieve more than $(n_a-1)b^2(1) - c$ since the gains are maximal and the fees are minimal. Thus, when $c > (n_a-1)b^2(1)$ every agent with at least one affiliation must have negative utility. Thus, the Empty Environment is the unique PE and therefore also the unique SE, and the proof of Part 2a is completed.

To prove Part 2b note that forming a new club can provide each agent with at most $u_i(G,c,b) = (n_a-1)b^2(1) - c$. Since $c > (n_a-1)b^2(1)$ no subset of agents wishes to form a new club. The "No Joining" and "No Leaving" conditions are vacuously satisfied since there are no populated clubs. Finally, since the maximal utility of an agent in this model is $u_i(G,c,b) = (n_a-1)b^2(1) - c$, then when $c > (n_a-1)b^2(1)$ in every non-empty environment there exists an agent that wishes to leave any of her affiliations. Therefore, every non-empty environment is not OCS. Hence, the Empty Environment is the unique OCS environment when $c > (n_a-1)b^2(1)$.

Part 3a is a direct conclusion from the uniqueness of the Empty Environment when $c > (n_a-1)b^2(1)$ (Part 2b) and from the stability of the Grand Club Environment when $c \leq (n_a-1)b^2(1)$ as demonstrated in Part 1b (when $c = (n_a-1)b^2(1)$ the Grand Environment is OCS, PE and SE, but the unique efficiency is lost).

To prove Part 3b, for every G and for every $c \in [0, n_a-1)$, we find an individual congestion function $b(\cdot)$ such that G is not OCS while the Grand Club Environment with n_a agents is OCS. Suppose G is the Empty Environment and let $b(1) > \sqrt{\frac{c}{n_a-1}}$ (since $c \in [0, n_a-1)$ such $b(1)$ always exists). Then, a deviation to form a new club that consists all agents is worthwhile and G is not OCS. Since $(n_a-1)b^2(1) > c$, by Part 1b the Grand Club Environment is OCS.

Let G be an environment where every agent maintains exactly one affiliation and there is more than one populated club (i.e. Partitioned Environment which is not the Grand Club Environment). Let s be one of the populated clubs in G, denote $|s| = l$ ($n_a-1 > l > 1$) and suppose that Agent i is a member of s. For the case where $c > 0$ note that for every individual congestion function where $\sqrt{\frac{c}{n_a-1}} < b(1) < \min\{1, \sqrt{\frac{c}{l-1}}\}$ (such $b(1)$ exists since $n_a-1 > l$ and $c \in [0, n_a-1)$) we get $u_i(G,c,b) = (l-1)b^2(1) - c < 0$ and therefore Agent i would find leaving Club s worthwhile and G is not OCS. However, since $c < (n_a-1)b^2(1)$ by Part 1b the Grand Club Environment is OCS. For the case where $c = 0$ note that $u_i(G,0,b) = (l-1)b^2(1)$. If Agent i joins another existing populated club s' of size $|s'| = k > 1$ her utility is $u_i(G + \{i,s'\}, 0, b) = (k+l-1)b(2)b(1)$. Note that for every individual congestion function such that $b(1) > b(2) > b(1)(1 - \frac{1}{n_a-1})$ it is worthwhile for Agent i to join club s' since it means that $b(2) > b(1)(1 - \frac{k}{k+l-1})$ (recall that $k > 1$ and $k + l \leq n_a$) which guarantees that $(k+l-1)b(2)b(1) > (l-1)b^2(1)$. Thus, we found an individual congestion function such that no Partitioned Environment with more than one populated club is OCS if $c = 0$. However, by Part 1b the Grand Club environment is OCS.

Now let G be an environment where every agent maintains at most one affiliation and there is at least one isolated agent. Let s be one of the populated clubs in G, denote $|s| = l$ ($n_a > l > 1$) and suppose that Agent i is a member of s. For the case where $c > 0$ note that for every individual congestion function where $\sqrt{\frac{c}{n_a-1}} < b(1) < \min\{1, \sqrt{\frac{c}{l-1}}\}$ (such $b(1)$ exists since $n_a > l$ and $c \in [0, n_a - 1))$ we get $u_i(G, c, b) = (l-1)b^2(1) - c < 0$ and therefore Agent i would find leaving Club s worthwhile and G is not OCS. However, since $c < (n_a - 1)b^2(1)$ by Part 1b the Grand Club Environment is OCS. For the case of $c = 0$, such an environment is never OCS since the isolated agent wishes to join any of the clubs (recall that $b(1) > 0$). Hence, G is not OCS. However, by Part 1b the Grand Club Environment is OCS.

Next, suppose that G includes at least one agent with more than one affiliation. With no loss of generality assume that Agent i is the agent with the highest number of affiliations and she maintains $k > 1$ memberships. Her utility is $u_i(G) = \sum_{j \in N, j \neq i} d(i, j|G) - kc$. Notice that the weight of every path between Agent i and another agent must be a multiplication of $b(k)$. Thus, let us denote $d(i, j|G) = b(k)d_{-i}(i, j|G)$ and therefore $u_i(G) = b(k) \sum_{j \in N, j \neq i} d_{-i}(i, j|G) - kc$. For Agent i, leaving Club $s \in S_G(i)$ is beneficial if $u_i(G) < u_i(G - \{i, s\})$ or

$$b(k) \sum_{j \in N, j \neq i} d_{-i}(i, j|G) - kc < b(k-1) \sum_{j \in N, j \neq i} d_{-i}(i, j|G - \{i, s\}) - (k-1)c$$

A sufficient condition for this inequality to hold is

$$b(k) \sum_{j \in N, j \neq i} d_{-i}(i, j|G) < b(k-1) \sum_{j \in N, j \neq i} d_{-i}(i, j|G - \{i, s\})$$

Or,

$$b(k) < b(k-1) \frac{\sum_{j \in N, j \neq i} d_{-i}(i, j|G - \{i, s\})}{\sum_{j \in N, j \neq i} d_{-i}(i, j|G)}$$

Denote,

$$F = \frac{\sum_{j \in N, j \neq i} d_{-i}(i, j|G - \{i, s\})}{\sum_{j \in N, j \neq i} d_{-i}(i, j|G)} \leq 1$$

Since $k > 1$, by leaving Club s Agent i is not isolated and $\sum_{j \in N, j \neq i} d_{-i}(i, j|G - \{i, s\}) > 0$. Hence, for every individual congestion function such that $b(k)$ is smaller than $b(k-1) \times F$, G is not OCS for any membership fees. Obviously, such an individual congestion function exists. In addition, let us set $b(1) = 1$. Then by Part 1b the Grand Club Environment with the same number of agents is OCS. □

A.17 Proposition 7

A.17.1 Lemma 8

Lemma 8. *Let* $G =< N, S, A >$*. In the individual congestion model if* $S_G(j) \cap S_G(i) \neq \emptyset$ *then* $d(i, j|G, b) = b(s_G(i)) \times b(s_G(j))$*.*

Proof. Note that the weight of every indirect path between Agent i and Agent j would be of the form $b(s_G(i)) \times \cdots \times b(s_G(j))$. If Agent i and Agent j are directly connected (since they share the same club), the value of the link between them in the induced network is $b(s_G(i)) \times b(s_G(j))$. Since $b(\cdot)$ is bounded from above by 1, then it must be that the direct link is the shortest path between Agent i and Agent j and therefore $d(i, j|G, b) = b(s_G(i)) \times b(s_G(j))$. □

A.17.2 Lemma 9

Lemma 9. *Let* $G =< N, S, A >$*, let* $m \subseteq N$ *and let* $i \in m$*. In the Individual Congestion model, if* $\forall j \in m$*:* $S_G(j) \cap S_G(i) \neq \emptyset$ *then for every individual congestion function* $b(\cdot)$*,* $u_i(G, b, c) \geq u_i(G + m, b, c)$ *and the inequality is strict for* $c > 0$*.*

Proof. In the individual congestion model, forming a new club is costly and weakly decreases the weights on the existing links of the members of the new club. Hence, a new club may improve the utility of the deviating agents only by forming shorter paths. Now, consider the formation of a new club by the set m and specifically Agent $i \in m$ such that $\forall j \in m$: $S_G(j) \cap S_G(i) \neq \emptyset$.
First, let us consider the connections between Agent i and Agent $k \in m$. By Lemma 8 the shortest path between Agent i and Agent k in both G and $G + m$ is their direct link and since both have an additional affiliation in $G + m$ we get $d(i, k|G, b) \geq d(i, k|G + m, b)$.
Now, let us consider the connections between Agent i and Agent $k \notin m$. Denote by $P(i, k|G)$ the shortest path between Agent i and Agent k in Environment G. Suppose that $P(i, k|G) \neq P(i, k|G + m)$. If the path $P(i, k|G + m)$ was available in G then its weight must have been not greater than $d(i, k|G, b)$ (otherwise $P(i, k|G + m)$ would have been the shortest path in G). Since the weights on the links did not increase in $G + m$, it must be that $d(i, k|G, b) \geq d(i, k|G + m, b)$. Next, we show that a path $P(i, k|G+m)$ that was not available in G can never be a shortest path in $G + m$. Suppose, in negation, that $P(i, k|G + m)$ was not available in G but is a shortest path in $G + m$. Then, it must be that at least one of its links is new, meaning it is a link between two agents in $m\backslash\{i\}$ - Agent j_1 and Agent j_2. However, since Agent i is directly linked with both Agent j_1 and Agent j_2 in $G + m$, by Lemma 8 a shorter path exists. contradiction. Thus, also for Agent $k \notin m$ we get $d(i, k|G, b) \geq d(i, k|G + m, b)$.
Thus, the formation of a new club by the set m that includes an Agent $i \in m$ such that $\forall j \in m$: $S_G(j) \cap S_G(i) \neq \emptyset$ does not improve any of her links. Hence,

$$\sum_{j \in N, j \neq i} d(i, j|G, b) \geq \sum_{j \in N, j \neq i} d(i, j|G + m, b)$$

And if $c > 0$,

$$\sum_{j \in N, j \neq i} d(i,j|G,b) - s_G(i) \times c > \sum_{j \in N, j \neq i} d(i,j|G+m,b) - (s_G(i)+1) \times c$$

Thus, we conclude that $u_i(G,b,c) \geq u_i(G+m,b,c)$ and the inequality is strict for $c > 0$. $\square$

A.17.3 Lemma 10

Lemma 10. *Let $G =< N, S, A >$, let $s \in S$ and let $i \notin s$. In the Individual Congestion model, if $\forall j \in s$: $S_G(j) \cap S_G(i) \neq \emptyset$ then for every individual congestion function $b(\cdot)$, $u_i(G,b,c) \geq u_i(G+\{i,s\},b,c)$ and the inequality is strict for $c > 0$.*

Proof. In the individual congestion model, joining an existing club is costly and weakly decreases the weights of the agent. Hence, joining a club may improve the utility of the agent only by forming shorter paths. Since $\forall j \in s$: $S_G(j) \cap S_G(i) \neq \emptyset$ the network induced by $G+\{i,s\}$ have the same links as the network induced by G. Moreover, the weights on the links in the network induced by $G+\{i,s\}$ are smaller or equal to the corresponding weights in G. Thus, for every agent k we get $d(i,k|G,b) \geq d(i,k|G+\{i,s\},b)$. Hence,

$$\sum_{j \in N, j \neq i} d(i,j|G,b) \geq \sum_{j \in N, j \neq i} d(i,j|G+\{i,s\},b)$$

And if $c > 0$,

$$\sum_{j \in N, j \neq i} d(i,j|G,b) - s_G(i) \times c > \sum_{j \in N, j \neq i} d(i,j|G+m,b) - (s_G(i)+1) \times c$$

Thus, $u_i(G,b,c) \geq u_i(G+\{i,s\},b,c)$ and the inequality is strict for $c > 0$. $\square$

A.17.4 The Proof

Proof. Since every agent in an m-Complete environment shares a club with any other agent, by Lemma 10, for every individual congestion function $b(\cdot)$ and for every membership fees, $u_i(G,b,c) \geq u_i(G+\{i,s\},b,c)$. Hence, no agent wants to join an existing club. For the same reason, by Lemma 9, for every individual congestion function $b(\cdot)$ and for every membership fees, no subset of agents wishes to form a new club.
Thus, m-Complete environments are not OCS if and only if there is an agent that wishes to leave any of her clubs. The utility of Agent i from an m-Complete environment given the individual congestion function $b(\cdot)$ and membership fees $c \geq 0$ is

$$u_i(G,b,c) = (n_a - 1)b^2(\gamma) - \gamma \times c$$

Recall that $n_a > m$, meaning that by leaving a club $m-1$ direct connections are replaced with indirect connections but all the remaining links of the agent are of higher quality and the total membership fees are lower. Specifically,

$$u_i(G-\{i,s\},b,c) = (n_a - m)b(\gamma)b(\gamma-1) + (m-1)b^3(\gamma)b(\gamma-1) - (\gamma-1) \times c$$

Thus, m-Complete environments are OCS if and only if

$$(n_a - 1)b^2(\gamma) - \gamma \times c \geq (n_a - m)b(\gamma)b(\gamma - 1) + (m - 1)b^3(\gamma)b(\gamma - 1) - (\gamma - 1) \times c$$

Or,

$$(n_a - m)b(\gamma)[b(\gamma) - b(\gamma - 1)] + (m - 1)b^2(\gamma)[1 - b(\gamma - 1)b(\gamma)] \geq c$$

Since we are only interested in positive membership fees, an m-complete environment is OCS if and only if

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

□

A.18 Proposition 8

A.18.1 Lemma 11

Lemma 11. *In the model with club congestion and individual congestion where $h(2) = 1$, $\forall m > 2 : h(m) = 0$, $b(k) = \frac{1}{2}[1 + \frac{1}{k}]$, $D = 1$ and $c = \frac{1}{4}$, $\forall g \in \mathbb{G}_n, \forall i \in N : u_i(G_g) = \frac{1}{4} \times u_i^{CA}(g)$.*

Proof. For every un-weighted network $g =< N, \bar{E} >$, the induced network of G_g denoted by $g =< N, E, W >$ is such that $E = \bar{E}$ and, by the choice of $h(\cdot)$ and $b(\cdot)$, each link between Agent i and Agent j has a weight of $\frac{1}{4}[1 + \frac{1}{n_i}][1 + \frac{1}{n_j}]$ or $\frac{1}{4} \times [\frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j}] + \frac{1}{4}$. Since $D = 1$ the agents benefit from direct connections only. Moreover, since the membership costs are $\frac{1}{4}$, the net utility of Agent i from the link to her club partner (and network neighbor) Agent j, is $\frac{1}{4} \times [\frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j}]$. Obviously, the net utility of Agent i from non-neighbors is zero. As a consequence, summing over all agents, $u_i(G_g) = \frac{1}{4} \times u_i^{CA}(g)$. □

A.18.2 The Proof

Proof. We suppose that $g \in CA(n)$ and show that $G_g \in OCS(\frac{1}{4}, n, h, b, 1)$. The "No Joining" condition holds since the utility from a club of size 3 is zero while the participation fees are positive. For the same reason, no coalition of size greater than two wishes to form a new club.
Next, consider two agents, i and j, that do not share a club in G_g. Then, by construction, Agent i and Agent j are not linked in g. Since g is pairwise stable, if $u_i^{CA}(g) < u_i^{CA}(g + \{i, j\})$ then $u_j^{CA}(g) > u_j^{CA}(g + \{i, j\})$. Alternatively, if $\frac{1}{4} \times u_i^{CA}(g) < \frac{1}{4} \times u_i^{CA}(g + \{i, j\})$ then $\frac{1}{4} \times u_j^{CA}(g) > \frac{1}{4} \times u_j^{CA}(g + \{i, j\})$. By Lemma 11, if $u_i(G_g) < u_i(G_{g+\{i,j\}})$ then $u_j(G_g) > u_j(G_{g+\{i,j\}})$. Denote by m_{ij} the coalition that includes only agents i and j. Then, note that $G_{g+\{i,j\}}$ is identical to $G_g + m_{ij}$ since both denote the addition of Club s that includes agents i and j to Environment G_g. Hence, if $u_i(G_g) < u_i(G_g + m_{ij})$ then $u_j(G_g) > u_j(G_g + m_{ij})$. Therefore, no coalition of size two wishes to form a new club and the "No New Club Formation" condition holds.
For the "No Leaving" condition, consider Agent i that participates, together with Agent j,

in Club s in G_g. Then, by construction, Agent i and Agent j are linked in g. Since g is pairwise stable $u_i^{CA}(g) \geq u_i^{CA}(g - \{i,j\})$. Alternatively, $\frac{1}{4} \times u_i^{CA}(g) \geq \frac{1}{4} \times u_i^{CA}(g - \{i,j\})$. By Lemma 11, $u_i(G_g) \geq u_i(G_{g-\{i,j\}})$. By Lemma 6, $u_i(G_g) \geq u_i(G_g - \{i,s\})$, meaning that this condition also holds. Therefore, $G_g \in OCS(\frac{1}{4}, n, h, b, 1)$.
For the other direction, we suppose that $G_g \in OCS(\frac{1}{4}, n, h, b, 1)$ and show that $g \in CA(n)$. First, consider Agent i that is linked with Agent j in g. By construction Agent i participates, together with Agent j, in Club s in G_g. Since G_g is OCS, $u_i(G_g) \geq u_i(G_g - \{i,s\})$. By Lemma 6, $u_i(G_g) \geq u_i(G_{g-\{i,j\}})$. By Lemma 11, $\frac{1}{4} \times u_i^{CA}(g) \geq \frac{1}{4} \times u_i^{CA}(g - \{i,j\})$. Thus, $u_i^{CA}(g) \geq u_i^{CA}(g - \{i,j\})$, meaning that no agent wishes to discard an existing link. Next, consider two agents, i and j, that are not linked in g. By construction agents i and j do not share a club in G_g. Since G_g is OCS, if $u_i(G_g) < u_i(G_g + m_{ij})$ then $u_j(G_g) > u_j(G_g + m_{ij})$. But, as mentioned above, $G_{g+\{i,j\}}$ is identical to $G_g + m_{ij}$. Therefore, if $u_i(G_g) < u_i(G_{g+\{i,j\}})$ then $u_j(G_g) > u_j(G_{g+\{i,j\}})$. By Lemma 11, if $\frac{1}{4} \times u_i^{CA}(g) < \frac{1}{4} \times u_i^{CA}(g + \{i,j\})$ then $\frac{1}{4} \times u_j^{CA}(g) > \frac{1}{4} \times u_j^{CA}(g + \{i,j\})$. Thus, if $u_i^{CA}(g) < u_i^{CA}(g + \{i,j\})$ then $u_j^{CA}(g) > u_j^{CA}(g + \{i,j\})$, meaning that no pair of agents wishes to form a new link. Therefore, $g \in CA(n)$ and the proof of the first part is completed.
For the second part note that since we assume that $\mathcal{G}_n$ includes only environments with distinct clubs, every environment $G \in \mathcal{G}_n \backslash \mathcal{G}_{\mathbb{G}n}$ includes at least one populated club of size greater than two. However, every agent that participates in a club of size greater than two wishes to leave the club since its benefits are zero (all induced links of such club are of weight zero) while the membership fees are positive. Therefore, $G \notin OCS(\frac{1}{4}, n, h, b, 1)$. $\square$

B DCV and Elasticity

B.1 The DCV of the Exponential Congestion Function

Lemma B.1 summarizes the club size that maximizes the DCV for various sets of parameters of the exponential congestion function. Some technical notations are required: Denote $b(\delta, n_a) = \frac{1}{n_a - 2}(\delta - (n_a - 1)\delta^{n_a - 1})$ and let $\delta^\star(n_a)$ be the unique $\delta \in (0,1)$ such that $b(\delta, n_a) = \delta(1 - 2\delta)$ and let $\hat{\delta}(n_a)$ be the unique $\delta \in (0,1)$ such that $b(\delta, n_a) = 0$.

Lemma B.1. *Let $n_a \geq 4$ and let $h(m)$ be an exponential club congestion function.*

1. *The club size that maximizes the DCV weakly increases with a.*
2. *If $a \in [0, \min\{b(\delta, n_a), \delta(1 - 2\delta)\})$ then the DCV is maximized at $m = 2$.*
3. *If $\delta \in (0, \delta^\star(n_a))$ and $a \in (b(\delta, n_a), 1 - \delta)$ then the DCV is maximized at $m = n_a$.*
4. *If $\delta \in (\delta^\star(n_a), \hat{\delta}(n_a))$ and $a \in (\max\{0, \delta(1 - 2\delta)\}, b(\delta, n_a))$ then the DCV is maximized at $m \in \{3, \ldots, n_a - 1\}$.*
5. *If $\delta \in [\frac{1}{2}, 1 - \frac{1}{n_a - 1}]$ and $a = 0$ then the DCV is maximized either at $m = \lfloor 1 - \frac{1}{\ln \delta} \rfloor$ or at $m = \lceil 1 - \frac{1}{\ln \delta} \rceil$.*
6. *If $\delta \in (1 - \frac{1}{n_a - 1}, 1)$ then the DCV is maximized at $m = n_a$.*

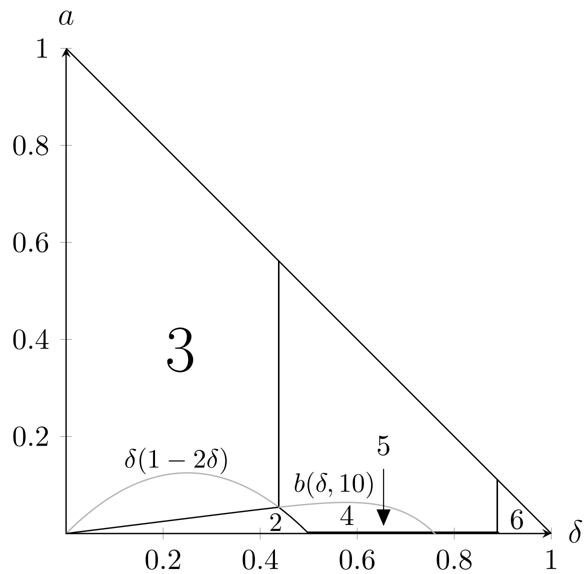


Figure 8: Lemma B.1 for $n_a = 10$. Each number lies within the area characterized by the corresponding statement in the lemma.

The case where $a = 0$ demonstrates the opposing effects of the club size on the DCV. When the number of agents in the club increases, congestion increases (for every δ) but more direct links are formed in the club (the multiplicative effect). We use Lemma 4 to show that when $\delta < \frac{1}{2}$ the congestion effect is dominant and the DCV is maximized when the club is small (Part 2). When δ increases the effect of congestion weakens and increasingly larger clubs maximize the DCV (Parts 5 and 6).

When the non-congested component of the exponential congestion function is introduced it reinforces the multiplicative effect since the aggregate benefit from $a > 0$ increases with the size of the club. Therefore, the club size that maximizes the DCV weakly increases with a (Part 1). Part 2 shows that for relatively low values of δ and a, the congestion component is still dominant and the DCV is maximized by the smallest club. But, when the non-congestion component increases (and δ is still low) then the DCV is maximized by the biggest club (Part 3).[49] Part 3 also makes use of the assertion in Part 1 to state that if the DCV is maximized by the biggest club for some a then it is maximized by the biggest club for any greater a (Part 6 uses the same assertion). If a is high enough (for $\delta > \frac{1}{2}$ its any value of a), a club of size two never maximizes the DCV since the multiplicative effect dominates the congestion component. Parts 4 and 5 show that for these values of δ, intermediate size clubs can maximize the DCV. Figure 8 demonstrates Lemma B.1 for the case of $n_a = 10$.

B.1.1 The Proof of Lemma B.1

Lemma B.2. *Let $h(m)$ be an exponential congestion function. Let $m > m'$ and suppose $k_h(m) > k_h(m')$ for a given parameter a. Then $k_h(m) > k_h(m')$ for every $\bar{a} \in [a, 1-\delta)$.*

[49] In the proof we use Lemma B.3 that shows that $k_h(m)$ has three parts - increasing, decreasing and increasing again. Therefore, to determine the club size that globally maximizes the DCV, the closest integer to the local maxima that separates the first two parts should be compared to the right-hand side limit, n_a.

Proof. For the given parameter a, $k_h(m) - k_h(m') > 0$. Therefore, $(m-1)(a+\delta^{m-1}) - (m'-1)(a+\delta^{m'-1}) > 0$ or written differently, $(m-m')a + (m-1)\delta^{m-1} - (m'-1)\delta^{m'-1} > 0$. Now suppose a increases to $\bar{a}$. Since $m > m'$, $(m-m')\bar{a} + (m-1)\delta^{m-1} - (m'-1)\delta^{m'-1} > 0$ and therefore $(m-1)(\bar{a}+\delta^{m-1}) - (m'-1)(\bar{a}+\delta^{m'-1}) > 0$. Hence, $k_h(m) - k_h(m') > 0$ given $\bar{a}$. □

Lemma B.3. *Let $n_a \geq 4$ and let $h(m)$ be an exponential club congestion function with $a > 0$. $k_h(m)$ has at most two extreme points, $\bar{m} < \bar{\bar{m}}$, where $\bar{m}$ is a local maximum and $\bar{\bar{m}}$ is a local minimum.*

Proof. Denote $g_h(m) = \frac{\eta_h(m)}{\eta_h(m+1)}$. To show that $g_h(m)$ is strictly increasing, it is helpful to rewrite it as

$$g_h(m) = \frac{\frac{\frac{h(m+1)-h(m)}{h(m)}}{\frac{1}{m}}}{\frac{\frac{h(m+2)-h(m+1)}{h(m+1)}}{\frac{1}{m+1}}}$$

Then,

$$g_h(m) = \frac{m}{m+1} \times \frac{h(m+1)}{h(m)} \times \frac{h(m+1)-h(m)}{h(m+2)-h(m+1)} = \frac{m}{m+1} \times \frac{h(m+1)}{h(m)} \times \frac{1}{\delta}.$$
$$g_h(m+1) = \frac{m+1}{m+2} \times \frac{h(m+2)}{h(m+1)} \times \frac{h(m+2)-h(m+1)}{h(m+3)-h(m+2)} = \frac{m+1}{m+2} \times \frac{h(m+2)}{h(m+1)} \times \frac{1}{\delta}.$$

Note that for every integer $m \geq 1$, $\delta \in (0,1)$ satisfies $\delta^{m-1} + \delta^{m+1} > 2\delta^m$. Therefore, $a^2 + 2a\delta^m + \delta^{2m} < a^2 + a\delta^{m-1} + a\delta^{m+1} + \delta^{2m}$ which can be rewritten as $h^2(m+1) < h(m) \times h(m+2)$. Hence, $\forall m \in \{2, \dots, n_a - 2\} : \frac{h(m+2)}{h(m+1)} > \frac{h(m+1)}{h(m)}$. Also, note that $\forall m \in \mathbb{N} : \frac{m+1}{m+2} > \frac{m}{m+1}$. Taken together, $\forall m \in \{2, \dots, n_a - 2\} : g_h(m+1) > g_h(m)$, meaning $g_h(m)$ is strictly increasing.
Since $\eta_h(m) \leq 0$ and since $g_h(m)$ is strictly increasing, there exists $m^\star$ such that for every $m < m^\star$ the club-size elasticity $\eta_h(m)$ is decreasing ($g_h(m) < 1$) while for every $m > m^\star$, $\eta_h(m)$ is increasing ($g_h(m) > 1$). Thus, generally, $\eta_h(m)$ is unimodal with a single minimum at $m^\star$.
Thus, generally, $\eta_h(m)$ can be divided to four parts in the following order:

(i) $\eta_h(m) > -1$ and $\eta_h(m)$ is decreasing.

(ii) $\eta_h(m) < -1$ and $\eta_h(m)$ is decreasing.

(iii) $\eta_h(m) < -1$ and $\eta_h(m)$ is increasing.

(iv) $\eta_h(m) > -1$ and $\eta_h(m)$ is increasing.

Therefore, by Lemma 4, $k_h(m)$ has at most three parts, the first increasing (corresponding to (i)), the second decreasing (corresponding to (ii) and (iii)) and the third increasing again (corresponding to (iv)). Hence, for $n_a \geq 4$, $k_h(m)$ has at most two extreme points, $\bar{m} < \bar{\bar{m}}$, where $\bar{m}$ is a local maximum and $\bar{\bar{m}}$ is a local minimum. □

Lemma B.4. *For every $n_a \geq 4$:*

1. $\hat{\delta}(n_a) > \frac{1}{2}$.

2. $\forall \delta \in (0, \hat{\delta}(n_a)) : b(\delta, n_a) > 0$.

3. $\arg\max_{\delta \in (0,1)} b(\delta, n_a) = \hat{\delta}^2(n_a)$.

Proof. First, note that $\hat{\delta}(n_a) = (\frac{1}{n_a-1})^{\frac{1}{n_a-2}}$ is the unique root of $b(\delta, n_a)$ that is real, positive and smaller than one. Next,

$$\frac{\partial \hat{\delta}(n_a)}{\partial n_a} = \frac{1}{n_a-2} \times (\frac{1}{n_a-1})^{\frac{1}{n_a-2}-1} \times \frac{-1}{(n_a-1)^2} + (\frac{1}{n_a-1})^{\frac{1}{n_a-2}} \times \ln \frac{1}{n_a-1} \times \frac{-1}{(n_a-2)^2}$$

Hence,

$$\frac{\partial \hat{\delta}(n_a)}{\partial n_a} = -\frac{1}{n_a-2} \times (\frac{1}{n_a-1})^{\frac{1}{n_a-2}} \times \big[\frac{1}{(n_a-1)} + \ln \frac{1}{n_a-1} \times \frac{1}{(n_a-2)}\big]$$

Since $n_a \geq 4$, $\frac{\partial \hat{\delta}(n_a)}{\partial n_a} > 0$ if and only if $\frac{1}{(n_a-1)} + \ln \frac{1}{n_a-1} \times \frac{1}{(n_a-2)} < 0$. Hence, $\frac{\partial \hat{\delta}(n_a)}{\partial n_a} > 0$ if and only if $\ln \frac{1}{n_a-1} < -1 + \frac{1}{n_a-1}$. Therefore, if $\ln \frac{1}{n_a-1} < -1$ then $\frac{\partial \hat{\delta}(n_a)}{\partial n_a} > 0$. This means that if $n_a > e+1$ then $\frac{\partial \hat{\delta}(n_a)}{\partial n_a} > 0$. Since $n_a \geq 4$ we showed that $\frac{\partial \hat{\delta}(n_a)}{\partial n_a} > 0$.
Note that $\hat{\delta}(4) = (\frac{1}{3})^{\frac{1}{2}} \approx 0.577$. Hence $\hat{\delta}(4) > \frac{1}{2}$. Since $\frac{\partial \hat{\delta}(n_a)}{\partial n_a} > 0$ we get that $\hat{\delta}(n_a) > \frac{1}{2}$ for every $n_a \geq 4$.
For every $n_a \geq 4$, $b(0, n_a) = 0$ and $b(\hat{\delta}(n_a), n_a) = 0$ and there is no other $\delta \in [0, \hat{\delta}(n_a)]$ such that $b(\delta, n_a) = 0$. Since $b(\delta, n_a)$ is continuous and its derivative with respect to δ at $\delta = 0$ is positive when $n_a \geq 4$ ($\frac{\partial b(\delta, n_a)}{\partial \delta}(0, n_a) = \frac{1}{n_a-2} > 0$), we infer that $\forall \delta \in (0, \hat{\delta}(n_a)) : b(\delta, n_a) > 0$ when $n_a \geq 4$.
Finally,

$$\frac{\partial b(\delta, n_a)}{\partial \delta} = \frac{1}{n_a-2} - \frac{(n_a-1)^2}{n_a-2} \times \delta^{n_a-2}$$

Thus, for a given n_a, the maximum of $b(\delta, n_a)$ is achieved at $\delta = (\frac{1}{n_a-1})^{\frac{2}{n_a-2}} = \hat{\delta}^2(n_a)$. $\square$

Proof. First, by Lemma B.2, if $m > m'$ and $k_h(m) > k_h(m')$ for a given parameter a then $k_h(m) > k_h(m')$ for every $\bar{a} \in [a, 1-\delta)$. Hence, if $m^\star$ is the club size that maximizes the DCV for a, then for every $\bar{a} \in [a, 1-\delta)$ the DCV is maximized by $m \geq m^\star$. Hence, the club size that maximizes the DCV weakly increases with a (Part 1).
Second, we show that if $a \in [0, \min\{b(\delta, n_a), \delta(1-2\delta)\})$ then the DCV is maximized at $m = 2$. We begin by considering the case of $h(m) = a + \delta^{m-1}$ where $\delta \in (0,1)$, $a \in (0, \min\{b(\delta, n_a), \delta(1-2\delta)\})$ and $a + \delta \in (0,1)$. By Lemma B.3, $k_h(m)$ has at most two extreme points, $\bar{m} < \bar{\bar{m}}$, where $\bar{m}$ is a local maximum and $\bar{\bar{m}}$ is a local minimum. If $k_h(2) > k_h(3)$, rewritten as $a < \delta(1-2\delta)$, then $m = 2$ must be the local integer maximum of $k_h(m)$. Therefore, in these cases the global integer maximum is either at $m = 2$ or at $m = n_a$. Hence, if also $k_h(2) > k_h(n_a)$, rewritten as $a < \frac{1}{n_a-2}\delta(1-(n_a-1)\delta^{n_a-2}) = b(n_a, \delta)$, then the global integer maximum is at $m = 2$. Thus, if $a \in (0, \min\big\{\delta(1-2\delta), b(n_a, \delta)\big\})$ then

$k_h(m)$ is maximized at $m = 2$. Note that $\delta(1-2\delta) > 0$ if and only if $\delta \in (0, \frac{1}{2})$. Hence, it is left to be shown that if $a = 0$ and $\delta \in (0, \frac{1}{2})$ then $k_h(m)$ is maximized at $m = 2$. In this case $h(m) = \delta^{m-1}$ and therefore the club-size elasticity is $\eta_h(m) = m(\delta - 1)$. Thus, the congestion function is elastic if $\delta < \frac{1}{2}$ since then $\eta_h(m) < -1$ for every club size. By Lemma 4, $k_h(m)$ is decreasing and therefore maximized at $m = 2$. Hence, if $a \in [0, \min\left\{\delta(1-2\delta), b(n_a, \delta)\right\})$ then the DCV is maximized at $m = 2$ (Part 2).

Recall that $\delta(1-2\delta)$ is positive if and only if $\delta \in (0, \frac{1}{2})$ and that its derivative with respect to δ at $\delta = 0$ is one ($\frac{\partial\delta(1-2\delta)}{\partial\delta}(\delta = 0) = 1$). Also recall that when $n_a \geq 4$ by Lemma B.4, $b(n_a, \delta)$ is positive when $\delta \in (0, \hat{\delta}(n_a))$ where $\hat{\delta}(n_a) > \frac{1}{2}$ and its derivative with respect to δ at $\delta = 0$ is $\frac{1}{n_a - 2} < 1$. Hence, these two function cross for some $\delta \in (0, \frac{1}{2})$ and since both are single peaked at this region, we denote it by $\delta^\star$ ($b(\delta^\star, n_a) = \delta^\star(1-2\delta^\star)$). Therefore, there is a unique $\delta^\star(n_a) \in (0, \frac{1}{2})$ such that $\forall\delta \in (0, \delta^\star) : \delta(1-2\delta) > b(n_a, \delta)$ and $\forall\delta \in (\delta^\star, \frac{1}{2}) : \delta(1-2\delta) < b(n_a, \delta)$. Consider the case where $\delta \in (0, \delta^\star)$ and $a \in (b(\delta, n_a), \delta(1-2\delta))$. In this range, $m = 2$ must be the local integer maximum of $k_h(m)$ (since $a < \delta(1-2\delta)$). However, the global maximum is $m = n_a$ since $a > b(\delta, n_a)$. Thus, for $\delta \in (0, \delta^\star)$ and $a \in (b(\delta, n_a), \delta(1-2\delta))$ the DCV is maximized at $m = n_a$. However, by Lemma B.2, by increasing a the club size that maximizes the DCV cannot decrease. Since n_a is the maximal size, then for $\delta \in (0, \delta^\star)$ and $a \in (b(\delta, n_a), 1-\delta)$ the DCV is maximized at $m = n_a$ (Part 3).

Next, consider the case where $\delta \in (\delta^\star, \hat{\delta})$ and $a \in (\max\{0, \delta(1-2\delta)\}, b(\delta, n_a))$. In this range, $m = 2$ is not the local integer maximum of $k_h(m)$ (since $a > \delta(1-2\delta)$). But, $k_h(2) > k_h(n_a)$ since $a < b(\delta, n_a)$. Therefore, the DCV is not maximized by $m = 2$ and it is not maximized by $m = n_a$. Therefore, the DCV is maximized at $m \in \{3, \ldots, n_a - 1\}$ (Part 4).

Next, consider the case where $\delta \in [\frac{1}{2}, 1 - \frac{1}{n_a - 1}]$ and $a = 0$. In this case the congestion function reduces to $h(m) = \delta^{m-1}$ where $\delta \in [\frac{1}{2}, 1 - \frac{1}{n_a - 1}]$. As a continuous function $k_h(m) = (m-1)\delta^{m-1}$ is single peaked and achieves its maximum at $m^\star = 1 - \frac{1}{\ln\delta}$. Therefore, the highest values achieved by integers are either in $\left\lfloor 1 - \frac{1}{\ln\delta} \right\rfloor$ or $\left\lceil 1 - \frac{1}{\ln\delta} \right\rceil$. Hence, when $a = 0$ and the club-size elasticity is indeterminate ($1 - \frac{1}{n_a - 1} \geq \delta \geq \frac{1}{2}$) the DCV is maximized either at $m = \left\lfloor 1 - \frac{1}{\ln\delta} \right\rfloor$ or at $m = \left\lceil 1 - \frac{1}{\ln\delta} \right\rceil$ (Part 5).

Finally, consider the case where $\delta \in (1 - \frac{1}{n_a - 1}, 1)$ and $a = 0$. Then the club congestion function becomes $h(m) = \delta^{m-1}$ where $\delta \in (1 - \frac{1}{n_a - 1}, 1)$. The club-size elasticity is $\eta_h(m) = m(\delta - 1)$ and the congestion function is inelastic since for $\delta > 1 - \frac{1}{n_a - 1}$ we get $\eta_h(m) > -1$ for every club size. By Lemma 4, $k_h(m)$ is increasing and therefore maximized at $m = n_a$. In addition, by Lemma B.2, by increasing a the club size that maximizes the DCV cannot decrease. Since n_a is the maximal size, then for $\delta \in (1 - \frac{1}{n_a - 1}, 1)$ and every legitimate value of a the DCV is maximized at $m = n_a$ (Part 6). □

B.2 The Stability of the Empty Environment

B.2.1 The Result

Proposition B.1. *Let E_n be the Empty Environment with n_a agents.*

1. *Let $n_a \geq 4$. Suppose $h(m)$ is the reciprocal congestion function. E_n is OCS if and only if $c \geq 1$.*

2. *Let $n_a \geq 4$. Suppose $h(m)$ is an exponential congestion function. Denote $b(\delta, n_a) = \frac{1}{n_a-2}(\delta - (n_a-1)\delta^{n_a-1})$ and let $\delta^\star(n_a)$ be the unique $\delta \in (0,1)$ such that $b(\delta, n_a) = \delta(1-2\delta)$.*

   (a) *Suppose $a \in [0, \min\{b(\delta, n_a), \delta(1-2\delta)\})$. E_n is OCS if and only if $c \geq a + \delta$.*

   (b) *Suppose that one of the following condition holds:*

      i. *$\delta \in (0, \delta^\star(n_a))$ and $a \in (b(\delta, n_a), 1-\delta)$.*

      ii. *$\delta \in (1 - \frac{1}{n_a-1}, 1)$.*

      E_n is OCS if and only if $c \geq (n_a-1)(a + \delta^{n_a-1})$.

   (c) *Suppose that $\delta \in [\frac{1}{2}, 1 - \frac{1}{n_a-1}]$ and $a = 0$. E_n is OCS if and only if $c \geq \max\{k_h(\lfloor 1 - \frac{1}{\ln \delta} \rfloor), k_h(\lceil 1 - \frac{1}{\ln \delta} \rceil)\}$.*

The third part of Proposition B.1 is a direct application of Lemma B.1. Under the exponential congestion function, each club size provides its members with different payoffs. The minimal costs for which the Empty Environment is OCS are determined by the most attractive deviation. In the case analyzed in Proposition B.1.2a, the congestion component is dominant and therefore the most attractive deviation is to the smallest club. If $\delta \in (0, \delta^\star(n_a))$, for the same δ when the non-congestion component is high enough, the grand club becomes the most attractive deviation.[50]

B.2.2 The Proof

Proof. The case of the reciprocal club congestion function is based on the DCV being a constant function that equals to 1. The case of the exponential club congestion function is based on Lemma B.1. The first case results from Part 2 and from $k_h(2) = a + \delta$. The second case is an implication of parts 3 and 6 (recall that $k_h(n_a) = (n_a - 1)(a + \delta^{n_a-1})$). The final case results from Part 5. □

C Existence of OCS m-Star Environments

C.1 One Analytic Result

Proposition 5 characterizes the membership fees for which an m-Star Environment is OCS. However, it does not provide a condition for the existence of such membership fees. Technically, for given n_a and m, Proposition 5 specifies upper and lower bounds on the membership fees for which an m-Star Environment is OCS but it does not guarantee that the upper bound is indeed greater than the lower bound. Claim C.1 identifies one case in which existence is guaranteed.

[50] One important implication of this discontinuity is on dynamic models where agents join the environment sequentially. Consider a dynamic model where the initial environment is the Empty environment, the clubs' rules follow the OCS rules and the membership fees are marginally high. Then, a tiny difference in the parameters of the congestion function or in the population size may lead to huge differences in the final environment's club composition.

Claim C.1. *Let $n_a > m \geq 2$ and let $h(\cdot)$ be a club congestion function. Denote $\gamma \equiv \frac{n_a-1}{m-1}$ and $l_h = \min\{k \in \mathbb{Z} | h(k) \leq h^2(m)\}$. Suppose $\gamma < m$. If*

$$J_h(m) \geq \max\{\max_{\gamma \geq k \geq 2} FNS_h(k,m), \max_{m \geq k > \gamma} FNI_h(k,m,n_a), \max_{\min\{l_h,n_a\} > k > m} FNL_h(k,m,n_a)\}$$

then a range of membership fees for which an m-Star Environment is OCS exists.

Proof. By Proposition 5, $J_h(m) = (m-1)[h(m+1) - h^2(m)]$. Since club congestion functions are assumed to be non-increasing, we get $h(m) \geq h(m+1) - h^2(m)$. Therefore, $(m-1)h(m) \geq (m-1)[h(m+1) - h^2(m)]$. Hence, $K_h(m) \geq J_h(m)$. If

$$J_h(m) \geq \max\{\max_{\gamma \geq k \geq 2} FNS_h(k,m), \max_{m \geq k > \gamma} FNI_h(k,m,n_a), \max_{\min\{l_h,n_a\} > k > m} FNL_h(k,m,n_a)\}$$

then by the second part of Proposition 5 a range of membership fees for which the m-Star Environment is OCS is guaranteed. $\square$

Claim C.1 provides a sufficient condition for the existence of membership fees for which an m-Star Environment is OCS in cases where the number of populated clubs is smaller than the size of the clubs (e.g. a 3-Star Environment with 5 agents). This condition is not vacant. Consider, for example, the case where $n_a = 13$ and $h(m) = 0.73 + 0.21^{m-1}$. We are guaranteed that a range of membership fees for which any 7-Star Environment is OCS exists since it can be shown that for a peripheral agent joining the other club is more attractive than any deviation to a new club.

C.2 Numerical Analysis

As noted above, Proposition 5 does not provide a condition for the existence of membership fees for which a given m-Star Environment is OCS. Figure 9 demonstrates the application of Proposition 5 to the question of existence of such membership fees in the case of 13 agents and an exponential club congestion function. In each of the six sub-figures, the shaded area presents the pairs of a (horizontal axis) and δ (vertical axis) for which the corresponding m-Star Environment is OCS for some membership fees (since $a + \delta < 1$ only the lower left triangle is relevant). In addition, in each sub-figure we indicate, in terms of the size of the new club, the deviation that determines the envelop of the area where no membership fees exists for which the corresponding m-Star Environment is OCS (restrictive intervals of the deviation are depicted as continuous while non-restrictive intervals are dotted).[51]
The upper leftmost sub-figure (the 2-Star Environment) summarizes Claim 4.4. The second part of Claim 5 could be recognized by the intersection of the shaded area with the Y-axis ($a = 0$) in the upper middle sub-figure (the 3-Star Environment) and the first part of Claim

[51]For each sub-figure (excluding the one for the Grand Club environment) we first calculated for each $k \in \{2, 3, ..., \min\{l_h, 13\}\}$ and for 1000 values of $\delta \in (0, 1)$ the set of as such that the upper bound is greater than the corresponding lower bound expression (using FNS_h, FNI_h or FNL_h). Claim C.1 guarantees that the calculation of $J_h(m)$ is unnecessary. Next, we calculated the intersection of all the sets derived in the first stage and presented it by the shaded area. The curves were derived similarly to the first stage procedure, except that the upper bound was set to be equal to the lower bound expression. For the Grand Club environment we repeated the same procedure using the lower bound specified in Proposition 4.

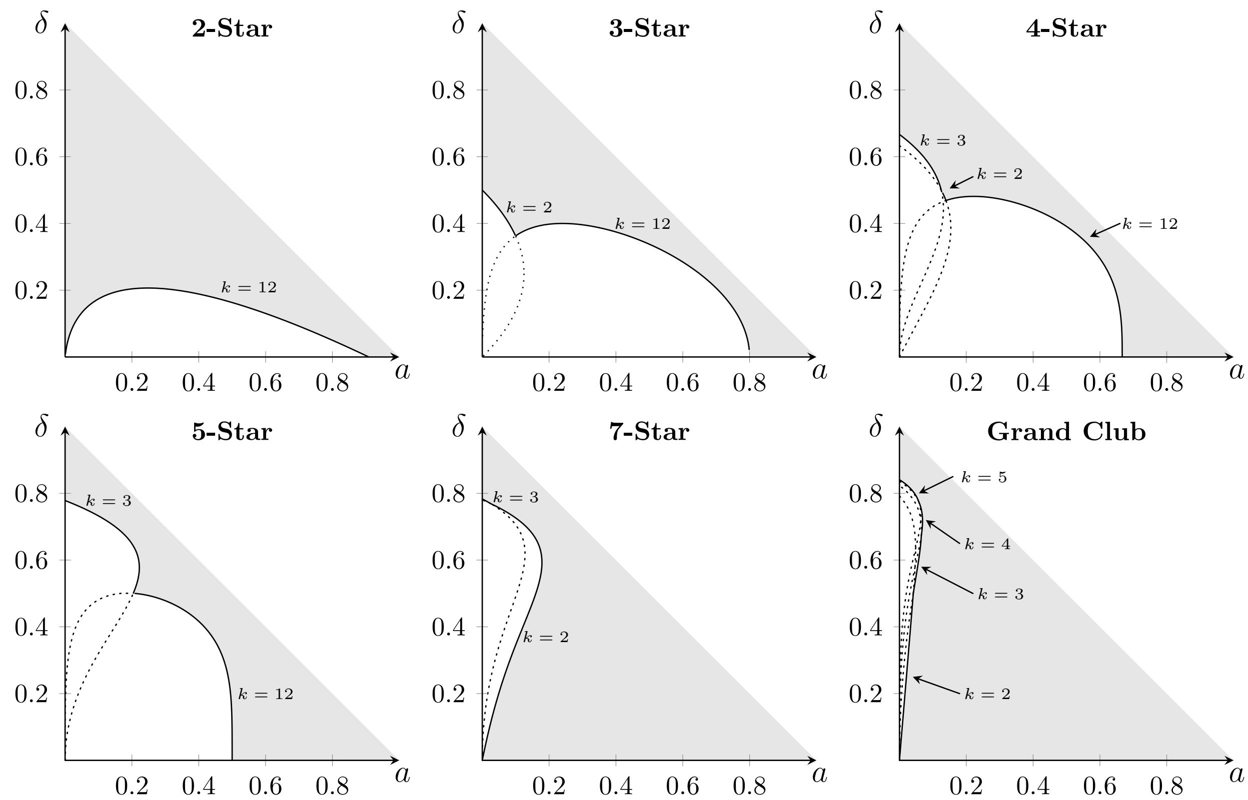


Figure 9: The existence of membership fees for which m-star environments are OCS when the club congestion function is exponential and $n_a = 13$.

3 could be recognized by the intersection of the non-shaded area with the Y-axis in the lower rightmost sub-figure (the Grand Club Environment).

The main insight provided by Figure 9 is that holding δ constant, the effect of a on stability is non-monotonic. This reflects the complicated lower bound conditions in Proposition 5, where the non-differentiable points denote changes in the effective lower bound. The cases of the 3-Star, 4-Star and 5-Star environments demonstrate the intuition very nicely. When the non-congested parameter is low, the effective bound is induced by a deviation of a small coalition since the effect of congestion is dominant and therefore should be minimized. However, when the non-congested parameter is high, the effective bound is a deviation of a large coalition, since congestion is relatively less important than the multiplicative effect introduced by a. Since the multiplicative effect strengthens with the size of the club, the most attractive deviation is to a club that includes all peripheral agents. Note that the first consideration is missing from the sub-figure of the 2-Star Environment since agents in this environment suffer no congestion. Similarly, the second consideration is missing from the sub-figure of the Grand Club Environment since the multiplicative effect is maximized (the reasoning is similar for the 7-Star Environment).