Structural holes in social networks

Sanjeev Goyal[*] Fernando Vega-Redondo[†]

Preliminary version: June 2004

Abstract

We consider a setting where every pair of players can exchange goods and that this exchange yields a fixed surplus. This exchange can take place only if the players have a connection. This connection can be direct in which case the two players split the surplus equally. We suppose that exchange can also occur if there is an indirect connection, in other words where there is one or more players in the path between the two players in question. In case of indirect connection, we require that every player *essential* to the transaction gets an equal share of the surplus.
This formulation gives rise to three types of incentives for forming links. One, individuals would like to join the network so as to create exchange possibilities which in turn create surpluses. Two, there are rewards from being essential: players would like to place themselves between unconnected others to extract rents from intermediation. The third incentive to form links arises out of the desire to circumvent intermediaries. Our main result is that, so long as costs of forming links are not very small, the unique equilibrium network is a star. In this network a single agent is essential for all bilateral exchange and there is significant inequality in payoff distribution.

---

[*] Department of Economics, University of Essex, sgoyal@essex.ac.uk.

[†] Department of Economics, Universidad de Alicante and University of Essex, vega@merlin.fae.ua.es.

1 Introduction

There is a growing consensus among social researchers that a proper understanding of many socio-economic phenomena cannot be dissociated from the underlying social network in which they are embedded. This pertains to a wide variety of problems ranging from labor markets (Granovetter, 1973), to industrial R&D (Gulati *et al.,* 2000), technological diffusion (Valente, 1996), scientific collaboration (Newman, 2001 and Goyal *et al.*, 2004), modern inter-industrial trade (Kranton and Minehart, 2001)), or even trade in historical times (Greif, 1994). In all these cases, the research has understanding how the underlying pattern of connections (i.e. the "architecture" of the social network) impinges on the outcome of the process – employment, trade, or the scope for technological diffusion.

Clearly, one of the features of network architecture that will have an important bearing on the network benefits accruing to the different individual concerns their relative importance in connecting others. Thus, for any particular agent, one would expect that the payoff she is able to muster would depend on how indispensable she is in bridging valuable connections among other individuals.

These ideas are at the heart of the notion of *structural holes* introduced by Ronald Burt (1992). Over the past decade, Burt and several others have elaborated on this notion in both theoretical and empirical terms. In particular, there is extensive empirical work done on this subject. This work may be summarized as follows: once we control for individual characteristics we are left with a substantial variance in individual performance in organizations (measured in terms of salaries and promotions). The differences in structural location of otherwise similar individuals – in particular whether they bridge structural holes in the social network – explains a significant part of this variance. Given the substantial rewards from network positioning, it seems natural to postulate that an individual will make investments to become a "bridge" (critical to the interconnectedness of the network) while other individuals have an incentive to prevent this from happening by forming alternative routes to interaction. In this paper, we ask what will be the shape of networks when (ex-ante identical) individuals can rationally form links and therefore create positional advantages for themselves while others have an equal incentive to prevent them from acquiring this positional power.

We develop a simple model of network formation to address this question. We consider a setting where every pair of players can exchange goods and that this exchange yields a fixed surplus. This exchange can take place only if the players have a connection. This connection can be direct in which case the two players split the surplus equally. We suppose that exchange can also occur if there is an indirect connection, in other words where there is one or more players in the path between the two players in question. In case of indirect connection, we have to formulate a procedure through which the surplus is divided among the players. This leads us to introduce the idea of *essential* players. We say that player k is essential to a exchange between i and j if he lies on every path between them in the network. In general, there will be more than one essential players on a path. It follows that each of the essential players have the power to block the exchange. This idea leads us to suppose that the surplus in the exchange between i and j is split equally between the essential players and i and j.

This formulation gives rise to three types of incentives. The first incentive is the standard access advantage: individuals would like to join the network so as to create exchange possibilities which in turn create surpluses. The second incentive is related to the rewards from being essential: players would like to place themselves between unconnected others to extract rents from intermediation. The third incentive arises out of the desire to not share surpluses with intermediaries; in other words, individuals will try and circumvent essential players to retain more of the surplus to themselves.

Our main result is that, as long as costs are not very low, *the unique equilibrium network is a star, with a single individual acting an essential intermediary to all bilateral connections.* There are two aspects of the result that we would like to note: one, the uniqueness of equilibrium and two, the significant inequality in payoffs between players who are ex-ante identical. Thus an extreme version of strategic positioning – with attendant payoff differences – is the only possible outcome in a setting with ex-ante identical individuals. We now briefly sketch the main arguments underlying this finding.

The main arguments derive from the incentives mentioned above. We first build on the access benefits idea to argue that an equilibrium network will either be empty or connected. We then examine the types of connected networks that are strategically stable. The third incentive we mentioned above suggests that players will wish to minimize the number of essential players that lie between them and others. This means that in minimal networks with long paths end-players want to disconnect from their current partners and instead link with players who are more central on the path. We then note that these central players are willing to form extra links with these end-players since this makes them even more central and yields them higher essentiality benefits. Thus there is a natural tendency for minimal networks to become agglomerated around central players. We then turn to networks in which there are cycles. In principle a cycle containing all players appears to be very attractive. No player is essential is such a network and since the network is connected everyone can transact with every other player. However, this network is vulnerable to the pressures of the second incentive: a desire of players to become central and essential. We show that two players who are furthest apart from each other in the cycle can deviate in a coordinated way and strictly increase their payoffs. This deviation involves the following strategy: they form a link with each other and delete one of their existing links so that the resulting network is a line, and they are the central players of this new network. In this line network, we show that while they derive lower access payoffs, the fall in access payoffs is more than compensated for by the increase is the essentialness payoffs. Thus this deviation is attractive and destroys the egalitarian cycle network. A similar argument destroys more complex networks which contain several cycles.

Our paper is related to research in economics as well as research in sociology. In recent years, the theory of network formation has been an active area of research.[1] The principal contribution of the present paper to this literature is a new model of network formation. In particular, the innovation lies in the introduction of two ideas: the first idea is that goods being exchanged in the network are rival goods, so that if one person gets more than other

---

[1] Early work includes Aumann and Myerson (1988); for more recent work see Bala and Goyal (1999), Jackson and Wolinsky (1996), and Dutta, van Den Nouweland and Tijs (1995)).

players get less while the second idea is the notion of *essentialness*. This notion of essentialness defines exactly which players are critical for a certain value generating transaction and the number of such players determines the payoff division between the different players in the network. These ideas are introduced within the general framework of the connections model (Bala and Goyal, 2000; Jackson and Wolinsky, 1996).

We now briefly relate our paper to the work in sociology on structural holes. The large body of empirical work on economic advantages of networks positions which we cited earlier generated a natural question: how is it that such advantages arise in organizations and that they remain stable over time in a setting with rational actors who can figure out the rewards from structural positions? Our paper shows that this state of affairs is consistent with a model of rational players who strategically seek to create positional advantages for themselves and also have an incentive in preventing others from becoming central. To the best of our knowledge our model is the first formal model to illustrate the strategic stability of structural holes and the large inequality in payoffs that go with it.

The paper is organized in four sections. The next section lays out the basic model. Section 3 presents the main results on equilibrium networks where both single players as well as any pair of players can implement any deviation. Section 3 briefly explores robustness of the main results to an alternative notion of strategic stability where bilateral commitment is not presumed. Section 4 concludes.

2 The Model

Link formation game: Let $N = \{1, 2, ..., n\}$ denote a finite set of ex-ante identical players. We shall assume that $n \geq 3$. Every player makes an announcement of intended links. An intended link $s_{i,j} \in \{0, 1\}$, where $s_{i,j} = 1$ means that player i intends to form a link with player j, while $s_{i,j} = 0$ means that player i does not intend to form such a link. Thus a strategy of player i is given by $s_i = \{\{s_{i,j}\}_{j \in N \setminus \{i\}}\}$. Let S_i denote the strategy set of player i. A link between two players i and j is formed if and only if $s_{i,j} = s_{j,i} = 1$. We denote the formed link by $g_{i,j} = 1$ and the absence of a link by $g_{i,j} = 0$. A strategy profile

$s = \{s_1, s_2, ..., s_n\}$ therefore induces a network $g(s)$. For expositional simplicity we shall often omit the dependence of the network on the underlying strategy profile. A *network* $g = \{(g_{i,j})\}$ is a formal description of the pair-wise links that exist between the players. We let $\mathcal{G}$ denote the set of all possible networks (i.e. all undirected networks with n vertices). We also let $\mathcal{N}_i(g) = \{j \in N : j \neq i,\ g_{ij} = 1\}$ be the set of players with whom player i has a link in the network g, and let $\eta_i(g) = |\mathcal{N}_i(g)|$ denote the cardinality of this set.

Given a network g, $g + g_{i,j}$ denotes the network obtained by replacing $g_{i,j} = 0$ in network g by $g_{i,j} = 1$, while $g - g_{i,j}$ denotes the network obtained by replacing $g_{i,j} = 1$ in network g by $g_{i,j} = 0$. There exists a *path* between i and j in a network g if either $g_{i,j} = 1$ or if there is a distinct set of players $\{i_1, \ldots, i_n\}$ such that $g_{i,i_1} = g_{i_1,i_2} = g_{i_2,i_3} = \cdots = g_{i_n,j} = 1$.

We shall say that a player k is essential for i and j if player k lies on every path that goes between i and j in the network. The indicator function $I_{i;j,k}$ takes value one if player i is essential to players j and k, and takes value zero otherwise. We also define $E(j,k) = \sum_{i \in N} I_{i;j,k} to$ be the essential players between j and k in network g. To illustrate, the number of essential players between j and k is zero if the players have a direct link; alternatively, if players are located around a circle then for every pair of players j and k the number of essential players is zero again. On the other hand, in a star every pair of peripheral players has a single common essential player, namely the center of the star.

There are two types of advantages that players get from being part of a network: there is the direct advantage from being able to access others and there is the indirect advantage that arises out of being an intermediary between others. This later advantage is captured in a simple way by the extent of essentialness of a player.

Given a strategy profile $s = \{s_1, s_2, ..., s_n\}$, the (net) payoffs to player i are given by:

$$\Pi_i(s_i, s_{-i}) = \sum_{j \in N_i(g)} \frac{1}{E(i,j;g) + 2} + \sum_{j,k \in N} \frac{I_{\{i;j,k\}}}{E(j,k;g) + 2} - \eta_i(g)c \tag{1}$$

We study the architecture of networks that are strategically stable and those that are efficient. Our notion of strategic stability is a refinement of Nash equilibrium and allows for coordinated two-person deviations.

Definition 2.1 *A strategy profile* $\mathbf{s}^*$ *is a (bilateral deviation proof) equilibrium if the following conditions hold:*

1. *For every* $i \in N$, $\Pi_i(s^*) \geq \Pi_i(s_i, s^*_{-i})$, *for all* $s_i \in S_i$.
2. *For any pair of players,* $i, j \in N$ *the following conditions hold:*
    - *There is no strategy pair* (s_i, s_j) *which satisfies:*

$$\Pi_i(s_i, s_j, \mathbf{s}^*_{\mathbf{-i-j}}) > \Pi_i(s^*_i, s^*_j, \mathbf{s}^*_{\mathbf{-i-j}})$$

and

$$\Pi_j(s_i, s_j, \mathbf{s}^*_{\mathbf{-i-j}}) > \Pi_j(s^*_j, s^*_j, \mathbf{s}^*_{\mathbf{-i-j}}).$$

This notion generalizes the original formulation of pair-wise stability due to Jackson and Wolinsky (1996) by allowing pairs of players to form and delete links simultaneously. We shall use the term 'equilibrium' to refer to bilateral deviation proof equilibrium. An equilibrium is said to be strict if the following holds.

Definition 2.2 *A strategy profile* $\mathbf{s}^*$ *is a strict equilibrium if the following conditions hold:*

1. *For every* $i \in N$, $\Pi_i(s^*) > \Pi_i(s_i, s^*_{-i})$, *for all* $s_i \in S_i$, $s_i \neq s^*_i$.
2. *For any pair of players,* $i, j \in N$ *the following conditions hold:*
    - *There is no strategy pair* $(s_i, s_j) \neq (s^*_i, s^*_j)$ *that satisfies:*

$$\Pi_i(s_i, s_j, \mathbf{s}^*_{\mathbf{-i-j}}) \geq \Pi_i(s^*_i, s^*_j, \mathbf{s}^*_{\mathbf{-i-j}})$$

and

$$\Pi_j(s_i, s_j, \mathbf{s}^*_{\mathbf{-i-j}}) \geq \Pi_j(s^*_j, s^*_j, \mathbf{s}^*_{\mathbf{-i-j}}).$$

Efficiency: The total welfare generated by a network g is given by $W(g) = \sum_{i \in N} \Pi_i(g)$. A network g is socially efficient if $W(g') \geq W(g)$, for all $g \in \mathcal{G}$.

2. Networks: A network is *connected* if there exists a path between any pair $i, j \in N$. A network, $g' \subset g$, is a *component* of g if for all $i, j \in g'$, $i \neq j$, there exists a path in g' connecting i and j, and for all $i \in g'$ and $k \in g$, $g_{i,k} = 1$ implies $k \in g'$. A component $g' \subset g$ is *complete* if $g_{i,j} = 1$ for all $i, j \in g'$.

Two networks g and g' are said to have the same architecture if one network can be obtained from the other by a permutation of the players' labels.

We shall say that a network is symmetric if every player has the same number of links, i.e., $\eta_i(g) = \eta(g)\ \forall i \in N$. We refer to η as the *degree* of the network. It should be noted that if the number of players is even, then a symmetric network of every degree is possible; this is not true when the number of players is odd. In some parts of the analysis, we explore symmetric networks and then we will be (implicitly) assuming that the number of players in the game is even. The *complete* network, g^c, is a symmetric network in which $\eta = n - 1$, $\forall i \in N$, while the *empty* network, g^e, is a symmetric network in which $\eta = 0$, $\forall i \in N$. We shall say that a network is asymmetric if there is at least one pair of players who have different number of links. Let $N_1(g), N_2(g), ..., N_m(g)$ be a partition of players corresponding to the number of links that players have, i.e., $i, j \in N_k(g)$, $k = 1, 2, ..m$, if and only if $\eta_i(g) = \eta_j(g)$. We note that k here refers to the order in the partition and not the precise number of links that players have. An inter-linked stars architecture has at least two members in the above partition, and the maximally and minimally linked groups, $N_1(g)$ and $N_m(g)$ respectively, satisfy the following two conditions: (i). $\eta_i(g) = n - 1$ for $i \in N_m(g)$ and (ii). ${}_i(g) = N_m(g)$ for $i \in N_1(g)$. The star network is a special case of such an architecture with $|N_m(g)| = 1$ and $|N_1(g)| = n - 1$.

A hybrid cycle-star network consists of a cycle of players $C = \{i_1, i_2, ...i_k\}$ with $g_{i_1,i_2} = g_{i_2,i_3} = ... = g_{i_k,i_1} = 1$, and a group of players $N \backslash C$, with the property that there is some player $i_x \in C$ such that $g_{j,i_x} = 1$ for all $j \in N \backslash C$.

3 Results

We start by characterizing efficient networks.

Proposition 3.1 *If* $c < n/4$ *then an efficient network is minimally connected, while if* $c > n/4$ *then an efficient network is empty.*

Proof: It is easy to see that an efficient network cannot be partially connected and that it must be minimal. The aggregate payoff from a minimally (fully) connected network is $(n-1)[n/2 - 2c]$, while the payoff from the empty network is 0. It then follows that a minimal connected network is efficient if and only if $c < n/4$, while an empty network is efficient if $c > n/4$. ■

We now turn to the study of equilibrium networks. Our first finding is that an equilibrium network is either connected or empty.

Proposition 3.2 *An equilibrium network is either empty or connected.*

The proof of this result makes use of the following technical result which is of independent interest. For any link $g_{i,j} = 1$ in g, define $M_i(g_{i,j}; g) = \Pi_i(g) - \Pi_i(g - g_{i,j})$.

Lemma 3.1 *Consider any network* g. *If* $g_{i,j} = 1$ *and the link is critical then* $M_i(g_{i,j}; g) = M_j(g_{i,j}; g)$.

Proof: By hypothesis $g_{i,j}$ is critical, and so it follows that i and j lie in different components in the network $g - g_{i,j}$. Let C_i be the component that contains i, and let C_j be the network that contains j. We can write $M_i(g_{i,j}; g)$ as follows:

$$
\begin{aligned}
& M_i(g_{i,j}; g) \\
&= \frac{1}{2} + \sum_{k \in C_j \setminus \{j\}} \frac{1}{E(i,k;g)+2} + \sum_{l \in C_i \setminus \{i\}} \sum_{k \in C_j \setminus \{j\}} \frac{1}{E(l,k;g)+2} + \sum_{l \in C_i \setminus \{i\}} \frac{1}{E(l,j;g)+2} - c
\end{aligned}
$$

where the first two terms refer to access benefits while the latter two terms refer to essentiality benefits. Similarly, we can write the marginal payoffs of player j form link $g_{i,j}$ as:

$$\begin{aligned} & M_j(g_{i,j};g) \\ = & \ \frac{1}{2} + \sum_{l \in C_i \setminus \{i\}} \frac{1}{E(j,l;g)+2} + \sum_{l \in C_i \setminus \{i\}} \sum_{k \in C_j \setminus \{j\}} \frac{1}{E(l,k;g)+2} + \sum_{k \in C_j \setminus \{j\}} \frac{1}{E(i,k;g)+2} - c \end{aligned}$$

It follows then that $M_i(g_{i,j};g) = M_j(g_{i,j};g)$ and the proof is complete. ■

Equipped with this lemma we can now complete the proof of the connectedness result.

Proof of Proposition 3.2: Let g be a non-empty equilibrium, and suppose it is not connected. Let C_1 be the largest component in g. If C_1 contains 2 players then it is immediate that both players in this component have a strict incentive to form a link with any other player not in the component. Given Lemma 3.1 it follows that every player outside C_1 in g also has an incentive to form a link with these players. Thus in the rest of the proof we concentrate on the case where C_1 contains more than 2 players. This also means that there is at least one player (say) $i \in C_1$ such that $\eta_i \geq 2$. Let $N_i^m(g) = |\{j \in N_i(g) : E(i,j) = m)\}|$ be the players whom i accesses via m essential players and let $\eta_i^m(g) = |N_i^m(g)|$. The payoffs of this player i in network g are then given by:

$$\frac{\eta_i^0(g)}{2} + \frac{\eta_i^1(g)}{3} + \frac{\eta_i^2(g)}{4} + ... + \frac{\eta_i^L(g)}{L+2} - \eta_i(g)c \tag{2}$$

for some $L \leq n-2$.

Since g is an equilibrium, it follows then

$$\frac{1}{\eta_i(g)} \left[\frac{\eta_i^0(g)}{2} + \frac{\eta_i^1(g)}{3} + \frac{\eta_i^2(g)}{4} + ... + \frac{\eta_i^L(g)}{L+2} \right] \geq c. \tag{3}$$

Now let us examine marginal returns for player $j \notin C_1$ from forming a link with some player i. Suppose, for simplicity, that player j is a singleton component. Then the *marginal returns* to j are given as follows:

$$\frac{\eta_j^0(g+g_{i,j})}{2} + \frac{\eta_j^1(g+g_{i,j})}{3} + \frac{\eta_i^2(g+g_{i,j})}{4} + ... + \frac{\eta_i^L(g+g_{i,j})}{L+2} - c. \tag{4}$$

Note now that for every $k \in C_1 \backslash \{i\}$, $E(j,k;g+g_{i,j}) = E(i,k;g)+1$. Thus $\eta_j^m(g+g_{i,j}) = \eta_i^{m-1}(g)$ for every $m \geq 1$ and $\eta_j^0(g+g_{i,j}) = 1$. Using these facts we can write the marginal returns of player j from the link with j as follows:

$$\frac{1}{2} + \frac{\eta_i^0(g)}{3} + \frac{\eta_i^1(g)}{4} + ... + \frac{\eta_i^L(g)}{L+3} - c. \tag{5}$$

Now we argue that

$$\begin{aligned} & \frac{1}{2} + \frac{\eta_i^0(g)}{3} + \frac{\eta_i^1(g)}{4} + ... + \frac{\eta_i^L(g)}{L+3} \\ & \quad > \frac{1}{2}\left[\frac{\eta_i^0(g)}{2} + \frac{\eta_i^1(g)}{3} + \frac{\eta_i^2(g)}{4} + ... + \frac{\eta_i^L(g)}{L+2}\right] \\ & \quad \geq \frac{1}{\eta_i(g)}\left[\frac{\eta_i^0(g)}{2} + \frac{\eta_i^1(g)}{3} + \frac{\eta_i^2(g)}{4} + ... + \frac{\eta_i^L(g)}{L+2}\right] \\ & \geq c \end{aligned}$$

The first inequality is immediate, while we use $\eta_i(g) \geq 2$ in deriving the second inequality and equation (3), in deriving the final inequality. We now apply Lemma 1 to conclude that player i also has a strict incentive to form a link with j, given that all existing links are retained. But note that given that link $g_{i,j}$ is formed, player i has no incentive to delete any of his erstwhile links since (roughly speaking) the marginal returns from each of these links has actually increased. Thus players i and j have a strict incentive to form an additional link and this link is a credible two-person deviation. These arguments extend directly to cover the case where j belongs to a non-singleton component. Thus g is not a equilibrium, a contradiction that completes the proof. ■

We now look more closely at the architecture of equilibrium networks. The following result gives examples of two simple networks that arise in equilibrium.

Proposition 3.3 *If* $1/6 < c < 1/2 + (n-2)/6$ *then a star is an equilibrium, while if* $c > 1/2$ *then the empty network is an equilibrium.*

Proof: We first prove that the star is an equilibrium for the parameter range specified. First consider the incentives of the center. The payoffs of the center are given by

$$\frac{n-1}{2} + \frac{1}{3}\frac{(n-1)(n-2)}{2} - (n-1)c. \tag{6}$$

Note that the payoffs from a subset of links k are given by

$$\frac{k}{2} + \frac{1}{3}\frac{(k)(k-1)}{2} - kc. \tag{7}$$

It can be verified that the marginal payoffs are increasing with respect to k, and so the optimal number of links is either zero or $n-1$. It follows then that center maintains all links if $c < 1/2 + (n-2)/6$. We next examine the incentives of peripheral players. Their payoff in the star is given by $1/2 + (n-2)/3 - c$. The payoffs from having two links are given by $1/2 + 1/2 + (n-3)/3 - 2c$. It then follows that the peripheral player will choose to maintain one link with the center if $1/6 < c < 1/2 + (n-2)/3$. This proves that the star is an equilibrium if $1/6 < c < 1/2 + (n-2)/6$. It is straightforward to see that the empty network is an equilibrium for $c > 1/2$. ■

The star is an equilibrium for a wide range of parameters and this is due to the fact that centrality generates large payoffs from essentialness. However, a star also exhibits an extreme form of essentialness, with only one player being essential for all pairs of players. This raises the question: are there other – perhaps more egalitarian – network architectures that can arise in equilibrium? The following results provide a complete answer to this question, for large societies.

Theorem 3.1 *Suppose that* n *is large. If* $1/6 < c < 1/2$ *then any strict equilibrium network is a star while if* $1/6 < c < 1/2 + (n-2)/6$ *then any strict equilibrium network is a star. Finally, if* $c > 1/2 + (n-2)/6$ *then the unique equilibrium network is empty.*

This result provides a complete characterization of equilibrium networks for large societies. The proof proceeds through five steps. The first step shows that any minimal equilibrium network is a star, for large n. The second step shows that there cannot be two or more cycles in an equilibrium network. The third step shows that there cannot be a single cycle containing all players in equilibrium. The fourth step shows that a non-empty equilibrium network with a cycle must be a hybrid cycle-star network. The final step shows that a hybrid cycle-star network cannot be an equilibrium for large n. This means that for large n the only candidates for equilibrium are the star and the empty network.

Lemma 3.2 *Suppose that* n *is large. Any equilibrium network that is minimally connected is a star.*

Proof: Consider g which is minimal and not a star. Let i and j be two end-players who are furthest apart; given our hypothesis it follows that $E(i,j;g) \geq 2$. Since i and j are end-players, $\eta_i(g) = \eta_j(g) = 1$. Let $g_{i,x} = g_{j,y} = 1$, where $x \neq y$ clearly. The payoff of player i in network g is given by

$$\pi_i(g) = \frac{1}{2} + \sum_{k=1}^{n-1} \frac{\eta_i^k(g|j)}{k+2} + \frac{1}{E(i,j;g)+2} - c \tag{8}$$

where $\eta_i^k(g|j)$ is the number of players, apart from j, each of whom has exactly k essential players in the unique path with i. Similarly, the payoff to player j from network g is given by

$$\Pi_j(g) = \frac{1}{2} + \sum_{k=1}^{n-1} \frac{\eta_j^k(g|i)}{k+2} + \frac{1}{E(i,j;g)+2} - c \tag{9}$$

Suppose without loss of generality that $\Pi_i(g) \geq \Pi_j(g)$. Now note that if player j were to delete his link with player y and instead form a link with player x then his payoff would be given by

$$\begin{aligned}
&\Pi_j(g - g_{j,y} + g_{j,x}) \\
&= \frac{1}{2} + \sum_{k=1}^{n-1} \frac{\eta_j^k(g - g_{j,y} + g_{i,x}|i)}{k+2} + \frac{1}{3} - c \\
&= \frac{1}{2} + \sum_{k=1}^{n-1} \frac{\eta_i^k(g - g_{j,y} + g_{j,x}|j)}{k+2} + \frac{1}{3} - c \\
&> \Pi_i(g) \\
&\geq \Pi_j(g)
\end{aligned}$$

The first equality arises out of the payoffs from direct link with player x, the payoffs from players further away and lastly the payoffs from indirect link with player i. The second equality follows from noting that players i and j are now symmetrically located in the network and so $\eta_j^k(g - g_{j,y} + g_{i,x}|i) = \eta_i^k(g - g_{j,y} + g_{j,x}|j)$, for each k. The first inequality follows from noting that $E(i,j;g) \geq 2$ and $E(i,j,g - g_{jy}, g + g_{jx}) = 1$, while the final inequality follows from the hypothesis above. We have therefore shown that player j has a strict incentive to switch links. We next show that player x also has strict incentives to form links with j if n is large enough.

Let $E = E(i,j;g) - 1$ be the number of essential players between x and j. There are three sources of gains for x from a link with j: he gains by virtue of reducing distance between j and x, by reducing distance between j and N_1, the set of players who access j via x in network g, and by becoming essential to all paths between j and players N_2, who could earlier access j without the aid of x. The first gain is given by $1/2 - 1/(E+2)$. Let $n_1 = |N_1|$ and $n_2 = |N_2| = n - n_1 - 1$. The second gain is bounded below by $n_1[1/3 - 1/(E+3)]$, while the third gain is bounded below by $n_2[1/(E+3)]$. Thus the marginal gain for player x is bounded below by

$$\left[\frac{1}{2} - \frac{1}{E+2}\right] + n_1 \left[\frac{1}{3} - \frac{1}{E+3}\right] + (n - n_1 - 1)\frac{1}{E+3}. \tag{10}$$

We now note that for any c there is an upper bound on the number of essential players $E(c)$ that can be supported in equilibrium between any two players. Suppose therefore that $1 \leq E < E(c)$. Now it follows from the above expression that for large enough n the marginal gain to player x from linking with j would exceed c. This completes the proof of the lemma. ■

We now consider non-minimal networks. The following result shows that there cannot be more than one cycle in a strict equilibrium.

Lemma 3.3 *There can be at most one cycle in a strict equilibrium network.*

Proof: Suppose g is an equilibrium network and there are two or more cycles in it. Let $C_1 = (i_1, i_2, ...i_n)$ be players in cycle 1, and let $C_2 = (j_1, j_2, ...j_m)$ be players in cycle 2. Since g is connected it follows that there are two possibilities: (1) cycles have common players and (2) cycles have no common players. We take these up in turn.

(1). Cycles have players in common: If there is a single common player i_1 in the two cycles then it is easy to see that the partners of i_1 (say) $i_2 \in C_1$ and $j_2 \in C_2$ have a strict incentive to delete their links with i_1 and instead form a link with each other. Consider next the case with two or more players in common. Let $(i_1, i_2, ...i_k)$ be the players in common. Suppose that $k \geq 3$; the case of $k = 2$ is simple and omitted. Then there exist players i_1, i_x, and j_y with the following properties: $i_x \in C_1$ but $i_x \notin C_2$, while $j_y \in C_2$ and $j_y \notin C_1$ and $g_{i_1,i_x} = g_{i_1,j_y} = g_{i_1,i_2} = ... = g_{i_{k-1},i_k} = 1$. Note also that like player i_1, i_k must again have links with a player who belongs to one of C_1 and C_2 only. It then follows that i_{k-1} and j_y

have at least a weak incentive to delete their current link with i_k and i_1 respectively and instead form a link with each other. It then follows that g is not a strict equilibrium.

2. Cycles have no common players. Since g is an equilibrium network it is connected and so there exists a path between the two cycles. Let $(i_1, i_2..i_k)$ be members of such a path with $i_1 \in C_1$ while $i_k \in C_2$. Suppose $g_{i_1,i_x} = 1$ and $g_{i_k,j_y} = 1$, where $i_x \in C_1$ and $j_y \in C_2$. Now it is easy to use a variant of the earlier argument for case 1 above to show that players i_x and j_y have a strict incentive to delete their link with i_1 and i_k and instead form a link with each other. The proof of the lemma is complete. ■

We now turn to the networks with a single cycle.

Lemma 3.4 *Suppose* $n \geq 4$.[2] *A cycle containing all players cannot be sustained in equilibrium.*

Proof: Consider two player i and j who are furthest apart in terms of geodesic distance in the cycle. Now consider the deviation in which each of the players deletes one link and they form a link with each other in such a way that they create a line. So, for instance, if n is even then there are $(n-2)/2$ players to one side of player i and $(n-2)/2$ players to the other side of player j in the line created. We now show that players i and j will strictly increase their payoff with this coordinated deviation.

We proceed in two steps: the first step is to show that individual payoffs are strictly increasing as we move toward the center in the line. The payoffs of an individual player consist of two components, the returns from accessing others and the returns from being essential on paths between pairs of other players. Number the players on a line as 1, 2, ...n. We show that the payoffs of player $l+1$ are greater than the payoffs of player l, where $2 \leq l < n/2$. The access returns to player l are given by

$$\frac{1}{l} + \frac{1}{l-1} + \ \dots \ + \frac{1}{2} + \frac{1}{2} + \ \dots \ \frac{1}{n-l+1} \tag{11}$$

---

[2] We note that for $n = 3$ a complete network can be sustained in equilibrium for $c < 1/6$.

while the access returns to player $l+1$ are given by

$$\frac{1}{l+1}+\frac{1}{l}+\ ...\ \frac{1}{2}+\frac{1}{2}+\ ...\ \frac{1}{n-l} \tag{12}$$

It now follows that access returns for player $l+1$ are larger than access returns for player l if $l < n/2$.

We now turn to the returns from being essential. The returns to player l from being essential consist of two parts: the returns from being essential to pairs of players i and j where $i < l$ and $j > l+1$, and a second part which consist of returns from being essential for pairs $i < l$ and player $l_1 + 1$. The essentialness payoff to player l can be written as follows:

$$\sum_{i=1}^{l-1}\sum_{j=l+2}^{n}\frac{1}{E(i,j;g)+2}+\sum_{i=1}^{l-1}\frac{1}{E(i,l+1;g)+2} \tag{13}$$

Similarly, the essentialness payoffs to player $l+1$ consist of two parts: those that arise out of linking players $i < l$ with players $j > l+1$ and those that arise out of linking player l with players $j > l+1$. These payoffs can be written as follows:

$$\sum_{i=1}^{l-1}\sum_{j=l+2}^{n}\frac{1}{E(i,j;g)+2}+\sum_{j=l+2}^{n}\frac{1}{E(l,j;g)+2} \tag{14}$$

The first part of the essentialness payoffs to the two players are equal, while the second part of the payoffs are greater for player $l+1$ if $l < n/2$.

Note that players $1 < l < n$ all have two links in a line network and hence have the same costs. Combining the payoffs from access and essentialness we can then conclude that payoffs to player $l+1$ are greater than payoffs to player l, so long as $1 < l < n/2$. A similar argument shows that payoffs to player k are larger than payoffs of player $k+1$ if $n/2 \leq k < n-1$. This completes our argument that payoffs are increasing as we move toward the center of the line.

The above argument shows that the central players i and j earn maximum payoffs among all non-corner players. We now show that these payoffs must exceed the payoffs that players i and j get in the cycle. Suppose not; then from the above argument, it follows that the total payoffs of all the non-corner players are smaller than the payoffs they get in the cycle. The payoffs to a corner player are

$$\sum_{k=2}^{n} \frac{1}{k} - c. \tag{15}$$

The total payoffs in the line are then

$$\begin{aligned}
& W(g^L) \\
\leq \; & (n-2)\left[\frac{n-1}{2} - 2c\right] + 2\left[\sum_{k=2}^{n} \frac{1}{k} - c\right] \\
= \; & (n-2)\left[\frac{n-1}{2}\right] + 2\left[\sum_{k=2}^{n} \frac{1}{k}\right] - c[2(n-2)+2] \\
= \; & (n-2)\left[\frac{n-1}{2}\right] + 2\left[\sum_{k=2}^{n} \frac{1}{k}\right] - 2c(n-1) \\
< \; & (n-2)\left[\frac{n-1}{2}\right] + 2\left[\frac{n-1}{2}\right] - 2c(n-1) \\
= \; & W(g^L)
\end{aligned}$$

where the first inequality follows from the above observations, the second inequality holds because the corner players earn a strictly lower (gross) payoff as compared to their (gross) payoff in the cycle while the final equality follows from direct computation. This generates a contradiction and so players i and j must earn a higher payoff in the line than in the cycle. This means they have a strict incentive to deviate from the cycle and that completes the proof of the lemma. ■

We now move to the fourth step of the proof and show that any equilibrium network with a cycle must be a hybrid star-cycle network.

Lemma 3.5 *Consider an equilibrium network which has a cycle. It must then have the hybrid star-cycle architecture for large enough n.*

Proof: Let C be the set of players in the cycle. We first show that for any pair of players $i, j \notin C$, $E(i, j; g) \leq 1$ in an equilibrium. The first step is to note that given a cost of link formation c there is a maximum number of essential players $E(c)$ that can be sustained in equilibrium where $E(c)$ solves:

$$\frac{E(c)+1}{2} - \left[\frac{1}{2} + \frac{1}{3} + ... + \frac{1}{E(c)+2}\right] = c \tag{16}$$

In other words, if $E(i, j; g) \geq E(c)$ then players i and j have a strict incentive form a direct link. Suppose therefore that g is an equilibrium network and focus on two furthest off end-players i and j such that $E(i, j; g) \leq E(c)$. If $2 \leq E(i, j; g)$ then the argument used in Lemma 3.2 implies that one of the players i or j has a strict incentive to switch. Suppose that player j wishes to switch and connect to player z where $g_{i,z} = 1$. We will now demonstrate that player z too has an incentive to form a link with j, if n is large enough. Let n_1 be the set of players such that z is inessential for their access to j, in network g. The marginal payoff to player z from a link with j is then bounded below by:

$$n_1 \frac{1}{E(c)+2} + (n - n_1 - 1)\left[\frac{1}{3} - \frac{1}{E(i, j; g)+2}\right]. \tag{17}$$

The first term is a lower bound on payoffs that z can get from being essential to the n_1 players for linking with j, players for whom he was inessential in g. The second term is the additional payoffs he gets from being essential for paths from j to other players $n - n_1 - 1$. Since $E(i, j; g) \geq 2$, it follows that for n large these additional payoffs will exceed any given

c. We have thus shown that for two players i and j outside the cycle, $E(i, j; g) \geq 2$ cannot sustained.

The above step shows that all players who are not part of the cycle must link to the same common player. We next consider a network with a cycle which does not contain the single player identified in the previous step. Let player a be this central player outside the cycle. This allows for two possibilities: one, that a forms a link with player j in the cycle and two, that a forms a link with b who in turn forms a link with j in the cycle. We take up case two. Suppose that $j_1 \in C$ and $g_{j_1,j} = 1$. We first observe that player j_1 has a strict incentive to delete his link with j and instead form a link with a. The costs remain the same while the gross payoffs increase strictly since player j is no longer essential and no new essential players have been created. We next show that player a also has an incentive to form a link with j_1, for sufficiently large n. Let n_1 is the number of players in the cycle. Then the marginal payoff to player a from forming a link with player j_1 is bounded below by:

$$(n_1 - 1).(n - n_1 - 2)\left[\frac{1}{3} - \frac{1}{5}\right] \tag{18}$$

Given any c, it follows that marginal returns to player a from linking with j_1 are greater than the cost, for large enough n. Case one is similarly shown to be unsustainable in equilibrium. Thus the single player who links with all non-cycle players must belong to the cycle as well. This completes the proof of the lemma. ■

The final step in the proof shows that the hybrid cycle-star network is not an equilibrium for large n.

Lemma 3.6 *Suppose that* $c > 1/6$. *The star is the unique non-empty equilibrium, for large* n.

Proof: Suppose that g is non-empty equilibrium network and it has the hybrid cycle-star structure. Let x be the number of peripheral players, in the star, while $y = n - x$ is the number of players in the cycle. Note that $y \geq 4$; $y = 3$ cannot be sustained in equilibrium given that $c > 1/6$. Next note that given a $c > 1/6$ there is a $y(c)$ such that $y \leq y(c)$ in equilibrium. Let player i be the center of the star and let i maintain links with j and k in the cycle, and let l be a peripheral player. Note that player j always has an strict incentive to switch his link from player i to player l. For player, forming an additional link with player j is only worthwhile if $(y-1)/6 > c$. So given a c, there is a $y(c)$ such that for $y \geq y(c)$ players l and j will form a link. Thus we can suppose that $y \leq y(c)$ in what follows.

We fix some c and look at the class of hybrid networks g in which $y \leq y(c)$. We show that players j and k have a strict incentive to form a link if the set of peripheral players $x \geq n - y(c)$ is sufficiently large. The payoffs of players j and k in the hybrid network g^H are given by

$$\pi_j(g^H) = \pi_k(g^H) = \frac{x}{3} + \frac{y-1}{2} - 2c \tag{19}$$

Now consider a deviation by players j and k in which player k deletes his link with player i and player j deletes his link with player m in the cycle and instead players j and k form a link with each other. The resulting network is a minimal network g' in which there are x peripheral players and a line starting with player i which consists of y players. The payoffs of player j in g' are given by

$$\pi_j(g') = \frac{x}{3} + \frac{1}{2} + \sum_{k=2}^{y-1} \frac{1}{k} + x \sum_{k=4}^{y+1} \frac{1}{k} + \sum_{k=3}^{y} \frac{1}{k} - 2c \tag{20}$$

These payoffs are bounded below by

$$\frac{x}{3} + x\frac{y-2}{y+1} - 2c = M \tag{21}$$

Next note that $M > x/3 + (y-1)/2 - 2c$ if

$$x \; > \; \frac{y-1}{2}\frac{y+1}{y-2} \tag{22}$$

Since $y \geq 4$, the right hand side is increasing in y and bounded above by $[y(c) - 1][y(c) + 1]/2[y(c) - 2]$. The final step is to note that $x + 1 = n - y + 1 \geq n - y(c) + 1$ and so is larger than this term for sufficiently large n. Thus player j has a strict incentive to switch links to player k for large n. We now turn to the incentives of player k.

The payoffs of player k in g' are given by

$$\pi_k(g') = \frac{x}{4} + \frac{1}{3} + \frac{1}{2} + \sum_{k=2}^{y-2} \frac{1}{k} + x \sum_{k=5}^{y+1} \frac{1}{k} + \sum_{k=4}^{y} \frac{1}{k} - 2c \tag{23}$$

These payoffs are bounded below by

$$\frac{x}{4} + x\frac{y-3}{y+1} - 2c = M' \tag{24}$$

$M' > x/3 + (y-1)/2 - 2c$ if

$$x \ > \ \frac{6(y-1)(y+1)}{11y - 37} \tag{25}$$

Since $y \geq 4$, the right hand side is positive and increasing in y and so is bounded above by $6[y(c) - 1][y(c) + 1]/[11y(c) - 37]$. Note that $x = n - y \geq n - y(c)$ is larger than this term for sufficiently large n. Thus player k has a strict incentive to switch links to player j, for sufficiently large n. The proof is complete. ■

4 Extension: Two-person coalition proof networks

In our analysis so far we have assumed that a deviation by two players is credible so long as it yields higher payoffs to both players. We have not looked at profitable deviations from the agreed upon deviation. In this section we will examine the implications of this further requirement. We start with a definition of coalition proof-ness in our context.

Definition 4.1 *A strategy profile* $\mathbf{s}^*$ *is a bilateral coalition proof equilibrium if the following conditions hold:*

1. *For every* $i \in N$, $\Pi_i(s^*) \geq \Pi_i(s_i, s^*_{-i})$, *for all* $s_i \in S_i$.

2. *For any pair of players,* $i, j \in N$ *at least one of the following two conditions hold:*

   - *Either there is no strategy pair* (s_i, s_j) *which satisfies:*

     $$\Pi_i(s_i, s_j, \mathbf{s}^*_{-\mathbf{i}-\mathbf{j}}) > \Pi_i(s^*_i, s^*_j, \mathbf{s}^*_{-\mathbf{i}-\mathbf{j}})$$

     and

     $$\Pi_j(s_i, s_j, \mathbf{s}^*_{-\mathbf{i}-\mathbf{j}}) > \Pi_j(s^*_j, s^*_j, \mathbf{s}^*_{-\mathbf{i}-\mathbf{j}}).$$

   - *OR for every such strategy pair* (s_i, s_j) *that satisfies above inequalities there exists a profitable individual deviation for either* i *or* j *from* (s_i, s_j, s^*_{-i-j}).

We shall use the short form – pcp-equilibrium – to refer to pair-wise coalition-proof equilibrium.

Proposition 4.1 *(i) If* $1/6 < c < 1/2 + (n-2)/6$ *then the star is a pcp-equilibrium (ii) Given any* $c > 0$, *there exists an* $n(c)$ *such that for all* $n \geq n(c)$, *a cycle containing all players is a pcp-equilibria.*[3]

Proof: The star is an equilibrium from Proposition 3.3; so it is clearly also a pcp-equilibrium. We prove that a cycle is a pcp-equilibrium. All players are symmetrically located in a cycle so we need only consider the incentives for a typical player, (say) player i. First note that the payoffs of player i in a cycle are $(n-1)/2 - 2c$. The payoffs from deleting both links are 0 and clearly for a given c, there is always a large enough n such that deleting both links is not optimal. Consider next the deviation of deleting one link. If player i deletes one link and keeps the other link as in the cycle then he becomes an end-player of the line network. In this network his payoffs is given by

$$\sum_{k=2}^{n} \frac{1}{k} - c \tag{26}$$

---

[3] The stability of the cycle which is not minimal is an instance of the inefficiency discussed in Jackson (2003) where players will over-connect to avoid positional disadvantages.

The payoff from both links in cycle are higher if

$$\frac{n-1}{2} - \sum_{k=2}^{n} \frac{1}{k} > c \tag{27}$$

Clearly this inequality holds for large enough n. We next consider the deviation in which player i deletes both links but forms a new link (say) with the center of the line network that arises. Again, a variation of the above argument shows that for large n this reduces payoffs.

We finally consider deviations by player i in which player i maintains one of the links as is but changes the partner of the other link. Suppose that in the cycle he has a link with $i-1$ and $i+1$. In the deviation, he maintains the link with $i-1$ but deletes the link with $i+1$ and instead forms a link with some player k. If player k retains all his links as in the cycle, it is easy to see that the payoffs of the player go down strictly since the costs remain the same (he maintains two links) while the gross payoffs decline since player k is essential for accessing at least one player, namely $i+1$. So this deviation is not profitable. If player k deletes both his links then the payoffs from the deviation are still lower and so it clearly not profitable. We turn to the final case, where players i and k coordinate and link up with each other but delete a link each so that the new network is a line. From Lemma 3.4 it follows that players such a deviation is profitable. However, now we need to check whether this deviation is credible: do the players have an incentive to actually delete one of their links? We show that for large n, at least on the players has a strict incentive to retain their links in the cycle. Number the players in the line as $1, 2...., n$, from left to right. So there are $i-1$ players to the left of player i while there are $n-k$ players to the right of player k in the line. Suppose without loss of generality that $i-1 \geq n-k$. Player k gets a payoff $\sum_{l=2}^{i+1} 1/l$ in the line network from these players $1, 2..., i$. The payoffs can be increased to $i/2$ if player k forms a link with player 1. Clearly, it is profitable for player k to deviate from the deviation and form a link with player 1, if n is large. It is similarly possible to show that player 1 would have an incentive to form a link with player k, if n is large. Thus the deviation in which i and k deviate to create a line is not two-person coalition proof. Finally, we note that starting from a cycle it is not profitable for player i to form any additional links. The proof for cycle being a pcp-equilibrium is complete. ■

5 Conclusion

This paper has studied a simple model of network formation where agents may exploit positional advantages if these provide them with the ability to block profitable bilateral interaction between two players who are not direct neighbors. We show that the strategic struggle for these advantages leads to a polarized star architecture where a single player becomes essential to connect every other pair of players. This represents a clearcut formalization of the notion found in the sociological literature that structural holes opens the potential for large benefits to those individuals who succeed in bridging them.

References

[1] Aumann, R. and R. Myerson (1988): "Endogenus formation of links between players and coalitions: an application to the Shapley Value" in *The Shapley Value*, A. Roth (ed.), 175-91, Cambridge University Press.

[2] Bala, V. and S. Goyal (1999): "A Non-Cooperative Model of Network Formation", *Econometrica* **68**, 1181-1230.

[3] Burt, R.S. (1982): *Structural Holes*, New York: Academic Press.

[4] Currarini, S. and M. Morelli (2000): "Network Formation with Sequential Demands", *Review of Economic Design* **5**, 229-250.

[5] Dutta, B., A. van Den Nouweland and S. Tijs (1995): "Link formation in cooperative situations", *International Journal of Game Theory* **27**, 245-256.

[6] Goyal, S., M. van der Leij, and J.L. Moraga (2003): "Economics: an emerging small world?,*Tinbergen Institute Discussion Paper 2004-1.*

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

[8] Greif, A. (1994): "On the Political Foundations of the Late Medieval Commercial Revolution: Genoa during the Twelfth and Thirteenth Centuries", *The Journal of Economic History* **54**, 271-287.

[9] Gulati, R., N. Nohria, and A. Zaheer (2000): "Guest editor's introduction to the special issue: strategic networks", *Strategic Management Journal* **21**, 199-201.

[10] Jackson, M. O. (2003), The stability and efficiency of economic and social networks, on *Networks and Groups: models of strategic formation*, eds. B. Dutta and M. O. Jackson. Springer Verlag. Berlin.

[11] Jackson, M. O. and A. Wolinsky (1996): "A Strategic model of economic and social networks", *Journal of Economic Theory* **71**, 44-74.

[12] Kranton, R. & D. Minehart (2001): "A theory of buyer-seller networks", *American Economic Review* **91**, 485-508.

[13] M. E. J. Newman, M. E. J. (2001): "The structure of scientific collaboration networks," *Proc. Natl. Acad. Sci. USA* **98**, 409-415.

[14] Rauch, J. E. (2001): "Business and Social Networks in International Trade" *Journal of Economic Literature* **39**, 1177-1203.

[15] Valente, T. (1996): "Social Network Thresholds in the Diffusion of Inovations", *Social Networks* **18**, 69-89.