w o r k i n g

p a p e r

20 01

Making Friends Meet: Network Formation with Introductions

Jan-Peter Siedlarek

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working Paper 20-01 January 2020

Making Friends Meet: Network Formation with Introductions

Jan-Peter Siedlarek

High levels of clustering—the tendency for two nodes in a network to share a neighbor—are ubiquitous in economic and social networks across different applications. In addition, many real-world networks show high payoffs for nodes that connect otherwise separate network regions, representing rewards for filling "structural holes" in the sense of Burt (1992) and keeping distances in networks short. This paper proposes a parsimonious model of network formation with introductions and intermediation rents that can explain both these features. Introductions make it cheaper to create connections that share a common node. They are subject to a tradeoff between gains from shorter connections with lower search cost and losses from lower intermediation rents for the central node. Stable networks are shown to have high levels of clustering at the same time that they permit substantial intermediation rents for nodes bridging structural holes.

JEL Classification: A14, D85.

Keywords: networks, network formation, clustering, intermediation, introductions.

Suggested citation: Siedlarek, Jan-Peter. 2020. "Making Friends Meet: Network Formation with Introductions." Federal Reserve Bank of Cleveland, Working Paper No. 20-01. https://doi.org/10.26509/frbc-wp-202001.

Jan-Peter Siedlarek is at the Federal Reserve Bank of Cleveland (jan-peter.siedlarek@clev.frb.org). He is grateful to Fernando Vega-Redondo for extensive discussions and advice and thanks Ben Craig, Peter Vida, and seminar participants at the European University Institute, University of Mannheim, PET2013 Lisbon and the EEA2013 Göteborg for many helpful comments and suggestions. This research was in part supported by funding from the DAAD (German Academic Exchange Service) and the DFG through grant SFB-TR15.

1 Introduction

The structure of relationships among economic actors performs an important role in a variety of social and economic contexts, influencing opportunities to cooperate, to trade, and to interact. The size and structure of such networks have been shown to influence phenomena as diverse as criminal activity, finding a job, being promoted within a company, the success of company R&D, and others.[1] Two prominent features that are frequently observed in real-world networks are high levels of clustering and returns for agents bringing otherwise distant parts of the network together. The first – high clustering – describes a tendency for pairs of nodes to share common neighbors. In many real-world networks that have been studied empirically, clustering coefficients are found to exceed levels expected from random link-formation processes. For example, studies of co-authorship networks, which record collaborations between researchers, exhibit a very high probability that a given pair of cooperating scientists also both cooperate with a common third researcher (see Newman (2003), Grossman (2002), and Goyal, van der Leij, and Moraga-González (2006)). Similar features are observed for the cooperation network of actors in movies (Watts, 2004) and the linking patterns of web sites (Adamic, 1999).

The second salient feature of real-world networks is returns to nodes facilitating the connection of others. Such connectors in many cases act as "bridging" agents that connect otherwise distant parts of the network and in the process take advantage of "structural holes" in the sense of Burt (1992), who coined the term studying the phenomenon in the context of managerial networks in companies. Such bridges earn rents on account of the position they inherit and create networks that have a small diameter and allow a high degree of interaction across the network. As shown by Watts and Strogatz (1998) such short distances can be caused by a small number of connections in the network that link otherwise distant parts of the network, resulting in "small worlds."[2]

In this paper I propose a parsimonious network-formation model to explain how networks with both high levels of clustering and bridging nodes with intermediation returns leading to short distances can arise. The model adapts a connections model with intermediation rents and starts from the general observation that just as the network structure can be important for a number of social and economic processes that occur on the network, it can also influence the conditions under which new links are formed in the network itself. Specifically, I propose to study the impact of an introduction mechanism that facilitates the formation of links between two nodes if they share a common neighbor in the existing network. Such link creation between nodes sharing a common neighbor is known as *triadic closure* in the literature, and it has been

[1] See Jackson (2008) for an overview of network theory in economics and applications as well as the more recent survey in Jackson, Rogers, and Zenou (2017).

[2] See the early article by Milgram (1967) describing an experiment in which letters crossed the United States through social networks via chains of first-name acquaintances.

observed to be of relevance in various applications: For example, in Mayer and Puller (2008) an existing common neighbor is a strong predictor of new links in the social networks of a university. Similar processes in which the three parties are involved in link creation can be seen in the formation of business relationships between firms and other organizations. For example, Uzzi (1996), in his study of the apparel industry in New York, describes how new relationships between two parties often result from referrals by common business partners:

> *In the firms I studied, third-party referral networks were often cited as sources of embeddedness. [...] One actor with an embedded tie to each of two unconnected actors acts as their go-between by using her common link to establish trustworthiness between them* (Uzzi, 1996, p. 679)

In the paper I study a variant of the connections model with intermediation rents, derived from Jackson and Wolinsky (1996) and Goyal, van der Leij, and Moraga-González (2006) to which I add a specific form of triadic closure, which I label "introduction": Two nodes that share a common neighbor can be invited to form a link by that neighbor. If all three parties agree, the new link is formed. Such a mediated connection has a lower cost than if the two parties were to seek each other out without assistance. However, by creating a direct connection it can cut off intermediary agents and thus threaten intermediation rents. The introduction mechanism thus combines a tendency for triadic closure with an explicit consideration of the incentives that arise from intermediation, generating a distinct trade-off for the introducing nodes. On the one hand, an introduction is an efficient way of creating connections as it reduces the costs involved in non-intermediated link formation. On the other hand, links created by introductions affect the distribution of payoffs and can expose the introducing player to circumvention, threatening intermediation payoffs received from being essential. The analysis of this trade-off and how it contributes to both high clustering and short distances with network bridges forms the core of the paper.

I first show that in this model efficient networks are either empty, star, or complete networks, reflecting the familiar patterns from the baseline connections model of Jackson and Wolinsky (1996). I then focus on decentralized incentives for link formation by studying the set of networks that are stable to deviations from link creation and destruction. I show that the set of efficient network configurations is not necessarily stable for a given parameter configuration, reflecting externalities in link creation and destruction. I then derive a lower bound on the clustering coefficient for a given node in stable networks, showing how it depends on the returns to intermediation as well as the cost advantage of introductions. Finally, I show that there is a limit on the distance between any pair of nodes given their degree and stable networks tend to be connected.

The contributions of my paper relate to the study of the introduction mechanism and the compensation of intermediate nodes. While intuitive, my approach to studying link creation as a process conditional on the network in place at a given time is relatively novel in the network-formation literature, in particular in the branch that studies strategic network formation.[3] The contributions closest to the present analysis are Jackson and Rogers (2007) and Jackson and Rogers (2005). In Jackson and Rogers (2007) the authors study a network-formation setup with a growing network in which new nodes first connect to a set of randomly chosen existing nodes and in a second process may connect to neighbors of those, i.e., connect to friends of friends. Connections created by the second process are similar to those formed in an introduction as studied here in that they close open triangles. Partly through that triadic closure mechanism, Jackson and Rogers (2007) are able to match a mean-field approximation of their model to the network properties of a number of applied settings, including high clustering coefficients. The present paper is complementary to Jackson and Rogers (2007) and deviates from it in at least two main ways. First, Jackson and Rogers (2007) model link creation via random meeting opportunities, which result in links if the pair of nodes to be linked finds it beneficial to connect. This ignores the strategic aspects of the key role that intermediary players perform in making the kind of introductions that facilitate the friends-of-friends meetings. In contrast, in this paper, I explicitly study intermediary incentives. Second, the underlying payoff model in Jackson and Rogers (2007) focuses on the returns of each connecting pair independent of implications for returns from other parts of the network.[4] My model puts such externalities at the center of the analysis, studying how incentives for intermediaries to protect their returns shape the resulting network. Beyond Jackson and Rogers (2007), an earlier contribution by Vázquez (2003) presents a mean-field analysis of graph dynamics with a nearest friend process. However, that contribution ignores the strategic considerations of the players involved and focuses instead on how local rules may motivate the preferential attachment hypothesis. More recently, local structures of the network such as triangles and the processes generating them have been exploited in empirical work on network data. This includes, for example, Chandrasekhar and Jackson (2018), who show how local structures such as triangles and other subgraphs can be used to study network-formation processes empirically. Recent surveys on this active work stream can be found in Graham (2019) as well as de Paula (2019).

The second closely related paper is Jackson and Rogers (2005). In their paper the authors offer a strategic network formation model that can predict small world properties. The key ingredients in their analysis are the benefits from indirect links as in the connections model of Jackson and Wolinsky (1996) and an "islands"

[3] The literature on network formation can usefully be grouped into two main categories: (i) random network formation and (ii) strategic network formation. Whereas the first approach analyzes the outcome of an exogenous stochastic process of link creation and aims to explain the observed features of real-world networks, the second explicitly studies the incentives of nodes to form links among themselves. Network formation is then the outcome of individual payoff maximization by nodes. The present paper falls into the second category, although I will in places refer to work from the random-network-formation tradition. I focus on the works most closely related to the model I study here. The reader is referred to the comprehensive overview provided in Jackson (2008).

[4] In extensions in Section V, Jackson and Rogers (2007) discuss the implications of different payoff specifications and formally show the effect of making payoffs dependent on degree. However, they do not offer a formal treatment of the externalities resulting from intermediation returns.

cost model, in which links connecting players on the same island are cheaper than links between islands. My paper similarly aims to explain both clustering and short distances but in addition analyzes the effect of intermediation benefits and the resulting incentives for agents to adopt and protect bridging positions. In addition, while in Jackson and Rogers (2005) the grouping of players into easily connected islands is given *exogenously*, in my paper the cost advantages arise *endogenously* through the introduction mechanism. This endogenous, network-based cost structure is one of the key contributions of the present paper over Jackson and Rogers (2005).

In addition, this paper offers a contribution to the literature on incentives for intermediary nodes and rents from intermediation. The potential benefits accruing to nodes in specific positions crucial to the network are discussed extensively by Burt (1992) in the context of organizations. The author investigates the rents available to individuals who bridge "structural holes" and the dynamics of jockeying for the positions required to access these rents. In the economics literature, a model that discusses issues most closely related to this paper is provided by Goyal and Vega-Redondo (2007). There, the authors consider network formation in the presence of intermediation benefits and analyze the interplay of three motivations: (i) access to the network, (ii) benefits from intermediation, and (iii) avoidance of sharing benefits with intermediaries. They find that in the absence of capacity constraints, a star emerges, in which a single node acts as intermediary for all transactions, receiving significant intermediation rents. Contrary to the present paper, in their model, Goyal and Vega-Redondo (2007) focus on direct link creation and do not implement an introduction mechanism. A similar model is also explored in Kleinberg et al. (2008), which again does not allow for introductions.

Finally, my paper is related to the literature on strategic network formation with transfers. The introduction mechanism I study requires a mechanism to facilitate compensation of the intermediating node. As a result, transfer payments and their ability to overcome externalities play a key role. The seminal contribution in this area is Currarini and Morelli (2000), who study a sequential setup with transfers and find that with relatively few restrictions on payoffs, efficient networks are formed in equilibrium. A more general analysis of various forms of transfers and their ability to implement efficiency in simultaneous network formation is found in Bloch and Jackson (2007). Neither paper studies the introduction process at the center of the present paper. My paper contributes to the literature by extending the standard game-theoretic tool of pairwise-stability (Jackson and Wolinsky, 1996) to a setting with more than two players.

In summary, this paper studies strategic network formation with introductions. The interplay of the benefits from intermediation and the cost advantages of introductions is shown to explain the coexistence of high clustering, intermediation rents, and short distances, thereby offering a simple joint explanation for three salient features in real-world networks. The remainder of the paper is structured as follows.

Section 2 presents the main model, including network notions and payoff structures. Using this model, Section 3 presents the introduction mechanism and characterizes efficient networks. Section 4 studies the characteristics of networks that are stable to deviations, including introductions, and contains the key results regarding clustering, intermediation benefits, and short distances. Finally, Section 5 concludes.

2 A Model of Network Formation with Intermediation and Introductions

I study a model of network formation with introductions and intermediation rents. It builds on the symmetric connections model of Jackson and Wolinsky (1996) in that nodes create payoffs from direct and indirect connections and there is a cost to maintaining links. I then introduce into this setting rents for link intermediation as well as rules for link creation and destruction. I allow for both direct link creation by pairs and introductions by sets of three nodes, thus providing an additional way for nodes to create connections, beyond the bilateral link creation and unilateral link destruction in Jackson and Wolinsky (1996).

2.1 A Connections Model with Intermediation Rents

To start, we set out some network notation. There is a finite set of players $N = \{1, 2, \ldots, n\}$ with $n > 3$. Players represent nodes that are linked in a network g. g is a set of links, that is, pairs of players that are connected to one another, with typical element ij representing a link between players i and j. Let g^N be the set of all pairs of players in N, describing the complete network. g^0 is the empty set and describes the empty network.

Payoffs are generated by every connection between a pair of players. Direct connections generate a normalized surplus of 1, indirect connections of length two (that is, with a single intermediate node) generate surplus of $\delta \leq 1$. Following Kleinberg et al. (2008) indirect connections of length greater than two generate zero surplus. This is consistent with empirical research on the benefits from connections in organizations showing that connections with more than one intermediate node provide almost no benefits (Burt, 2007, 1992).

The surplus generated by direct connections without an intermediary is split equally between the two nodes involved, that is, each receives a payoff of $1/2$. In addition, link maintenance costs are c for each direct link, again, split equally between the two nodes involved.

When indirect connections are facilitated by an intermediary, that node can capture a share of the surplus created by the indirect connection. These rents can exist, for example, in networked markets with intermediaries competing for opportunities to facilitate trade. They can also be part of the "structural holes" argument of Burt (1992), in which network nodes jockey for positions of "brokerage." The value of intermediation rents depends on the extent to which the intermediary is indispensable for the connection to work: Where many alternatives exist, rents available to each individual intermediary will be lower. Formally, total rents for intermediation between i and j are denoted $\gamma(r_{ij})$, which is a decreasing function of r_{ij}, the number of intermediaries, that is, nodes that are connected to both i and j.

In this model I follow Goyal and Vega-Redondo (2007) and assume that intermediary nodes capture a positive share of the total surplus if and only if they are *essential* to the connection, that is, if they are on every path connecting two nodes and thus $r_{ij} = 1$. This assumption is consistent with intermediation rents being competed away in the spirit of Bertrand competition when there is more than one intermediary. It also serves to bring into focus the key strategic concern of intermediaries to avoid being circumvented, which is of primary interest in this paper.[5]

In summary, in the connections model with intermediation rents, total payoffs for player i from network g are given by the following function.

$$\pi_i(g) = d_i(g)\frac{1-c}{2} + \sum_{\substack{j\neq i:\\ \ell_{ij}=2}} \delta\frac{1 - I_{\{r_{ij}=1\}}\gamma}{2} + \sum_{\substack{j,k\neq i:\\ \ell_{jk}=2\wedge ij\in g\wedge ik\in g}} I_{\{r_{jk}=1\}}\delta\gamma \tag{1}$$

where $d_i(g)$ denotes the degree of player i in network g. Payoffs are the sum of the following components: (i) benefits from direct connections net of link maintenance costs, (ii) benefits from indirect connections net of intermediation rents paid, and (iii) intermediation rents received.

2.2 Network Formation with Introductions

This section describes the process for creating and deleting links. There are two ways of creating links in this model. In addition to bilateral link creation, that is, nodes i and j agreeing to form a link between themselves, the model allows for link creation by introduction, that is, via a third-party intermediary. Links can also be destroyed unilaterally.

Bilateral link creation allows any unconnected pair i, j to create a new link between themselves if both agree. Link creation incurs a fixed cost $f/2 \geq 0$ for each node or f in total. This captures, for example, investments in efforts to engage with a counter-party and develop a sufficient level of trust for a functional relationship. As a result, the cost is sunk and not recovered when the connection is destroyed.

[5] The assumption is consistent with a model of bargaining without replacement in the limit when bargaining frictions disappear (Siedlarek, 2015). See Kleinberg et al. (2008) for a different approach in which intermediation benefits decay with the number of alternative paths in a gradual way.

In addition to bilateral link creation, I allow for intermediated link creation via introductions, that is, creation of a new connection between two unconnected nodes i, j that share a common neighbor k. This also creates the link ij but requires the agreement of the introducing node k. Introductions have a cost advantage of bilateral link creation as they take advantage of an existing indirect connection and incur a sunk cost of $(1-\epsilon)f$, with $\epsilon \in [0,1]$. This creates a trade-off for the use of introductions: They are cheaper than regular bilateral link creation but require the agreement of the introducing node.

Any link ij can be unilaterally destroyed by either node i or node j. The destruction of a link does not recover any search cost incurred in creating it, consistent with the notion that such costs are sunk. To reflect the sunk cost nature of link creation costs, I do not show them in the static payoff function $\pi_i(g)$ in equation 1 above. Instead they come into play when considering the implications of a transition between different networks as part of an analysis of efficiency (Section 3) and stability (Section 4). I begin the analysis of the model with a characterization of efficient networks.

3 Efficient Networks with Introductions

Efficiency considerations in the model apply to both the process of network formation and the resulting structure of relationships. We thus define efficiency on the basis of both.

Definition 1. *A network structure g is efficient if it maximizes the sum of total payoffs across nodes net of the minimum cost incurred in creating it.*

We start by considering the cost of link creation. Generally speaking, introductions are more efficient than link creation without introductions due to their discount ϵ on fixed cost f. However, the requirement for nodes to share a common neighbor before introductions are feasible implies some role for non-mediated, bilateral link creation. Proposition 2 then shows the interaction of the net benefits from connections and the costs from link creation. The structure of the proof follows Jackson and Wolinsky (1996) and is provided together with all other proofs in the Appendix.

Note that different from Jackson and Wolinsky (1996), the result requires a characterization not just the final network structure but also of the most efficient way of creating the network through a combination of search and introductions.

Proposition 2. *The unique efficient network structure in the model with introductions is:*

1. *The* empty network *if*

$$1 - c < f - \frac{n-2}{2}\delta, \textit{ and}$$
$$1 - c < \left[1 - \frac{n-2}{n}\epsilon\right] f$$

2. *The* star network *if*

$$1 - c > f - \frac{n-2}{2}\delta, \textit{ and}$$
$$1 - c < \delta + (1 - \epsilon) f$$

   The star network includes all n players connected via $n-1$ links created by bilateral link creation.

3. *The* complete network *if*

$$1 - c > \left[1 - \frac{n-2}{n}\epsilon\right] f, \textit{ and}$$
$$1 - c > \delta + (1 - \epsilon) f$$

   The complete network consists of $n-1$ links created by bilateral link creation with all remaining links created through introductions.

The efficient network structure is unique up to a permutation of players.

Here, as in subsequent analysis, the characterization focuses on the term $1 - c$, which captures the net benefits of a direct connection, that is, total benefits less link-maintenance costs for a direct connection. Figure 1 illustrates the parameter regions characterized in Proposition 2 for the case $\epsilon = 1$, that is, when introductions have zero fixed cost relative to bilateral link formation. As link formation costs increase relative to maintenance costs c, the efficient network is either complete or empty: Once link formation is productive at all, it pays to make maximum use of introductions and form the complete network.

At $f = 0$ the efficient structures and parameter ranges correspond to those in Jackson and Wolinsky (1996). However, as fixed costs f increase, differences between Jackson and Wolinsky (1996) and the setting with fixed costs emerge. The parameter range in which the empty network is efficient grows, reflecting the additional costs to be incurred for a given benefit to be obtained. As f increases from zero, at first it only affects the range in which the star network is efficient. Once f increases above $\frac{n}{2}\delta$, the star is never efficient and the parameter range for which the complete network is efficient begins to shrink as well. The result emphasizes the extent to which the cost advantage of introductions pushes efficient structures toward

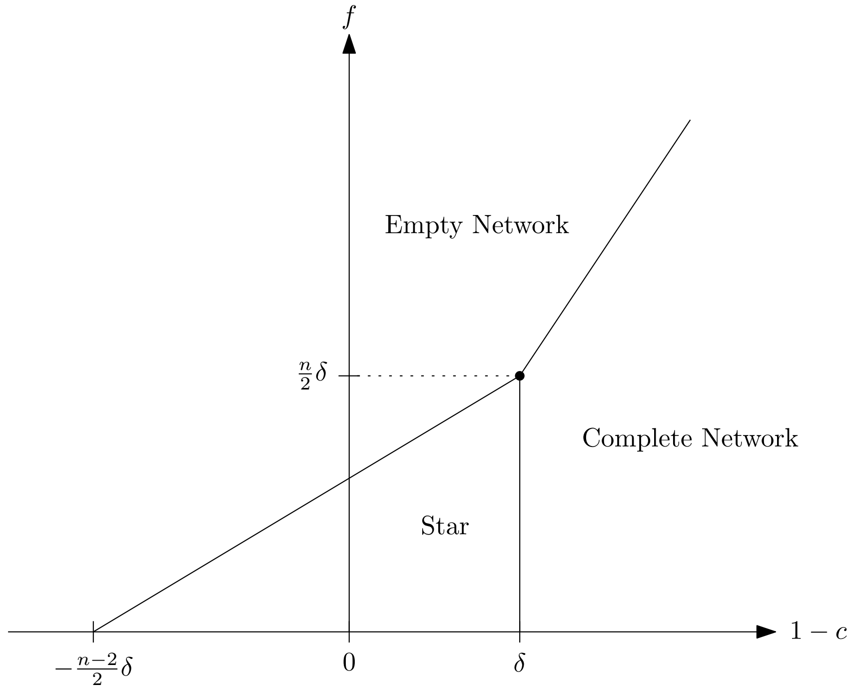


Figure 1: Efficient Network Configurations ($\epsilon = 1$)

high density and high levels of clustering. The efficient configuration is independent of the intermediation rent parameter γ as it presents a transfer between agents without efficiency implications in the sense of Definition 1.

The results given here provide a benchmark for subsequent analysis, which studies the incentives for players to create and remove link from the network.

4 Pairwise Stable Networks with Introductions

This section presents an analysis of networks that are stable when introductions are available. I apply a myopic stability concept based on Jackson and Wolinsky (1996) that incorporates a suitable extension of pairwise stability. This section analyzes the strategic incentives for network formation in the connections model with introductions using a version of pairwise stability (Jackson and Wolinsky, 1996) suitably extended to allow for transfers and the dynamic introduction process that is the subject of this paper.

An analysis of introductions requires the inclusion of transfers to the introducing player in order to compensate that player for potentially lost intermediation benefits. For illustration, consider the payoff implications of the new link that is created. The new link shortens the distance between the two players being introduced to one step, yielding the resulting benefits; however, the new link does not generate additional benefits for the introducing player as he is already connected to both. Indeed, the new link results in the introducing player no longer being essential for the connection between the players introduced, and thus he will lose out. Introductions by themselves are at best payoff neutral for the introducing player, and transfer payments are necessary to make any introduction profitable for the introducer. As a result, the stability concept I use here is one that considers deviations with transfers as proposed in Bloch and Jackson (2006) and is extended to accommodate the introduction process.

Definition 3 (Myopic stability under introductions)**.** *A network g is myopically stable with introductions if:*

a. (Destruction) $\forall\ ij \in g$,

$$\pi_i(g) + \pi_j(g) \geq \pi_i(g - ij) + \pi_j(g - ij)$$

b. (Bilateral Link Formation) $\forall\ ij \notin g$,

$$\pi_i(g) + \pi_j(g) \geq \pi_i(g + ij) + \pi_j(g + ij) - f$$

c. (Introduction) $\forall\ \left\{i, j, k : ik \in g \wedge jk \in g \wedge ij \notin g\right\}$,

$$\sum_{v \in \{i,j,k\}} \pi_v(g) \geq \sum_{v \in \{i,j,k\}} \pi_v(g + ij) - (1 - \epsilon) f$$

The first two conditions correspond to those used for networks that are *pairwise stable with transfers*, as analyzed in Bloch and Jackson (2006), adapted for the introductions model. In particular, the conditions introduce the fixed cost f, which is sunk in the sense that it is incurred when bilateral links are created and not recovered by players if a link is destroyed. In addition, the definition includes a third condition, which requires that in a stable network there be no opportunities for profitable introductions, that is, a triplet of nodes that form an open triangle jointly benefiting from adding the missing link, with a lower fixed cost $(1 - \epsilon)f$.

All three conditions allow for transfers between the players involved by considering the *sum of payoffs* rather than individual payoffs. This also applies to link destruction in order to maintain symmetry between link creation and link destruction.[6] Allowing for transfers in link destruction reduces the set of profitable deviations of this type: All link removals that are jointly profitable necessarily involve at least one player for whom it is unilaterally profitable; however, if one player loses out from the removal of the link, Definition 3 requires that the damage done to the other side involved in the link not be too high. In this sense, the

[6] See footnote 5 in Bloch and Jackson (2006) for a discussion.

solution concept with transfers is weaker than that without transfers as far as link destruction is concerned.[7]

In the following I consider (i) the stability properties of the efficient configurations (Section 4.1) and (ii) characterize the properties of stable networks in general (Sections 4.2 and 4.3).

4.1 Stability of Efficient Networks

Here I derive the stability properties of the configurations that are possible efficient arrangements for some parameter ranges. The analysis starts with one of the possible efficient configurations – the empty network, the star network, and the complete network – and characterizes the parameter restrictions necessary for each configuration to be stable. The details of the derivation of the conditions have been relegated to the Appendix, Section A.2.

Proposition 4. *The three efficient network structures are myopically stable with introductions if the following conditions hold:*

1. *The* empty network *is stable if*

$$1 - c \leq f$$

2. *The* star network *is stable if*

$$1 - c \geq -\frac{1+\gamma}{2}(n-2)\,\delta, \textit{ and}$$
$$1 - c \leq \min\left\{f + (1-\gamma)\delta, \delta + (1-\epsilon)f\right\}$$

3. *The* complete network *is stable if*

$$1 - c \geq \delta$$

The parameter thresholds are illustrated in Figure 2 for the case $\epsilon = 1$. Table 1 lists the conditions next to the corresponding thresholds for efficiency, which are derived in Section 3. In some areas, efficiency and stability are aligned. For example, whenever a complete network is efficient, it is also stable. However, in other areas stability and efficiency diverge reflecting (i) the presence of externalities in payoffs as well as (ii) the sunk nature of link formation costs. The first of these is familiar from other papers on strategic network formation such as Jackson and Wolinsky (1996): When considering whether to create or destroy a link,

[7] Note also that Definition 3 includes transfers implicitly by comparing sums of payoffs rather than specifying explicitly the amounts exchanged between players. The implicit approach is more concise and sufficient for the myopic case.

the players involved assess the impact only on their own payoffs and disregard the effect on other players. The second effect is novel to the present paper and a key difference to the setting studied in Jackson and Wolinsky (1996). There, players can recover all costs from linking by removing a connection. In the setting with introductions studied here, link formation costs are sunk and cannot be recovered. Thus, in assessing the myopic stability of a given network, a link can be unprofitable to remove without recovery of fixed costs, even though it might be profitable to do so if the costs could be recovered. This discrepancy opens up an additional wedge between stability and efficiency that is not present in the model of Jackson and Wolinsky (1996).

The shaded area in Figure 2 represents the parameter space in which none of the three configurations is stable for any $n > 3$. The complete network is not robust to deviations in which a pair of players deletes a link, because they prefer an indirect connection intermediated by at least two competing middlemen, who incur no intermediation rents. The star network is not stable either: Here, a deviation by two peripheral players creating a new direct connection via bilateral link formation would be profitable, as it not only gives them the benefit of a direct connection but also allows them to circumvent the intermediation rents earned by the center node in the star network.

The derived conditions also imply that for certain parameter ranges, multiple configurations can be myopically stable. For example, if the surplus decays slowly such that $1 - c \leq \delta$ and in addition bilateral link formation costs are sufficiently high ($f \geq 1 - c$), then both the empty and the star network are stable. As shown in Section 3, only one of these would be efficient at the same time. The multiplicity of stable networks is a generic feature of the model and a function of both the myopic stability concept and the sunk cost feature of the search cost f.

Network	Efficient	Myopically Stable
Empty	$1 - c < f - \frac{n-2}{2}\delta$, and $1 - c < \left[1 - \frac{n-2}{n}\epsilon\right] f$	$1 - c \leq f$
Star	$1 - c > f - \frac{n-2}{2}\delta$, and $1 - c < \delta + (1 - \epsilon)f$	$1 - c \geq -\frac{1+\gamma}{2}(n-2)\delta$, $1 - c \leq \delta + (1 - \epsilon)f$, and $1 - c \leq f + (1 - \gamma)\delta$
Complete	$1 - c > \left[1 - \frac{n-2}{n}\epsilon\right] f$, and $1 - c > \delta + (1 - \epsilon)f$	$1 - c \geq \delta$

Table 1: Parameter Ranges for Efficient and Myopically Stable Networks

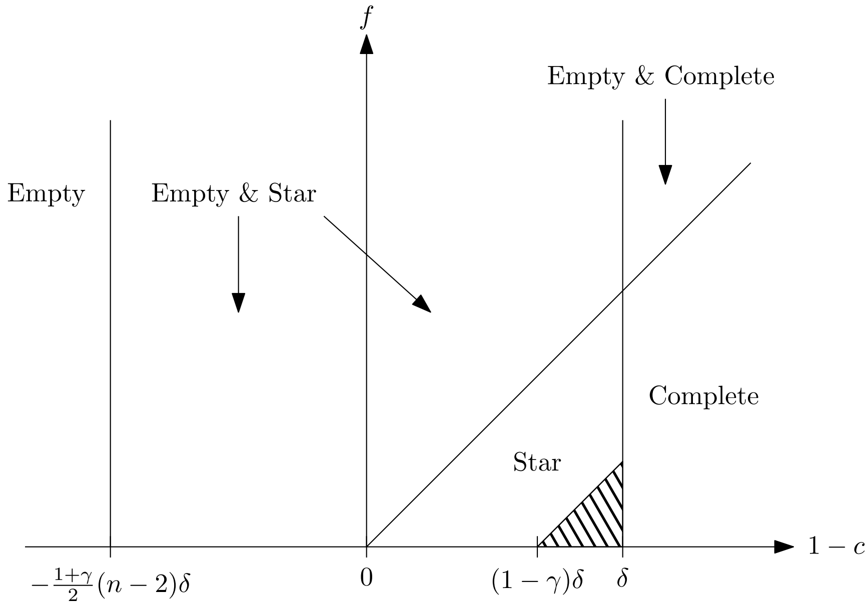


Figure 2: Myopic Stability of Empty, Star, and Complete Networks ($\epsilon = 1$)

4.2 Stable Networks Exhibit Minimum Level of Clustering

In this section I consider the impact of introductions on the clustering properties and the connectedness of myopically stable networks. First, I ask which possible introductions will remain unused in stable networks even if introductions are efficient. That is, I focus on the case in which $1 - c - (1 - \epsilon)f > \delta$ and direct connections generate a surplus larger than indirect ones, after accounting for the costs of introduction.

First, Proposition 5 shows that the payoff from an introduction of i and j by k depends on the local network environment only through nodes that are neighbors of either i or j but not both.

Proposition 5. *Consider a network g such that there is an opportunity for k to introduce i and j. The change in total payoffs to i, j, k from creating the introduction in the sense of Definition 3 is*

$$\begin{aligned}\Delta\pi_{ijk} =& 1 - c - (1-\epsilon)f - \delta \\ &+ \delta \sum_{\substack{u\neq v\neq w:\\ v,w\in\{i,j\},\, u\notin\{i,j,k\},\\ uv\in g,\, uw\notin g,\, uk\notin g}} \left[I_{\{\ell_{uw}>2\}}\frac{1+\gamma}{2} + I_{\{\ell_{uw}=2\,\wedge\, r_{uw}=1\}}\frac{\gamma}{2}\right] \\ &+ \delta \sum_{\substack{u\neq v\neq w:\\ v,w\in\{i,j\},\, u\notin\{i,j,k\},\\ uv\in g,\, uw\notin g,\, uk\in g}} \left[I_{\{r_{uw}=1\}}\frac{-\gamma}{2}\right]\end{aligned}$$

The result shows that (i) nodes not directly connected to k and connected to exactly one of the nodes to be introduced i, j can create *positive* payoffs from a new indirect connection if none existed before or from avoiding intermediation rents if paid previously. In addition, (ii) nodes connected directly to the introducing node k and exactly one of the nodes to be introduced i, j can generate a *negative* payoff from lost intermediation benefits to k if k was essential. It follows immediately that we can bound the payoff from an introduction from below by focusing on the number of neighbors of i or j that can generate a negative payoff for $\{i, j, k\}$ jointly.

Proposition 6. *Consider a network g such that there is an opportunity for k to introduce i and j. Let μ be the number of nodes u such that (i) $uk \in g$, (ii) $ui \in g$ or $uj \in g$, but not both, and (iii) k is essential for the indirect connection between u and j (or u and i), respectively. If g is myopically stable with introductions then*

$$\mu \geq \mu^* \equiv \left\lfloor 2\frac{1-c-(1-\epsilon)f-\delta}{\delta\gamma} \right\rfloor$$

Note that for μ^* to be positive requires an introduction to be profitable in isolation, that is, in a setting without any neighbor to ijk such that $1-c-(1-\epsilon)f-\delta > 0$.[8] Conditional on this term being positive, μ^* increases as intermediation profits γ become less important, as the loss of intermediation benefits to k is the friction that prevents introductions from being profitable. In addition, μ^* is increasing in ϵ, the cost advantage of introductions relative to direct bilateral link creation.

Proposition 6 connects any unused introduction opportunity to a minimum number of nodes that form a closed triangle. We can then establish the following result bounding the local clustering coefficient of any node in a myopically stable network based on the underlying parameters of the model.

Proposition 7. *Let g be a network that is myopically stable with introductions. Define $\lambda^* = \lfloor \frac{\mu^*}{2} \rfloor + 1$. Then the local clustering coefficient c_k of any node k with degree $d_k \geq 2$ in the network is bounded below by $\underline{c}(d_k)$ such that*

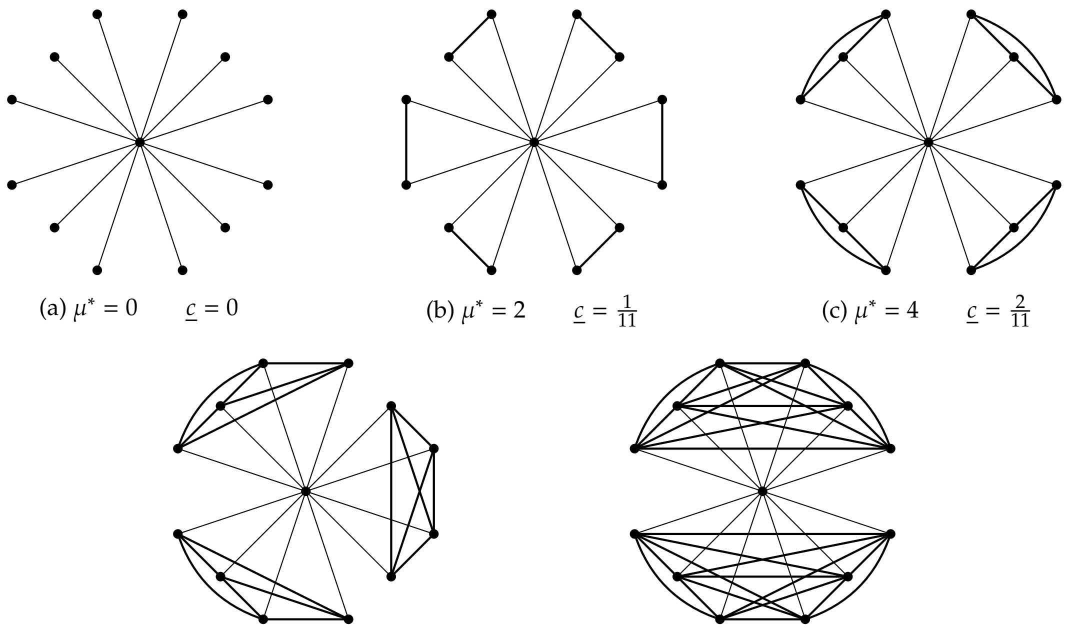
$$c_k \geq \underline{c}(d_k) \equiv \begin{cases} 1 & \text{if } d_k \leq \lambda^* \\ \left\lfloor \frac{d_k}{\lambda^*} \right\rfloor \frac{(\lambda^*-1)\lambda^*}{(d_k-1)d_k} & \text{if } d_k > \lambda^* \end{cases}$$

The bound is derived by assigning all neighbors of node k to cliques, that is, fully connected subnetworks (excluding links to k) of size λ^* nodes, which implies that each pair of unconnected nodes together have at least μ^* neighbors each that they share with k but that the nodes do not share with each other. This insures that any open triangles between cliques in the neighborhood of k are myopically stable by Proposition 6. The bound can be reached if μ^* is even and d_k is divisible by λ^* so that all neighbors of k are part of a clique of minimum size. If μ^* is odd, then stability requires that at most one clique can be of exactly size λ^* and all remaining ones have to be of size $\lambda^* + 1$. Furthermore, if d_k is not divisible by λ^*, the remaining nodes have

[8] This also implies $1-c-\delta > 0$ and thus that the complete network is stable.

to be added to existing cliques to ensure stability.

The higher μ^*, the higher the lower bound on local clustering. That is, as introductions become more profitable, minimum clustering increases. As the degree of k increases, the bound converges toward zero. Figure 3 illustrates Proposition 7 for a simple example of a central node with degree 12.[9]



(a) $\mu^* = 0$ $\underline{c} = 0$

(b) $\mu^* = 2$ $\underline{c} = \frac{1}{11}$

(c) $\mu^* = 4$ $\underline{c} = \frac{2}{11}$

(d) $\mu^* = 6$ $\underline{c} = \frac{3}{11}$

(e) $\mu^* = 10$ $\underline{c} = \frac{5}{11}$

Figure 3: Example of Minimum Clustering with Cliques of Even Size $\frac{\mu^*}{2} + 1$

The driving force behind the minimum clustering in this result is the introduction mechanism and not the presence of intermediation rents alone. This is in contrast to the results in Goyal and Vega-Redondo (2007), who focus on the incentives of disconnected players to create a new link bilaterally in order to avoid paying intermediation rents. In my model, the incentive for triadic closure arises from the benefits that a direct link offers relative to an indirect link if the intermediating player can be sufficiently compensated for the loss of intermediation benefits. Indeed, in Goyal and Vega-Redondo (2007) a higher intermediation rent parameter would likely increase the tendency of links to be formed,[10]whereas, in Proposition 7, higher intermediation rents per connection are associated with a lower μ^* and thus a lower minimum clustering coefficient and consequently higher realized intermediation payoffs for the central node.

9 Möhlmeier, Rusinowska, and Tanimura (2016) generate similar windmill shaped networks from very different mechanics in their model of network formation with both positive and negative externalities.

10 In their paper the authors hold the share of surplus captured by essential intermediaries fixed.

4.3 Stable Network Maximum Distance and Connectedness

Next, I study the maximum distance of networks that are pairwise stable with introductions. I identify an upper bound on the number of connections of the two highest degree nodes that are a distance of more than three apart.

Proposition 8. *Let g be a network that is pairwise stable with introductions. Let i, j be the pair of nodes with highest sum of degrees $d_i(g) + d_j(g)$ such that $\ell_{ij} > 3$. Then:*

$$d_i(g) + d_j(g) \leq \frac{2\left[f - (1 - c)\right]}{\delta(1 + \gamma)} \tag{2}$$

Corollary 9. *If search costs are sufficiently small such that $f < 1 - c + \delta\frac{1+\gamma}{2}$, any non-empty network that is a pairwise stable network with introductions will be connected.*

Note that the result leverages both the returns from indirect connections and the intermediation rents for nodes that bring together otherwise disconnected parts of the network. The threshold degree is falling as payoffs for indirect connections δ and returns for intermediation γ increase. In contrast, if $\delta \to 0$ the threshold tends to infinity if search costs exceed the direct benefits of a link ($f > (1 - c)$).

4.4 Discussion

In summary, the myopic stability analysis of the model with introductions reveals how the possibility of introductions creates a lower bound on the level of clustering, thereby matching the observed regularity of significant clustering in real-world networks. In addition, it finds that stable networks tend to be connected in a single component, which is consistent with the short distances often found in the data. The combination of clustering with the existence of bridging nodes as illustrated in Figure 3 relies on having both intermediation rents and introductions in the model. If introductions do not offer a cost advantage ($\epsilon = 0$), then whether or not an open triangle is closed depends only on the bilateral net benefit $1 - c - f - \delta$ and thus this mechanism will either close all open triangles or not. Likewise, if there are no intermediation rents ($\gamma = 0$), then there is no loss for introducing agents in the sense of Proposition 5. This implies that the number of closed triangles required to support any open triangle in the sense of Proposition 6 goes to infinity, resulting in all introductions taking place. There would be no structural holes remaining.

Also note that while the tendency for connected components and increased clustering through introductions captures important features of real-world networks, the analysis also brings out limits to this process. Specifically, the process of link creation stops with bridge agents that connect otherwise disconnected parts of the network. If these agents earn sufficiently high intermediation rents, then their incentives will be such so as to prevent connections forming between the two parts on either side of the bridge. The myopic stability analysis thus provides an explanation for connectedness, high clustering and the long-run existence of bridging agents that connect across structural holes.

5 Conclusion

This paper analyzes network formation when the creation of new links is mediated by existing connections. Specifically, I consider a setting in which players that are unconnected but share a direct neighbor can be "introduced" by that neighbor. The paper presents an analysis of a suitable adapted notion of pairwise stability with introductions to study the incentives involved in introductions and the impact on network outcomes.

I find that myopically stable networks include efficient configurations for some parameter ranges. As in other models, the myopic stability concept suffers from significant multiplicity, but nonetheless, predictions can be derived regarding the level of connectedness. In addition, the analysis sheds light on the extent to which introductions in the model are used by nodes to create networks with high levels of clustering, mirroring properties observed in real-world networks.

The paper offers insights into how simple local dynamics can lead to high levels of local clustering, as players coordinate among more than two players to take advantage of payoff advantages from introductions. In applications, such network-based coordination may explain the high levels of clustering observed in real-world networks, even in settings where it may be difficult for two players to connect independently. More generally, the analysis suggests how the involvement of additional players and transfers may help to deal with externalities that tend to push equilibrium outcomes toward over- or under-connected networks. Future research may consider how these insights generalize to more general settings, such as coalitions of more than three players forming connected cliques.

References

Adamic, **Lada A.** (1999). “The Small World Web.” In: *Research and Advanced Technology for Digital Libraries*. Ed. by **Serge Abiteboul** and **Anne-Marie Vercoustre**. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 443–452. DOI: 10.1007/3-540-48155-9_27.

Bloch, **Francis** and **Matthew O. Jackson** (2006). “Definitions of Equilibrium in Network Formation Games.” *International Journal of Game Theory* 34.3 (Oct. 2, 2006), pp. 305–318. DOI: 10.1007/s00182-006-0022-9.

Bloch, **Francis** and **Matthew O. Jackson** (2007). “The Formation of Networks with Transfers among Players.” *Journal of Economic Theory* 133.1 (Mar. 2007), pp. 83–110. DOI: 10.1016/j.jet.2005.10.003.

Burt, **Ronald S.** (1992). “Structural Holes: The Social Structure of Competition.” *Cambridge: Harvard*.

Burt, **Ronald S.** (2007). “Secondhand Brokerage: Evidence on the Importance of Local Structure for Managers, Bankers, and Analysts.” *Academy of Management Journal* 50.1 (Feb. 2007), pp. 119–148. DOI: 10.5465/amj.2007.24162082.

Chandrasekhar, **Arun** and **Matthew O. Jackson** (2018). *A Network Formation Model Based on Subgraphs*. June 2018. arXiv: 1611.07658.

Currarini, **Sergio** and **Massimo Morelli** (2000). “Network Formation with Sequential Demands.” *Review of Economic Design* 5.3 (Sept. 1, 2000), pp. 229–249. DOI: 10.1007/PL00013689.

De Paula, **Aureo** (2019). *Econometric Models of Network Formation*. Sept. 2019. arXiv: 1910.07781.

Goyal, **Sanjeev**, **Marco J. van der Leij**, and **José L. Moraga-González** (2006). “Economics: An Emerging Small World.” *Journal of Political Economy* 114.2 (Apr. 2006), pp. 403–412. DOI: 10.1086/500990.

Goyal, **Sanjeev** and **Fernando Vega-Redondo** (2007). “Structural Holes in Social Networks.” *Journal of Economic Theory* 137.1 (Nov. 2007), pp. 460–492. DOI: 10.1016/j.jet.2007.01.006.

Graham, **Bryan S.** (2019). *Dyadic Regression*. Aug. 23, 2019. arXiv: 1908.09029.

Grossman, **Jerrold W.** (2002). “The Evolution of the Mathematical Research Collaboration Graph.” *Congressus Numerantium*, pp. 201–212.

Jackson, **Matthew O.** (2008). *Social and Economic Networks*. Princeton University Press. 519 pp. DOI: 10.1515/9781400833993.

Jackson, **Matthew O.** and **Brian W. Rogers** (2005). “The Economics of Small Worlds.” *Journal of the European Economic Association* 3 (2/3, 2005). DOI: 10.1162/jeea.2005.3.2-3.617.

Jackson, **Matthew O.** and **Brian W. Rogers** (2007). “Meeting Strangers and Friends of Friends: How Random Are Social Networks?” *American Economic Review* 97.3 (May 2007), pp. 890–915. DOI: 10.1257/aer.97.3.890.

Jackson, **Matthew O.**, **Brian W. Rogers**, and **Yves Zenou** (2017). "The Economic Consequences of Social-Network Structure." *Journal of Economic Literature* 55.1 (Mar. 2017), pp. 49–95. DOI: 10.1257/jel.20150694.

Jackson, **Matthew O.** and **Asher Wolinsky** (1996). "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory* 71.1 (Oct. 1, 1996), pp. 44–74. DOI: 10.1006/jeth.1996.0108.

Kleinberg, **Jon**, **Siddharth Suri**, **Éva Tardos**, and **Tom Wexler** (2008). "Strategic Network Formation with Structural Holes." *ACM SIGecom Exchanges* 7.3 (Nov. 1, 2008), pp. 1–4. DOI: 10.1145/1486877.1486888.

Mayer, **Adalbert** and **Steven L. Puller** (2008). "The Old Boy (and Girl) Network: Social Network Formation on University Campuses." *Journal of Public Economics* 92.1-2 (Feb. 2008), pp. 329–347. DOI: 10.1016/j.jpubeco.2007.09.001.

Milgram, **Stanley** (1967). "The Small-World Problem." *Psychology Today* 1.1.

Möhlmeier, **Philipp**, **Agnieszka Rusinowska**, and **Emily Tanimura** (2016). "A Degree-Distance-Based Connections Model with Negative and Positive Externalities." *Journal of Public Economic Theory* 18.2, pp. 168–192. DOI: 10.1111/jpet.12183.

Newman, **Mark E. J.** (2003). "The Structure and Function of Complex Networks." *SIAM Review* 45.2 (Jan. 2003), pp. 167–256. DOI: 10.1137/S003614450342480. arXiv: cond-mat/0303516.

Siedlarek, **Jan-Peter** (2015). "Intermediation in Networks." *Cleveland Fed Working Papers* (WP 15-18 Oct. 2015). DOI: 10.26509/frbc-wp-201518.

Uzzi, **Brian** (1996). "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61.4 (Aug. 1996), p. 674. DOI: 10.2307/2096399.

Vázquez, **Alexei** (2003). "Growing Network with Local Rules: Preferential Attachment, Clustering Hierarchy, and Degree Correlations." *Physical Review E* 67.5 (May 7, 2003). DOI: 10.1103/PhysRevE.67.056104.

Vega-Redondo, **Fernando** (2007). *Complex Social Networks*. Cambridge University Press, Jan. 8, 2007. 316 pp. DOI: 10.1017/cco10521857406.

Watts, **Duncan J.** (2004). *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton Studies in Complexity. OCLC: 248841323. Princeton, NJ: Princeton Univ. Press. 262 pp.

Watts, **Duncan J.** and **Steven H. Strogatz** (1998). "Collective Dynamics of Small-World Networks." *Nature* 393.6684 (June 1998), pp. 440–442. DOI: 10.1038/30918.

Appendix

A Proofs

A.1 Proof of Proposition 2 – Efficient Network Configurations

Proof. The proof adapts the efficiency results for the standard connections model described in Jackson (2008, Chapter 6.3). An important difference here is the distinction between link creation through search and introduction, which requires a discussion of link costs. For any component in a given network g, Lemma 10 provides a lower bound on the number of links that need to be created by search. The proof is immediate and omitted here.

Lemma 10. *The number of links created through search to create a component of k nodes is at least $k-1$. The minimum cost to form a component of k nodes with $m \geq k-1$ nodes is $(k-1)\,f + (m-(k-1))\,(1-\epsilon)f$.*

Having established the minimum-cost way to form any network g, we can proceed to the proof of efficient network structure.

First, consider the case with $1-c < \delta$. I will argue that the star is the efficient configuration to connect k nodes in this case. A star network of k nodes incurs search costs of exactly $(k-1)f$ and generates a net benefit including costs of link formation of:

$$(k-1)\,[1-c] + \frac{(k-1)(k-2)}{2}\delta - (k-1)f \tag{3}$$

Now any other configuration connecting k nodes with $m \geq k-1$ links will generate at most

$$m\,[1-c] + \left(\frac{k(k-1)}{2} - m\right)\delta - (k-1)f - (m-(k-1))\,(1-\epsilon) \tag{4}$$

where the first component represents the benefits from direct links and the second component reflects the upper bound of the benefits that can be derived from any indirectly connected nodes. The third component reflects the minimum amount of search costs incurred. Subtracting the second equation from the first and rearranging yields the payoff advantage for the star relative to any other configuration of k nodes:

$$(k-1-m)\left[1-c-\delta-(1-\epsilon)f\right] \tag{5}$$

As $1-c < \delta + (1-\epsilon)f$ and $m \geq k-1$, this implies that the payoff advantage is minimized at $m = k-1$.

It is thus established that if $1-c < \delta + (1-\epsilon)f$, efficient networks consist of stars and isolated nodes. Next, we restrict the set of candidate efficient networks for $1-c < \delta + (1-\epsilon)f$ further by establishing that there is either a single star of all nodes or an empty network.

Assume a candidate network consisting of two stars with $k_1 \geq 1$ and $k_2 \geq 2$ nodes yielding positive utility each. Then total payoff is:

$$(k_1 - 1)\left[1 - c - f + (k_1 - 2)\frac{\delta}{2}\right] + (k_2 - 1)\left[1 - c - f + (k_2 - 2)\frac{\delta}{2}\right] \tag{6}$$

Reconfiguring the nodes into a single star yields:

$$(k_1 + k_2 - 1)\left[1 - c - f + (k_1 + k_2 - 2)\frac{\delta}{2}\right] \tag{7}$$

Now, subtracting the first equation from the second and simplifying yields:

$$\left[1 - c - f\right] + (2k_1 k_2 - 2)\frac{\delta}{2} \tag{8}$$

which is strictly positive if each separate star yields positive utility as $2k_1k_2 > k_1$ and $2k_1k_2 > k_2$. Thus, a network with more than one separate stars (including the case where all but one star are single disconnected nodes) yields lower utility than one in which the nodes involved are combined into a single star.

The treatment of the case where $1 - c < \delta + (1 - \epsilon)f$ is concluded by comparing payoffs of a single star involving n nodes and the empty network. As the latter derives zero utility, the star is the unique efficient structure if:

$$(n - 1)(1 - c) + \frac{(n-1)(n-2)}{2}\delta - (n - 1)f > 0 \tag{9}$$

which reduces to:

$$1 - c > f - \frac{(n-2)}{2}\delta \tag{10}$$

Next consider the case with $1 - c \geq \delta + (1 - \epsilon)f$. In this case, adding a link by introduction weakly increases utility. Thus, any link created by introductions creates positive utility. For a component of k nodes, the minimum number of links created through search is $k - 1$. Given $k - 1$ links created by search, the structure that maximizes the number of introductions is the star, which allows a fully connected component of size k to be formed with $k - 1$ links created by search and all remaining links formed by introduction.

Now, for the case where there are more than one such fully connected components, an argument analogous to that for two stars shows that utility increases in a single connected component. Thus, if

$1 - c \geq \delta + (1-\epsilon)f$, the efficient network is either complete or empty. Utility from the complete network with $n-1$ links created by search and all remaining links by introduction is strictly higher if:

$$(n-1)\frac{n}{2}[1-c] - (n-1)f - (n-1)\frac{(n-2)}{2}(1-\epsilon)f > 0 \tag{11}$$

which reduces to

$$1 - c > \left[1 - \frac{n-2}{2}\epsilon\right] f \tag{12}$$

□

A.2 Proof of Proposition 4 – Stability of Efficient Network Configurations

This section derives the parameter restrictions for the stability of the three network configurations that can be efficient, including the empty network, the star network, and the complete network.

Empty Network In the empty network, no links exist implying $\pi_i(g) = 0 \; \forall \; i$ and thus the only condition to verify is bilateral link creation, yielding the stability condition:

$$1 - c \leq f \tag{13}$$

Thus the empty network is stable if net benefits from a single new link do not outweigh the search costs.

Star Network In the star network, the set of possible deviations to consider applies to two types of nodes (hub and spoke) and there are opportunities to destroy links as well as to create links through search and introductions:

(a) The star network is stable against *link destruction* by the hub and one peripheral node if:

$$1 - c \geq -(n-2)\frac{1+\gamma}{2}\delta \tag{14}$$

(b) The star network is stable against *link creation by search* of two peripheral nodes if:

$$1 - c \leq f + (1-\gamma)\delta \tag{15}$$

(c) The star network is stable against *introduction* of a pair of peripheral nodes through the hub if:

$$1 - c \leq \delta + (1 - \epsilon) f \tag{16}$$

Complete Network Total net payoffs from the complete network before search costs and transfers are given by:

$$\pi(g) = \frac{n-1}{2}(1 - c) \tag{17}$$

As all possible links are in place, the only deviation to consider is whether a pair of players would find it profitable to destroy a link. In this case, a single direct link is replaced with a set of $n - 2$ indirect connections, one through each of the other players that remain connected to i. For $n > 3$, this implies that there are at least two intermediaries and the two disconnecting players jointly capture δ. The complete network is thus stable against link destruction if:

$$1 - c \geq \delta \tag{18}$$

A.3 Proof of Proposition 5

Proof. The introduction of i and j by k creates a new link ij. The direct effect on payoffs of i, j and k jointly results from replacing an indirect connection between i and j with a direct connection at the cost of the introduction. The direct effect on payoffs is thus given by

$$1 - c - (1 - \epsilon) f - \delta$$

In addition, the new link ij has implications for payoffs generating connections with other nodes. These effects only involve nodes that are neighbors of exactly one of i and j. This is because the new link ij affects payoffs arising from another node $u \notin \{i, j, k\}$ only if either (i) link ij shortens an existing path or (ii) provides a new path of length two (thereby creating connection benefits or cutting intermediation rents). Neither is possible if node u is at a distance of at least two from both i and j. Furthermore, if a node u is connected to both i and j, then the new link ij does not affect any connection where u acts as an end node. In addition, even though the link ij replaces intermediation via u with a direct link on the connection between i and j, this has no payoff implications as the existence of such u implies that there are at least two intermediaries for the i, j connection: u and k.

We thus focus on the changes to payoffs for ijk arising from nodes u that are neighbors of exactly one of

i and j. There are then two cases to consider depending on whether or not u is connected to k as well.

1. $u \neq v \neq w : v, w \in \{i, j\}, u \notin \{i, j, k\}, uv \in g, uw \notin g, uk \notin g$

   In this case, u is not connected to k. The new link ij created by the introduction will create a new connection of length two between u and i or j. If prior to the introduction there was no such connection of length two, then the payoff to ijk will be $\delta\frac{1+\gamma}{2}$ as i and j will act as both an end node and as a unique intermediary to the link with u. If there already exists a connection of length two between u and one of i or j (intermediated by a neighbor of i, j, respectively) and this connection had one intermediary only, then the new link ij will create an alternative indirect connection and reduce the intermediation rents that have to be paid. Thus, payoffs to ijk will increase by $\frac{\gamma}{2}$.

2. $u \neq v \neq w : v, w \in \{i, j\}, u \notin \{i, j, k\}, uv \in g, uw \notin g, uk \notin g$

   In this case, u is connected to k and thus both i and j will already have a connection of length less than two to u, one with a direct link and one via intermediation through k. Thus, the new link ij created by the introduction will not provide a completely new connection. However, if k is essential to the connection of u and i or j, then the new link ij will create an alternative connection of length two and thus the intermediation benefits captured by k will be lost. Only part of this will be captured by i and j and thus the overall loss to ijk is $\delta\frac{\gamma}{2}$

Combining direct and indirect effects yields the expression in Proposition 5.

□

A.4 Proof of Proposition 6

Proof. The proof is by contradiction. Assume that network g is myopically stable with introductions and there exists an unused introduction opportunity for k to introduce i and j and μ, that is, the number of nodes u connected to k and either i (or j) such that k is essential for the connection between u and j (or i, respectively) is strictly less than $\mu^* = \left\lfloor 2\frac{1-c-(1-\epsilon)f-\delta}{\delta\gamma} \right\rfloor$.

Substituting the expression in Proposition 5 and dropping the non-negative middle term, the payoff change to ijk is then at least

$$\begin{aligned}\Delta\pi_{ijk} \geq & 1 - c - (1-\epsilon)f - \delta - \mu\frac{\delta\gamma}{2} \\ > & 1 - c - (1-\epsilon)f - \delta - \mu^*\frac{\delta\gamma}{2} \\ = & 1 - c - (1-\epsilon)f - \delta - \left\lfloor 2\frac{1 - c - (1-\epsilon)f - \delta}{\delta\gamma} \right\rfloor \frac{\delta\gamma}{2} \\ \geq & 1 - c - (1-\epsilon)f - \delta - \left(2\frac{1 - c - (1-\epsilon)f - \delta}{\delta\gamma}\right) \frac{\delta\gamma}{2} \\ = & 1 - c - (1-\epsilon)f - \delta - \left(1 - c - (1-\epsilon)f - \delta\right) \\ = & 0\end{aligned}$$

The introduction is thus jointly profitable to ijk, delivering the contradiction.

□

A.5 Proof of Proposition 7

Proof. The proof is by contradiction. Let g be myopically stable with introductions and with node k with degree d_k such that the clustering coefficient c_k is strictly below $\underline{c}(d_k) = \lfloor \frac{d_k}{\lambda^*} \rfloor \frac{(\lambda^*-1)\lambda^*}{(d_k-1)d)_k} \leq \frac{\lambda^*-1}{d_k-1}$.

First some notation. Let the set of nodes that are neighbors of k in g be denoted by K. Define by g_K the subnetwork of g formed by the set of nodes K and the links that are links between nodes in K in g. The degree of node i within g_K is labeled d_i^K and the average degree across K is labeled d_K. Note that all the links in g_K are between neighbors of k and thus these are the links that are required for stability against introductions in the sense of Proposition 6.

We distinguish between two cases based on the number of nodes in the neighborhood of k:

Case 1: $d_k \leq \lambda^*$: If $d_k \leq \lambda^*$, then $\underline{c}(d_k) = 1$, that is, the neighbors of k form a fully connected subnetwork. Thus if $c_k < \underline{c}(d_k)$ there exists at least one pair i, j in the neighborhood of k with $ij \notin g_K$. Then the maximum number of links of both i and j within g_K is strictly less than $\lambda^* - 1 = \frac{\mu^*}{2}$ and thus the total number of neighbors that the pair i and j share with k but not with each other is strictly less than μ^*. By Proposition 6 this implies that g is not myopically stable with introductions, delivering the contradiction.

Case 2: $d_k > \lambda^*$**:** If $d_k > \lambda^*$, then $\underline{c}(d_k) = \left\lfloor \frac{d_k}{\lambda^*} \right\rfloor \frac{(\lambda^*-1)\lambda^*}{(d_k-1)d_k}$. Now, assume that this condition is violated and thus

$$\begin{aligned} c_k &< \underline{c}(d_k) \\ &= \left\lfloor \frac{d_k}{\lambda^*} \right\rfloor \frac{(\lambda^*-1)\lambda^*}{(d_k-1)d_k} \\ &\leq \frac{(\lambda^*-1)}{(d_k-1)} \end{aligned}$$

As the total number of triangles around k is given by $\frac{(d_k-1)d_k}{2}$, the condition on clustering c_k implies that the total number of closed triangles around k, and thus links, in network g_K is at most $(\lambda^*-1)\,\frac{d_k}{2}$. With each link being shared by two nodes and a total of K nodes in g_K, it follows that the average degree d_K is strictly less than $\lambda^* - 1 = \frac{\mu^*}{2}$.

We can now show that $\min_{ij \notin g_K} \left[d_i^K + d_j^K \right] < \mu^*$, that is, there is at least one pair of nodes with no link between them whose degrees in g_K add up to less than μ^*. By Proposition 6 this implies that the network is not myopically stable with introductions.

$$\min_{ij:i,j\in K, ij\notin g_K} \left[d_i^K + d_j^K \right] \leq \frac{1}{|\{ij : i,j \in K, ij \notin g_K\}|} \sum_{ij:i,j\in K, ij\notin g_K} \left[d_i^K + d_j^K \right] \tag{19}$$

$$\leq \frac{1}{|\{ij : i,j \in K\}|} \sum_{ij:i,j\in K} \left[d_i^K + d_j^K \right] \tag{20}$$

$$= \frac{2}{(d_k-1)} \sum_{ij:i,j\in K} \left[d_i^K + d_j^K \right] \tag{21}$$

$$= \frac{2}{d_k(d_k-1)} \sum_{i\in K} \sum_{j\in K:j>i} \left[d_i^K + d_j^K \right] \tag{22}$$

$$= \frac{1}{d_k(d_k-1)} \sum_{i\in K} \sum_{j\in K:j\leq i} \left[d_i^K + d_j^K \right] \tag{23}$$

$$= \frac{1}{d_k(d_k-1)} \left\{ \sum_{i\in K} \sum_{j\in K:j\neq i} d_i^K + \sum_{j\in K} \sum_{i\in K:i\neq j} d_j^K \right\} \tag{24}$$

$$= \frac{1}{d_k(d_k-1)} \left\{ \sum_{i\in K} (d_k-1) d_i^K + \sum_{j\in K} (d_k-1) d_j^K \right\} \tag{25}$$

$$= \frac{1}{d_k} \left\{ \sum_{i\in K} d_i^K + \sum_{j\in K} d_j^K \right\} \tag{26}$$

$$= 2d_K < \mu^* \tag{27}$$

The first step uses the fact that minimum is less than the average. In the second step we exploit the fact that the average of the sum of degrees across absent links is weakly less than the average across all links. This follows from the fact that in this expression conditioning on the absence of a link underweights high degree nodes relative to low degree nodes.[11]

□

A.6 Proof of Proposition 8

Proof. The proof is by contradiction. Assume that network g is myopically pairwise stable with introductions. Further, assume that there are i, j violating 2 such that $\ell_{ij} > 3$ and $d_i(g) + d_j(g) > \frac{2[f-(1-c)]}{\delta(1+\gamma)}$.

Consider a new connection being formed between i and j. The new link will create a direct connection with direct net benefit $1 - c$ at search cost f. In addition there are $d_i(g) + d_j(g)$ new connections of length two and creating total benefit δ. For each such connection, i and j form an end node and, as $\ell_{ij} > 3$, an essential intermediary, respectively, and thus capture a benefit of $\frac{1+\gamma}{2}\delta$.

Total payoff from the connection to i and j is thus

$$\Delta\pi_{ij} = \left[d_i(g) + d_j(g)\right]\delta\frac{1+\gamma}{2} + \left[1 - c - f\right] \tag{28}$$

$$> -\left[1 - c - f\right] + \left[1 - c - f\right] \tag{29}$$

$$=0 \tag{30}$$

and thus the additional link is jointly profitable and the network g is not pairwise stable, delivering the contradiction.

□

[11] See Vega-Redondo (2007, p.44f) for a discussion of this point.