Federal Reserve Bank of Cleveland Working Paper Series

Making Friends Meet: Network Formation with Introductions

Jan-Peter Siedlarek

Working Paper No. 20-01R2

June 2022

Suggested citation: Siedlarek, Jan-Peter. 2022. "Making Friends Meet: Network Formation with Introductions." Working Paper No. 20-01R2. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202001r2.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Making Friends Meet: Network Formation with Introductions*

Jan-Peter Siedlarek†

June 2022

Abstract

This paper proposes a parsimonious model of network formation with introductions in the presence of intermediation rents. Introductions allow two nodes to form a new connection on favorable terms with the help of a common neighbor. The decision to form links via introductions is subject to a trade-off between the gains from having a direct connection at lower cost and the potential losses for the introducer from lower intermediation rents. When nodes take advantage of introductions, stable networks tend to exhibit a minimum amount of clustering. At the same time, intermediary nodes have incentives to protect their position, and stable networks can exhibit nodes exploiting structural holes, that is, bridges across otherwise unconnected parts of the network earning intermediation rents.

JEL Classification: A14, D85

Keywords: networks; network formation; clustering; structural holes; intermediation; introductions

*Acknowledgments: I am grateful to Fernando Vega-Redondo for extensive discussions and advice. I received many helpful comments and suggestions from the associate editor and two anonymous referees, as well as Ben Craig, Michael Gofman, Roberto Pinheiro, Peter Vida, and seminar audiences at the European University Institute, the University of Mannheim, PET 2013 Lisbon and the EEA 2013 Göteborg. I thank Jovial Clayton for excellent research support. All remaining errors are mine. This research was in part supported by funding from the DAAD (German Academic Exchange Service) and the DFG through grant SFB-TR15.

†Federal Reserve Bank of Cleveland, janpeter.siedlarek@clev.frb.org, https://www.jsiedlarek.net *Disclaimer:* The views stated herein are those of the author and are not necessarily those of the Federal Reserve Bank of Cleveland or of the Board of Governors of the Federal Reserve System.

1 Introduction

When new connections are being formed in business or social networks, existing connections can play an important role in facilitating new links. A simple illustration of such network-based linking involves an introduction or referral, which creates a new connection between two previously unconnected nodes, thanks to a common acquaintance that brings them together. This paper proposes a model to study how incorporating introductions in a strategic network formation model affects the characteristics of the networks that result.

Link creation between nodes sharing a common neighbor is known as triadic closure, and it has been documented to be an important dynamic in the formation of networks in a number of different settings.[1] Uzzi (1996), in an in-depth study of the apparel industry in New York, describes how new relationships in business networks 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, emphasis added)

Also in the context of business networks, Kogut and Walker (2001) investigate the co-ownership network of German firms and identify a role for firms they call "brokers" that facilitate new acquisitions based on their existing connections and tend to be placed in central positions. In a different setting, Chaney (2014) shows that new connections in international trade networks can be explained with a dynamic model in which exporting firms create new trade relationships by exploiting their existing export contacts in the geographic vicinity of the desired new destination. More generally, business referrals can play an important role in establishing new relationships between market participants, thereby enabling trading to take place. Similar dynamics with network-based link creation have been observed in social networks. For example, Mayer and Puller (2008) study the dynamics of a network of relationships at a university. They show that an existing common neighbor within the network is a strong predictor of the formation of a new link between two nodes.

A key aspect of introductions as a way to create new relationships is the active role played by the go-between in the creation of the new connection. Without the agreement of the referring party, the new connection could not access the benefits of the introduction. This paper studies such introductions including the incentives for the parties involved and their implications for the formation of social and business networks.

Why would the presence of a common neighbor support the efforts of a pair of nodes to create

[1] See, for example, Easley and Kleinberg (2010, Chapter 3) for an introduction.

a connection? As noted in the passage from Uzzi (1996) quoted above, a common neighbor can facilitate new links and help to build trust between two potential partners where this would be difficult without a go-between. In addition, links supported by common neighbors may be able to access greater benefits than connections without such support, for example to help enforce risk agreements among individuals (Jackson, Rodriguez-Barraquer, and Tan, 2012), or by offering redundancy advantages (Renou and Tomala, 2012). In sociology, Coleman (1988) makes a strong argument for the benefits derived by individuals that are embedded in densely connected networks and that he labels "network closure". The model in this paper captures the advantages of links created by introduction on the cost side, assigning them a lower cost than links created bilaterally.

Aside from these benefits, introductions also come with potential downsides to the nodes involved. The key one for the purposes of this paper concerns the redistribution of intermediation rents. Specifically, where a network node in a structurally important position between two other unconnected nodes earns rents for intermediating an indirect connection between them, such a go-between potentially risks losing these rents when new links are created among the intermediary's neighbors. Intermediation rents have been studied extensively in sociology and economics, in particular in the context of business and organizational networks. In his work on network advantage in such contexts, Burt (1992) labels such positions "structural holes" and provides evidence for the payoffs that individuals in these positions can earn. Since this earlier study, follow-up work has found additional evidence of the value of holding a central position in the network in various settings, including intrafirm organizational networks and interfirm R&D collaboration (Podolny and Baron, 1997; Ahuja, 2000; Mehra, Kilduff, and Brass, 2001; Owen-Smith and Powell, 2004). In economics, returns for intermediaries in networks have been studied by, among others, Blume et al. (2009), Condorelli, Galeotti, and Renou (2017), Farboodi, Jarosch, and Menzio (2017), and Manea (2018).

This paper presents a parsimonious model to study network formation with introductions in networks with intermediation rents, such as in the business networks studied in Burt (1992) and the related papers above. The model is a variant of the connections model of Jackson and Wolinsky (1996) with intermediation rents added, in the spirit of Goyal, van der Leij, and Moraga-González (2006). Network connections generate payoffs for the parties connected, directly or indirectly. Furthermore, as in Goyal and Vega-Redondo (2007) and Kleinberg et al. (2008), there are returns for intermediaries that connect otherwise disconnected nodes.

New links can be formed bilaterally or through introductions. Introductions can create a new link if the new link connects two nodes that share a common neighbor and all three parties agree. Introductions have a cost advantage over bilateral link formation, but they potentially threaten the rents for the central go-between.

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 network features such as high clustering, short distances, and network bridges forms the core of the paper.

The paper shows first that efficient network structures are either empty networks, star networks, or complete networks, reflecting the familiar patterns from the baseline connections model of Jackson and Wolinsky (1996). It then focuses on incentives for decentralized link formation by studying the set of networks that are stable to deviations by link creation or destruction. The analysis shows that the set of efficient network configurations is not necessarily stable for a given parameter configuration, revealing externalities in link creation and destruction. More generally, if the benefits to linking are sufficiently high, stable networks tend to be connected and there exists a limit on the distance between any pair of nodes, given their degree. Next, the paper analyzes the use of introductions. It derives a lower bound on the clustering coefficient for any node in stable networks, with introductions providing the impetus for closing open triangles. The bound depends on the strength of returns to intermediation as well as the cost advantage of introductions over bilateral link formation. If parameters are such that nodes can take advantage of introductions, the predicted network exhibits features of a small world (Watts and Strogatz, 1998) with high clustering and short distances, a structure that has been identified in business networks (Kogut and Walker, 2001). What limits the closing of all open triangles through introductions is the incentive for essential intermediaries to protect the rents they receive from connecting otherwise disconnected parts of the network. These incentives imply that where opportunities for introductions remain unused in stable networks they act as bridges across densely connected subnetworks. Thus, under suitable conditions, links bridging otherwise disconnected parts of the network and earning substantial returns to intermediation can coexist with high clustering in stable networks, capturing the structural holes of Burt (1992) as documented in empirical settings, for example, in Owen-Smith and Powell (2004).

The paper contributes to the literature on network formation related to the study of the introduction mechanism and the compensation of intermediate nodes. While intuitive, studying link creation as a process conditional on the network in place at a given time is relatively novel in the literature on strategic network formation.[2] The two papers closest to this paper are Jackson and

[2] 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.

Rogers (2007) and Jackson and Rogers (2005). In Jackson and Rogers (2007) the authors study a random 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 nodes, that is, 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. Using a combination of both processes, 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 different 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 the implications for returns from other parts of the network.[3] My model puts such externalities at the center of the analysis, studying how incentives for intermediaries to protect their returns shape the resulting network and can lead to outcomes combining both high clustering and bridging across structural holes.

Beyond Jackson and Rogers (2007), an earlier contribution by Vázquez (2003) presents a mean-field analysis of graph dynamics with a nearest-friend link creation process. That contribution ignores strategic considerations 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. For example, Chandrasekhar and Jackson (2018) 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) and 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" cost model, in which links connecting players on the same island are cheaper than links between islands. The present model similarly explains both clustering and short distances but, in addition, analyzes the effect of intermediation benefits and the resulting incentives for individuals to adopt and protect bridging positions. In Jackson and Rogers (2005) the grouping of players into easily connected islands is given exogenously; in my paper, the cost

[3] 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 include an analysis of the externalities resulting from intermediation returns.

advantages arise endogenously through the introduction mechanism. This endogenous, network-based cost structure is one of the key differences between the present paper and Jackson and Rogers (2005).

In strategic network formation, Priazhkina and Page (2018) study referrals or market-sharing by sellers in buyer-seller networks. Although this setting is different from the connections model studied here, referrals are related to introductions in that they require an existing seller-buyer relationship in order for the seller to be able to execute a referral. In their paper, the authors study the incentives for referrals and the implications for market functioning.

In addition, this paper offers a contribution to the literature on incentives for intermediary nodes and rents from intermediation. The benefits accruing to nodes in specific positions crucial to the connectivity of 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. Subsequent papers have studied the tension between triadic closure and structural holes. In the economics literature, Goyal and Vega-Redondo (2007) 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. Other related papers in this vein include Buskens and van de Rijt (2008), Kleinberg et al. (2008) and Kossinets and Watts (2006).

Ambrus and Elliott (2021) consider stable and efficient configurations of risk-sharing networks. They find a trade-off between stability and inequality: The most stable networks tend to be the most unequal as well, with central agents connecting more peripheral ones and capturing most of the benefits of the network. As in the present paper, Ambrus and Elliott (2021) find that stable networks tend to both exhibit short distances and include some nodes that bridge others and earn high returns. In their main model, closed triangles have no advantage over open triangles and indeed a closed triangle would be inefficient. In an extension Ambrus and Elliott (2021) study the implication of requiring agents to have a common friend for risk-sharing to work. This requirement implies that at least three nodes need to be involved for risk-sharing, and for this reason Ambrus and Elliott (2021) also modify their initially pairwise stability notion to allow for deviations by triplets. This concept of a triplet-wise deviation mirrors the notion of deviations by introduction used in this paper. In that version of the model, network structures that are similar to the windmill graphs in Figure 5 arise as efficient and stable configurations.

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 the 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.

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 short distances, intermediation benefits, and minimum clustering. Finally, Section 5 concludes.

2 A Model of Network Formation with Intermediation and Introductions

This section introduces a model of network formation with introductions and intermediation rents. It builds on the symmetric connections model of Jackson and Wolinsky (1996): Nodes create payoffs from direct and indirect connections and there is a cost to maintaining links. On top of this, the model in this paper introduces rents for link intermediation as well as a new way of creating links via introductions.

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. The network 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 connections between pairs of players. Direct connections generate a normalized surplus of 1 less a linking cost c_{ij}. Indirect connections that are intermediated by another node generate a surplus of $\delta \leq 1$. Following Kleinberg et al. (2008), I assume that connections requiring more than two steps generate a zero surplus. While there may be other contexts where the decay of returns from connections is slower, the assumption is consistent with the empirical evidence on returns to intermediation and brokerage in organizational networks as reported by Burt (1992, 2007). Burt (2007) studies the returns to secondhand brokerage, that is, where more than one intermediary is required to connect two parties, in three different organizational networks. He finds substantial returns to direct information brokerage and documents a lack of returns to secondhand brokerage in all three groups, concluding that in the context of organizational networks, benefits for nodes that are greater than two steps away are negligible, and argues in support of an analytical framework that ignores such higher-order intermediation.

The costs and benefits generated by direct connections without an intermediary are split equally between the two nodes involved; that is, each receives a payoff of $\frac{(1-c_{ij})}{2}$.

When indirect connections are facilitated by an intermediary, that node can capture a share of the surplus created by the connection, akin to the returns to brokerage identified in Burt (1992) and subsequent work. Such intermediation profits can also capture the benefits for intermediaries in trading networks as studied in Condorelli, Galeotti, and Renou (2017), Farboodi, Jarosch, and Shimer (2017), Choi, Galeotti, and Goyal (2017), and Manea (2018).

The value of intermediation rents depends on the extent to which the intermediary is indispensable for the connection: Where many alternatives exist, rents available to an intermediary will be lower. Formally, total rents for intermediation between i and j are denoted $\gamma(r_{ij}(g))$, which is a function of $r_{ij}(g)$, the number of intermediaries, that is, nodes that are connected to both i and j in g. Following Goyal and Vega-Redondo (2007) I 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 $\gamma(r_{ij}(g)) = \gamma$ if $r_{ij}(g) = 1$ and $\gamma(r_{ij}(g)) = 0$ otherwise. This assumption captures intermediation rents being competed away in the spirit of Bertrand competition as soon as there is more than one intermediary.[4] The assumption further reflects the experimental evidence presented in Choi, Galeotti, and Goyal (2017). They study trading efficiency and surplus division across different network configurations and find that intermediaries that are on all paths between a buyer and a seller node extract a large share of the surplus, while intermediaries in positions where alternatives exist receive payoffs that are close to zero.

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

$$\tilde{\pi}_i(g) = \underbrace{d_i(g)\frac{1-c_{ij}}{2}}_{\text{direct connections}} + \underbrace{\sum_{\substack{j\neq i:\\ \ell_{ij}(g)=2}} \delta\frac{1-I_{\{r_{ij}(g)=1\}}\gamma}{2}}_{\text{indirect connections}} + \underbrace{\sum_{\substack{j\neq i,k\neq i:\\ \ell_{jk}(g)=2\wedge ij\in g\wedge ik\in g}} I_{\{r_{jk}(g)=1\}}\delta\gamma}_{\text{intermediation rents}} \tag{1}$$

where $d_i(g)$ denotes the degree of player i in network g and ℓ_{ij} is the length of the shortest path between nodes i and j. Payoffs are the sum of the following components: (i) benefits from

[4] Aside from incorporating Bertrand competition between intermediaries, zero rents in the case of two or more intermediaries can be derived as a prediction of a model of bargaining in networks without replacement in the limit when bargaining frictions disappear. See Siedlarek (2015). Kleinberg et al. (2008) adopt a different approach in which intermediation benefits decay with the number of alternative paths in a more gradual way.

Figure 1: Two Forms of Link Formation: Bilateral and Introduction

(a) Bilateral (b) Introduction

Notes: A dashed line indicates a newly formed link. A solid line indicates a pre-existing link.

direct connections net of link maintenance costs, (ii) benefits from indirect connections net of intermediation rents paid, and (iii) intermediation rents received.

2.2 Link Costs and Introductions

The model permits two different ways of creating new connections. First, under *bilateral link creation* two nodes agree to form a link between themselves. Second, under *link creation by introduction* a third-party intermediary can facilitate a new connection between two nodes that are not connected with each other but are both connected to the intermediary. Such connections have a cost advantage over connections created by regular bilateral link formation.

Bilateral link creation allows any unconnected pair i, j to create a new link between themselves if both agree. See Figure 1a. Links created in this way incur the full cost $\frac{c_{ij}}{2} = \frac{c}{2} \geq 0$ for each node or $c_{ij} = c$ in total. This cost captures, for example, investments in efforts to screen a potential partner and develop and maintain a sufficient level of trust for a functional relationship.

In addition to bilateral link creation, I introduce link formation via introductions. An introduction occurs when a new connection is created between two unconnected nodes i, j that share a common neighbor k and that neighbor acts as intermediary. See Figure 1b for an example in a simple network of three nodes. Just like bilateral link formation, this process creates the link ij, but in contrast to bilateral link creation by i and j independently, it requires the agreement of the introducing node k.

To capture the benefits of leveraging a common partner in the new connection, introductions in the model have a cost advantage over bilateral link creation and create links that cost $c_{ij} = (1-\epsilon)c$, 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.

Allowing for a cost difference between the two types of links requires tracking the way in which a link was created. I therefore partition the set of links g as follows: g_B contains links created by bilateral link formation, g_I those created by introduction. As a partition the sets satisfy $g_B \cup g_I = g$ and $g_B \cap g_I = \emptyset$. For a network to be feasible, any link in g_I has to be supported by a node that is a common neighbor connected by links in g_B, that is, $\forall ij \in g_I \ \exists\, k : ik \in g_B \wedge jk \in g_B$. We denote by G_I the set of all feasible networks $g = (g_B, g_I)$ with introductions that satisfy this condition.

Payoffs for player i for any given network $g = (g_B, g_I)$ can then be written as shown in equation 2.

$$\begin{aligned}\pi_i(g) &\\ = \ &\underbrace{d_i(g)\frac{1-c}{2}}_{\text{direct connections}} + \underbrace{d_i(g_I)\epsilon\frac{c}{2}}_{\text{introductions}} + \underbrace{\sum_{\substack{j\neq i:\\ \ell_{ij}(g)=2}} \delta\frac{1-I_{\{r_{ij}(g)=1\}}\gamma}{2}}_{\text{indirect connections}} + \underbrace{\sum_{\substack{j\neq i,k\neq i:\\ \ell_{jk}(g)=2\wedge ij\in g\wedge ik\in g}} I_{\{r_{jk}(g)=1\}}\delta\gamma}_{\text{intermediation rents}}\end{aligned} \tag{2}$$

3 Efficient Networks with Introductions

Efficiency considerations in the model apply to both the process of network formation and the resulting structure of relationships. The definition of efficiency incorporates both by tracking the mode of link creation through the payoff equation (Eq. 2).

Definition 1. *A network structure $g^* = (g_B^*, g_I^*)$ is efficient among the set of feasible networks G_I if it maximizes the sum of total payoffs across nodes net of the minimum cost incurred in creating it. That is,*

$$g^* = \underset{g\in G_I}{\operatorname{argmax}} \sum_{i\in N} \pi_i(g)$$

We start by considering the cost of link creation. Generally speaking, introductions are more efficient than link creation without introductions due to their cost advantage ϵ. However, the requirement for nodes to share a common neighbor for introductions to be feasible implies some role for non-mediated bilateral link creation. For example, a closed triangle of three nodes requires at least two links created by bilateral link formation. The third link closing the triangle can then be formed by introduction. Without this restriction taking account of the dynamics of the introduction process, only networks with all links in g_I could ever be efficient.

Proposition 2 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 of the links of the final network structure but also of the most efficient way of creating the network through a combination of bilateral link creation and introductions.

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

1. *The* empty network (g^0, g^0) *if*

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

2. *The n-player* star network (g^S, g^0) *if*

$$c < 1 + \frac{n-2}{2}\delta, \text{ and}$$
$$c > \frac{1-\delta}{1-\epsilon}$$

   The star network is formed by $n-1$ links created by bilateral link creation.

3. *The n-player* complete network $(g^S, g^N \setminus g^S)$ *if*

$$c < \frac{1}{1 - \left(1 - \frac{2}{n}\right)\epsilon}, \text{ and}$$
$$c < \frac{1-\delta}{1-\epsilon}$$

   To be feasible the complete network consists of a star network g^S with $n-1$ links created by bilateral link creation with all remaining links created through introductions.

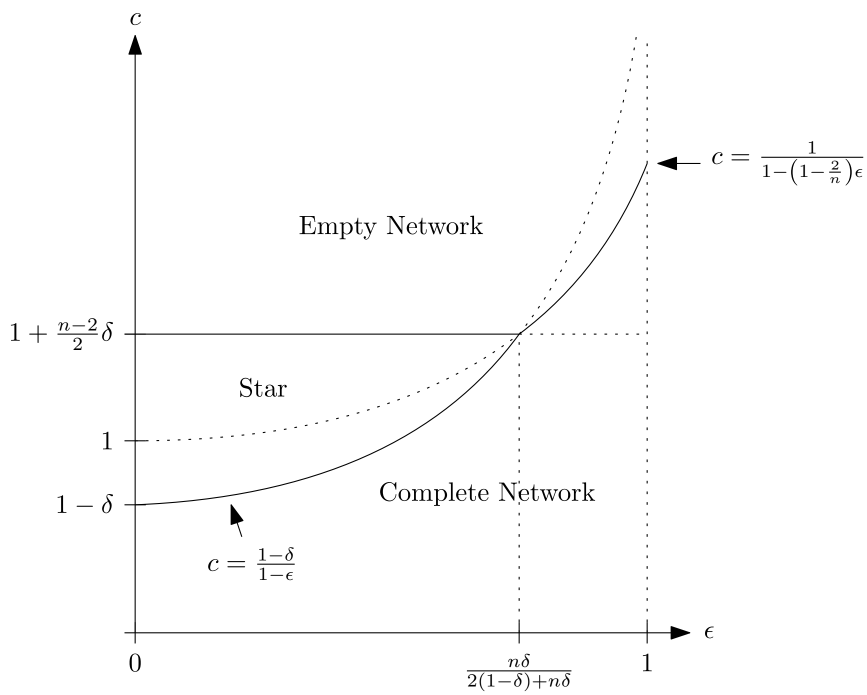
The efficient network structure is unique up to a permutation of players.

The characterization focuses on the relationship between the cost per link c and the efficiency of introductions ϵ. Figure 2 illustrates the parameter regions characterized in Proposition 2. As the cost efficiency of introductions increases, 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 $\epsilon = 0$ introductions do not provide any advantage over bilateral link creation. As a result, the efficient structures and parameter ranges correspond to those in Jackson and Wolinsky (1996). However, as the cost advantage of introductions ϵ increases, differences between Jackson and Wolinsky (1996) and the setting with introductions emerge.

Initially, as ϵ increases from zero, it becomes more advantageous to close open triangles and the cost range for which the star network is efficient decreases, while the range for the complete network increases. Above a certain level of cost efficiency, the star network is not efficient for any level of linking cost anymore and only either the empty network or the complete network remains efficient. The result shows the extent to which the cost advantage of introductions pushes efficient structures toward high density and high levels of clustering. The efficient configuration is

Figure 2: Efficient Network Configurations (ϵ = 1)



Notes: This chart illustrates the efficient network configurations for different parameter ranges. Solid lines separate the areas where different configurations are efficient. Dotted lines are auxiliary and added to facilitate labeling and reading of axis labels.

independent of the intermediation rent parameter γ as it only presents a transfer between players, 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 links from the network.

4 Stable Networks with Introductions

This section presents results on networks that are stable when introductions are available. I build on the pairwise stability concept of Jackson and Wolinsky (1996) by incorporating a suitable extension that allows for introductions and transfers.

The analysis of introductions requires the inclusion of transfers to the introducing player in order to compensate that player for the potential loss of 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 an increase in connection benefits. However, the new link does not generate any 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 needed 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). Their stability concept recognizes unilateral link destruction and bilateral link creation. I extend their setting with an additional stability condition to account for link creation via 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) $\forall\ ij \notin g$,

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

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

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

The first two conditions correspond to those used for networks that are *pairwise stable with transfers*, as analyzed in Bloch and Jackson (2006). Note that destruction of a link in g_B can imply the destruction of additional links in g_I if the destroyed link renders these introductions infeasible. The definition includes a third condition, which requires that in a stable network there be no opportunities for profitable introductions. That is, there cannot be a triplet of nodes that form an open triangle jointly benefiting from adding the missing link, with a lower cost $(1 - \epsilon)$.

Note that the third condition — introduction — allows a deviation by a coalition of three agents. This kind of deviation arguably requires an additional degree of coordination beyond the usual bilateral deviations familiar from Jackson and Wolinsky (1996) and other papers. In this paper, such deviations are permitted if they reflect introductions; that is, the three nodes are already connected on a subnetwork induced by themselves and are forming the missing link. In the context of international trade, Chaney (2014) shows that such a process can successfully explain the dynamics of how trade networks develop over time.[5]

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

[5] See also the discussion of triplet-wise deviations in Section 6.2 of Ambrus and Elliott (2021). They allow deviations by three nodes as long as they create the new links among themselves and thus facilitate risk-sharing among themselves.

symmetry between link creation and link destruction.[6] Note that such transfers do not need to involve a monetary exchange at the time of the introduction. For example, a debt might be incurred in the form of owing a favor to another party in the future (Jackson, Rodriguez-Barraquer, and Tan, 2012).

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

The following section analyzes networks that are stable when introductions are feasible by (i) considering the stability properties of the efficient configurations (Section 4.1) and (ii) characterizing the properties of stable networks in general (Sections 4.2 and 4.3).

4.1 Stability of Efficient Networks

This section derives 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 (g^0, g^0) *is stable if*

$$c \geq 1$$

2. *The* star network (g^S, g^0) *is stable if*

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

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

3. *The* complete network $(g^S, g^N \setminus g^S)$ *is stable if*

$$c \leq \min\left\{\frac{1-\delta}{1-\epsilon}, \frac{1}{1-\left(1-\frac{2}{n}\right)\epsilon}\right\}$$

Table 1: Parameter Ranges for Efficient and Myopically Stable Networks

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

Note: This table lists the parameter conditions under which the empty, star, and complete networks are efficient and myopically stable.

Combining the results derived above, Table 1 lists the parameter ranges for stability next to the corresponding thresholds for efficiency, which are derived in Section 3.

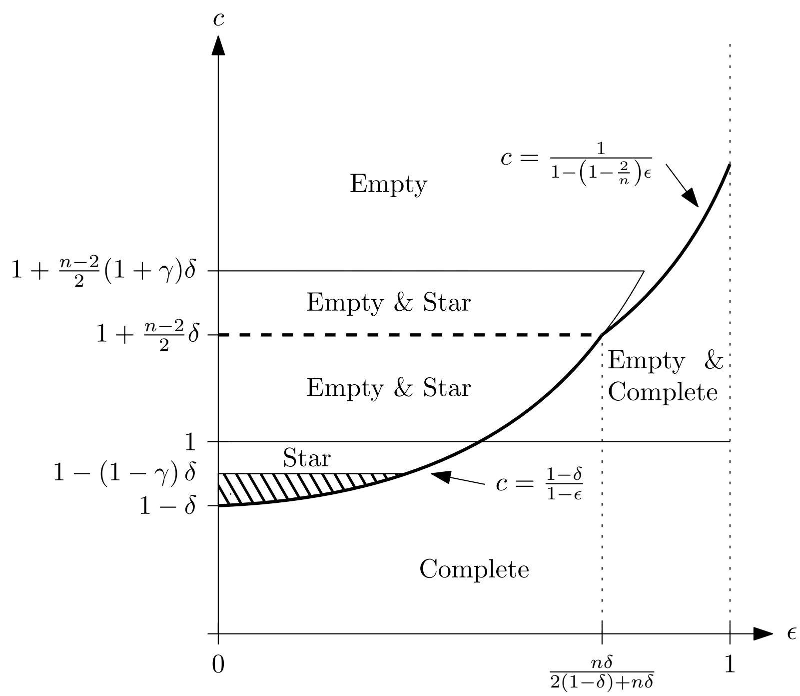
The parameter thresholds for stability are illustrated as solid lines in Figure 3. The chart replicates the efficiency thresholds derived above and shown in Figure 2 to allow answering of the question whether or not the efficient network is stable across the parameter space.

The chart shows that whenever the empty network is efficient, then that network is also myopically stable with introductions. Likewise, whenever the complete network is efficient ($c \leq \min\left\{\frac{1-\delta}{1-\epsilon}, \frac{1}{1-(1-\frac{2}{n})\epsilon}\right\}$), that network is also stable. However, note that while for the complete network the reverse is also true — the complete network is efficient whenever it is stable — this is not the case for the empty or the star network. Both types of network can be stable in areas of the parameter space where they would not be the efficient network.

Of note is the shaded area toward the bottom left of Figure 3. In this area, none of the three efficient structures is stable and, in particular, the star network, which is the efficient network structure in this parameter range, is not stable. Thus, in this area of the parameter space any stable network will necessarily be inefficient.

In addition, Figure 3 shows that there are parameter ranges where multiple networks can be stable. For example, in the center-left of the chart, both the star network and the empty network are stable; in the center-right, both the complete network and the empty network are stable. There may be other networks that are also stable. Multiplicity implies that even if an efficient structure is stable, it does not necessarily follow that this is the network outcome that is achieved. Agents would need to coordinate to achieve the efficient network and inefficient, but stable, network structures can be formed instead.[7]

Figure 3: Myopic Stability of Empty, Star, and Complete Networks



Notes: This chart illustrates the myopic stability of the empty, star, and complete networks. Solid lines separate the areas where different configurations are stable. Dotted lines are auxiliary and added to facilitate labeling and reading of axis labels. The thick lines (dashed and solid) are the solid lines replicated from Figure 2 and separate the parameter ranges for which different network configurations are efficient. The horizontal thick dashed line at $c = 1 + \frac{n-2}{2}\delta$ indicates the boundary between the parameter ranges where the star network (below) and the empty network are stable (above). The shaded area marks the parameter range where none of the three networks is stable. In this area, the star network is efficient but not stable.

The divergence between efficiency and stability is due to externalities and the myopic stability notion employed. Externalities exist where players that form or destroy a link do not experience the full change in payoff that this generates. For example, in a star network, periphery nodes forming a new bilateral link benefit from reducing intermediation rents paid to the center node, but do not account for the fact that this is a loss for the node in the center and thus merely a transfer from the perspective of efficiency. This particular externality is behind the shared area in Figure 3 in which the efficient network is not stable. The misalignment between efficiency and stability is familiar from the literature on strategic network formation such as Jackson and Wolinsky (1996): When considering whether to create or destroy a link, the players involved assess the impact only on their own payoffs and disregard the effect on other players, which can lead to both too many and too few links (Bloch and Jackson, 2007).

[7] I thank an anonymous referee for making this point.

Myopia in the stability notion is a further source of inefficiency. For example, when considering deviations from the empty network, agents only assess the immediate payoff changes generated by the first link created, ignoring possible future deviations creating larger networks. This effect means that the empty network is efficient if $c > 1+\frac{n-2}{2}\delta$, but stable at the lower cutoff $c > 1$, covering parts of the parameter space where other network structures would be efficient. Farsighted stability notions can help overcome this friction (Dutta, Ghosal, and Ray, 2005; Herings, Mauleon, and Vannetelbosch, 2009).

4.2 Stable Networks' Maximum Distance and Connectedness

Next, I briefly 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 steps apart.

Proposition 5. *Let g be a network that is pairwise stable with introductions. Let (i, j) be the pair of nodes with the highest sum of degrees $d_i(g) + d_j(g)$ such that $\ell_{ij} > 3$. Then:*

$$d_i(g) + d_j(g) \leq -\frac{2\left[(1-c)\right]}{\delta(1+\gamma)} \tag{3}$$

The result shows that even in settings where an isolated link is not profitable ($1 - c < 0$), sufficiently high intermediation benefits can exert a force toward limiting the diameter of a network by encouraging higher degree nodes that are far away from each other to connect and benefit from intermediating a number of indirect connections.

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 link costs exceed the direct benefits of a link ($c > 1$). Finally, once link costs are sufficiently small, the result implies that any stable network will be empty or fully connected.

Corollary 6. *If link costs are sufficiently small such that $c < 1 + \delta\frac{1+\gamma}{2}$, any non-empty network that is a pairwise stable network with introductions will be connected.*

4.3 Stable Networks Exhibit Minimum Level of Clustering

This section considers the impact of introductions on the clustering properties of stable networks. Clustering is a measure of the local cliquishness or cohesiveness of a network and measures the extent to which a triplet of nodes i, j, k in which k is connected to both i and j forms a fully connected triangle; that is, there is a link that connects j to k. By definition, wherever introductions take place, they create such closed triangles and thereby contribute to higher clustering. I therefore ask which, if any, possible introductions will remain unused in stable networks. I then extend the analysis to derive a lower bound for local clustering coefficients in stable networks.

Lemma 7 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.

Lemma 7. *Consider a network g such that there is an opportunity for k to introduce i and j. Let $\Delta\pi_{ijk}$ be the change in total payoffs to i, j, k from creating the introduction. Then*

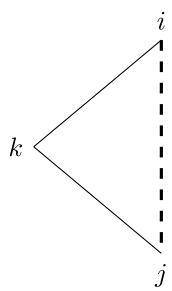
$$\begin{aligned}\Delta\pi_{ijk} =& 1-(1-\epsilon)c-\delta \\ &+\delta \sum_{\substack{u\notin\{i,j,k\},\\ v\in\{i,j\},w\in\{i,j\}:\\ uv\in g\wedge uw\notin g\\ ku\notin g}} \left[I_{\ell_{uw}(g)>2}\frac{1+\gamma}{2} + I_{\{\ell_{uw}(g)=2\wedge r_{uw}(g)=1\}}\frac{\gamma}{2}\right] \\ &-\delta \sum_{\substack{u\notin\{i,j,k\},\\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

The result decomposes the overall effect into several constituents based on the local network structure around the introduction: first, the payoff effect from the *new direct link* between i and j that replaces the indirect connection via k with a direct link (Figure 4a); second, the positive payoff effect from the *new indirect link* between any nodes not directly connected to k and connected to exactly one of the nodes to be introduced i, j, where no indirect link previously existed (Figure 4b); third, the positive payoff effect from such a newly created indirect link where that new link helps *avoid intermediation rents* previously paid to a third node (Figure 4c); and fourth, the negative payoff effect from such a newly created indirect link where that link means that the introducing node k *loses intermediation rents* previously captured (Figure 4d). Note that all but the first of these components relate to nodes that are connected to exactly one of the nodes to be introduced. Lemma 7 shows that this decomposition fully captures the payoff implications for the three nodes involved in an introduction.

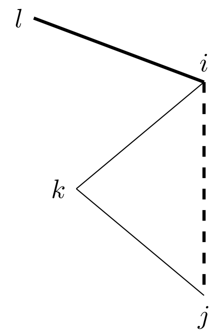
It then follows from Lemma 7 that we can bound the payoff from an introduction from below by focusing on the number of neighbors of i or j that can generate a negative payoff for $\{i, j, k\}$ jointly.

Proposition 8. *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*

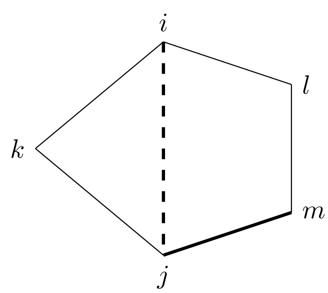
Figure 4: Payoff Implications of an Introduction



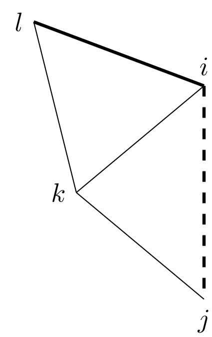
(a) New Link Direct



(b) New Link Indirect (+ve)



(c) Avoided Intermediation Rents (+ve)



(d) Lost Intermediation Rents (-ve)

Notes: This chart illustrates the payoff implications from creating a new link by introduction for the three nodes involved, here i, j and k, as derived in Lemma 7. A dashed line indicates the newly formed link by introduction. A solid line indicates a pre-existing link. *+ve* and *-ve* indicate the sign of $\Delta\pi_{ijk}$, the joint payoff change for nodes i, j, and k, where it is unambigious. Figure 4a illustrates the new direct link between nodes i and j, replacing the indirect connection via node k. Figure 4b shows a new indirect connection of length two between nodes j and l via node i, which previously had a shortest path of length three. Both nodes j and i earn increased payoff from the new connection. Figure 4c shows the avoidance of intermediation rents. Prior to the new link ij being formed, nodes j and l are connected via essential intermediary m, paying intermediation rent. The new link ij creates a second path of length two and thus no more rent is paid, increasing payoffs for j. Figure 4d illustrates the loss of intermediation rents for k. The new link ij creates a second path of length two between j and l, removing the intermediation rents paid to k before. Some of the savings are captured by node l and, thus, lost to the three nodes of interest i, j, and k.

then

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

Intuitively, this result describes a necessary condition for the existence of open introductions in stable networks in the form of a lower bound on μ, the number of nodes that are connected to the introducing node k and exactly one of the nodes to be introduced (i or j). Such nodes generate intermediation payoffs for k that would be partially lost by closing the open triangle (see Figure 4d). Conversely, the result implies that any introduction where this condition is not met will be profitable to conduct. As we will see below, this leads to high degrees of clustering.

In terms of the application to business networks, this result shows that introductions and referrals tend to be profitable for the business involved and can help create networks with closed triangles. Where introductions are feasible but not conducted, they must protect significant intermediation rents for the introducing business, so that it is unprofitable to create the introduction.

Note that for the lower bound μ^* to be positive requires an introduction to be profitable in isolation, that is, in a setting without any neighbor to ijk such that $1-(1-\epsilon)c-\delta > 0$. We will focus on this case in the subsequent discussion. If introductions by themselves are not profitable to the three nodes involved even without any of the additional effects identified above, the results below do not apply. Conditional on isolated introductions being profitable, μ^* increases as intermediation profits γ become less important, because the loss of intermediation benefits to the introducing node is the friction that can prevent introductions from being profitable. In addition, μ^* is increasing in ϵ, the cost advantage of introductions relative to direct bilateral link creation.

Proposition 8 connects any unused introduction opportunity to a minimum number of nodes that form a closed triangle. We can now establish a lower bound on the local clustering coefficient of any node in a myopically stable network based on the underlying parameters of the model. To define terms, let the individual clustering coefficient C_k be the share of triplets formed by node k and two of its neighbors, such that the neighbors themselves are directed.

$$C_k := \frac{|i \neq j \neq k : ki \in g, kj \in g, ij \in g|}{|i \neq j \neq k : ki \in g, kj \in g|}$$

Then the following result holds.

Proposition 9. *Let g be a network that is myopically stable with introductions and let $\lambda^* := \lfloor \frac{\mu^*}{2} \rfloor + 1$.*

For any node $k \in N$ with degree $d_k \geq 2$ the local clustering coefficient C_k is bounded below by $\underline{C}(d_k)$

such that

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

Figure 5 illustrates Proposition 9 for a simple example of a central node with degree 12.[8] At the bound, the local neighborhood around node k forms a subnetwork that maximizes the number of cliques with at least λ^* nodes each. In such a subnetwork, each node is connected to every other node belonging to the same clique, but not connected to any of the neighbors belonging to other cliques. As a consequence, any open triangle of links around k will involve two nodes that are neighbors of k and that each have at least μ^* neighbors that they share with k and that they do not share with each other. By Proposition 8, any such open triangles between cliques in the neighborhood of k is myopically stable against introductions that would create a closed triangle.

The bound can be reached exactly if μ^* is even and d_k is divisible by λ^* so that all neighbors of k are part of a clique of exactly size λ^*. If μ^* is odd, then stability requires that at most one clique can be of exactly size λ^*, and all remaining cliques have to be of size $\lambda^* + 1$. Furthermore, if d_k is not divisible by λ^*, the remainder of the nodes will have to be distributed across the cliques to ensure stability.

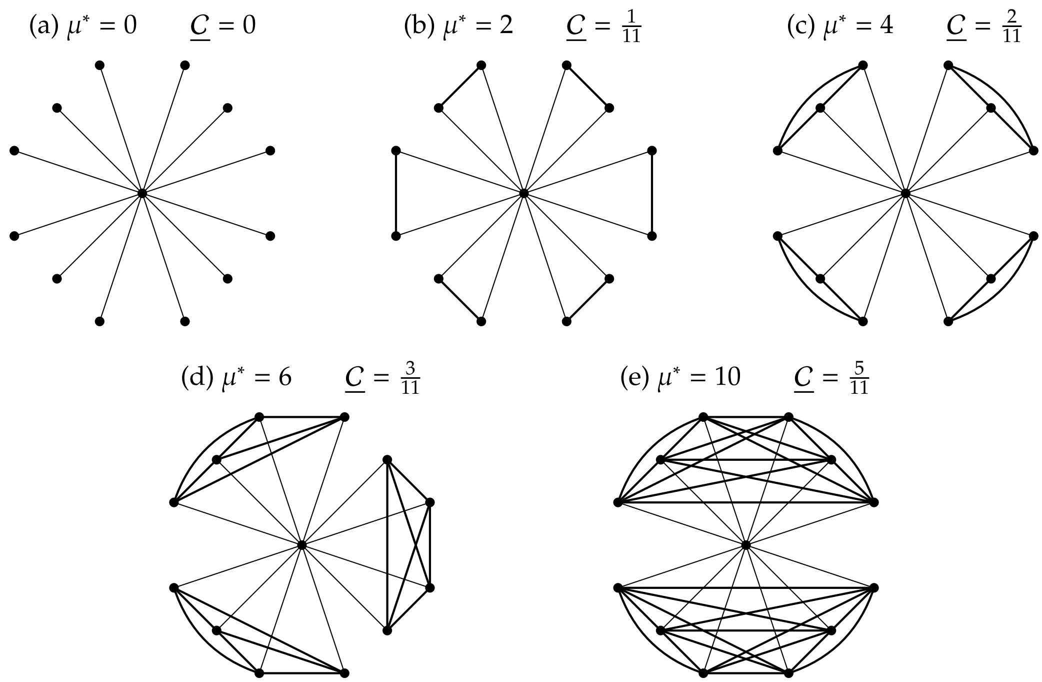
The higher μ^*, the higher the lower bound on local clustering $\underline{C}(d_k)$. That is, as introductions become more profitable, minimum clustering increases. Depending on the underlying parameters, minimum clustering might be zero as in Figure 5(a) or relatively high as in Figure 5(e). In addition, as the degree of k increases, the lower bound on clustering converges toward zero.

The driving force behind the minimum clustering in this result is the introduction mechanism and not the presence of intermediation rents alone. This mechanism 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 for links to be formed,[9] whereas in Proposition 9 higher intermediation rents per connection are associated with a lower μ^*. This implies a lower minimum clustering coefficient and consequently higher realized intermediation payoffs for the central node.

[8] Möhlmeier, Rusinowska, and Tanimura (2016) generate similar windmill-shaped networks or friendship graphs from very different mechanics in their model of network formation with both positive and negative externalities. In Ambrus and Elliott (2021), friendship graphs are efficient structures for risk-sharing if common friends are required to enforce risk-sharing agreements.

[9] In their paper the authors hold the share of surplus captured by essential intermediaries fixed.

Figure 5: Example of Minimum Clustering with Cliques of Even Size $\frac{\mu^*}{2} + 1$



Notes: This chart illustrates node allocation with minimum clustering as per Proposition 9 for different values of μ^*. To achieve the lower bound on clustering, nodes form cliques around the central node.

4.4 Stable Networks Protect Structural Holes

The previous sections showed that stable networks in the link formation model with introductions tend to be connected and show high degrees of clustering. In this section, I will focus on incentives for highly connected nodes to protect structural holes. The concept of structural holes as described in Burt (1992, 2007) refers to "a gap between two individuals" (Burt, 1992) with positive implications for a third node that fills the gap and facilitates brokerage. In the context of the connections model with intermediation analyzed in this paper, brokerage opportunities are created when a node intermediates an indirect connection between two other nodes and acts as the unique such intermediary.

A useful measure of how important an intermediary node is for the connection between pairs of nodes in the network is called betweenness (Jackson, 2008, see, e.g., p. 39). The betweenness centrality of a node is the average fraction of shortest paths between all other pairs of nodes in the network on which the node acts as an intermediary. For the purposes of this model, we employ a modified *local* betweenness measure that reflects the assumption represented in the model's payoff function. The measure restricts attention to paths of up to length two and in addition only values connections where the node is the only intermediary. It thus captures the share of connections for which the intermediary extracts an intermediation rent. Formally, let $\mathcal{B}_k$ be the share of paths between pairs of nodes in the neighborhood of node k for which node k is essential.

$$\mathcal{B}_k := \frac{|i \neq j \neq k : ki \in g, kj \in g, ij \notin g, r_{ij}(g) = 1|}{|i \neq j \neq k : ki \in g, kj \in g|}$$

The measure $\mathcal{B}_k$ captures the share of connections involving neighbors of k for which it extracts an intermediation rent for brokering the connection between them. It serves as a measure of structural holes around k for which it profitably acts as a broker. Then the following result holds.

Proposition 10. *Let g be a network that is myopically stable with introductions and let $1-(1-\epsilon)c-\delta > 0$ so that $\mu^* > 0$. Then,*

1. *For an open triangle of nodes i, j, and k such that i and j are connected to k and not directly connected themselves, node k is the essential intermediary between i and j.*

2. *Furthermore,*

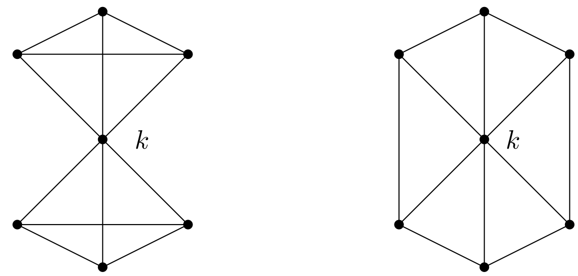
$$\mathcal{B}_k = 1 - C_k$$

The result follows from Lemma 7, which shows that the payoff change from an introduction is strictly positive unless it threatens a sufficient number of indirect connections for which the introducing agent is essential. Thus, in stable networks every pair of nodes in the neighborhood of another node has either a direct connection or an indirect connection for which the third node acts as essential intermediary. Importantly, there are no indirect connections for which the intermediary is not essential. This partition into triangles that are either closed or open and involve essential intermediaries causes the direct relationship between the clustering coefficient and the betweenness measure. In an application to business networks, this result suggests that relationship networks can exhibit nodes acting as bridges connecting otherwise unconnected parts of the network and earning returns on exploiting their position as in Burt (1992). Evidence for such returns has been documented in different settings, for example, in networks of R&D collaboration (Ahuja, 2000; Owen-Smith and Powell, 2004).

The result implies that given a level of clustering, the number of connections for which intermediation rents are captured is maximized in stable networks. Figure 6 illustrates this point by comparing two neighborhoods around node k with the same level of clustering. In Figure 6(a), node k acts as essential intermediary for every open triangle, with k bridging the maximum number of structural holes. In Figure 6(b), while the number of links among neighbors is the same, there are no open triangles for which k earns intermediation rents and thus no structural holes. Proposition 10 shows that the second structure cannot be stable in the model with introductions. Indeed, given the clustering coefficient, only the first neighborhood can be stable.

Figure 6: Example of Neighborhood With and Without Structural Holes

(a) $C_k = \frac{6}{15}$ $\mathcal{B}_k = \frac{9}{15}$ (b) $C_k = \frac{6}{15}$ $\mathcal{B}_k = \frac{0}{15}$



Notes: This chart shows a simple example of a neighborhood around node k with a given level of clustering. The subchart (a) shows a high level of structural holes as measured using the betweenness of node k, while the second subchart (b) has zero betweenness with the same level of clustering. The point the chart makes is that the presence of structural holes is a network property that is independent of clustering. Proposition 10 shows that in the model with introductions only the first configuration in which structural holes are protected is stable.

4.5 Comparison with Jackson and Rogers (2007)

The preceding analysis explains how the combination of introductions and intermediation rents can result in networks with high degrees of clustering (Section 4.3) while protecting structural holes (Section 4.4). The finding of high degrees of clustering is not a unique prediction. Other models of network formation are also able to generate high degrees of clustering. Closely related to the present paper, Jackson and Rogers (2007) show in a random graph model of network formation how a process of linking to friends-of-friends can result in significant clustering. They find that a suitable combination of fully random and friends-of-friends linking can result in average clustering coefficients that match well those of a number of example networks from different domains, including the World Wide Web, academic citations, coauthorship, high school romance, and others.

The key conceptual difference between Jackson and Rogers (2007) and the present paper is that the former is a model of network formation based on a stochastic process of network growth, whereas the latter adopts a strategic approach that explicitly models the payoffs to individuals from any given network structure. Despite the differences between the two approaches, both provide useful insights and are complementary in nature. Relative to the random network formation approach, the strategic approach permits insights into the incentives for link formation resulting from an underlying structural description of costs and benefits. The strategic approach also permits a direct characterization of efficient network structures, whereas in random graph models, efficiency can be inferred only indirectly.

In addition, by definition random graph models of network formation generate a random network, that is, a distribution of networks, which can be characterized to show certain stochastic properties, like the mean degree or the probability of a giant component, often using mean-field approximation. For example, Jackson and Rogers (2007) are able to derive their results on network clustering by characterizing the expected share of transitive triples or the clustering coefficient of the average node (Jackson and Rogers, 2007, Theorem 2). By contrast, the clustering characterization of Proposition 9 in this paper applies to each node and does not describe an average over the whole network.

The key contribution of this paper beyond the model in Jackson and Rogers (2007) relates to predictions concerning network bridges and structural holes. As shown in Proposition 10, the model in this paper explicitly connects both clustering and the existence of structural holes bridging otherwise unconnected parts of the network. As a result, the model can combine densely connected subnetworks with structural holes generating substantial intermediation rents for the intermediary nodes. As illustrated in Figure 6, this combination of clustering and structural holes is not a given. A key insight from the model with introduction is that it provides the incentives for intermediaries to exploit introductions where profitable, but otherwise protect their brokerage opportunities.

Jackson and Rogers (2007) do not target structural holes in their analysis and their model does not provide results regarding the betweenness or other centrality measures that could proxy for the structural holes idea. I argue that the random nature of the linking process suggests that we can expect fewer structural holes bridged by essential intermediaries than in the present model. The random linking process in Jackson and Rogers (2007) achieves high clustering through a link formation mechanism combining fully random meetings with friends-of-friends meetings. When a link is created via the friends-of-friends mechanism, it automatically generates a closed triangle, similar to the introductions in the present model. However, that link is placed randomly among the available friends and randomly among their neighbors. Thus, this process does not preferentially create closed triangles where the intermediary node is not essential, nor, vice versa, does it preferentially protect open triangles where they support significant intermediation rents. In terms of Figure 6, we can therefore expect an outcome that is on average somewhere in between the two extreme cases shown. This is in contrast to the model in this paper, which includes incentives for nodes to create the structure in Figure 6(a) with structural holes and high intermediation rents.

5 Conclusion

This paper analyzes network formation with introductions in a connections model with intermediation rents. Specifically, the paper considers a model of network formation in which players that are unconnected but share a direct neighbor can be "introduced" by that neighbor. Links created by an introduction offer payoff advantages over links created bilaterally, but they can threaten intermediation rents for the introducing node. The trade-off between the costs and benefits limits the extent to which nodes take advantage of introductions to create highly clustered networks via triadic closure. The paper shows first that efficient network structures are either empty, star, or complete networks, as in the baseline connections model of Jackson and Wolinsky (1996). The paper then conducts a stability analysis using a suitable adapted notion of pairwise stability with introductions, to study the incentives involved in introductions and the impact on network outcomes. As expected there can be a gap between efficiency and stability due to the externalities involved in network formation.

The myopic stability analysis of the model with introductions finds that stable networks tend to be connected in a single component, which is consistent with the short distances often found in the data. In addition, it shows how the possibility of introductions creates a lower bound on the level of clustering, providing a plausible mechanism for the observed high levels of clustering in many real-world networks. Finally, it shows that where clustering is not complete and there are open triangles in stable networks, these protect intermediation rents generated by structural holes. Jointly, under certain parameters, stable networks can exhibit both small world properties (Watts and Strogatz, 1998) and bridging agents protecting structural holes (Burt, 1992).

The combination of clustering with the existence of bridging nodes that confer substantial intermediation returns to the nodes operating them as illustrated in Figure 5 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 $1 - c - \delta$ plus changes to intermediation returns. Likewise, if there are no intermediation rents ($\gamma = 0$), then by Lemma 7 there is no loss for introducing players. This implies that the lower bound of closed triangles required to support any open triangle in the sense of Proposition 8 goes to infinity, resulting in all feasible introductions taking place. There would be complete clustering and no structural holes remaining.

Note that while the tendency for introductions to lead to connected components and increased clustering captures features of real-world organization and business networks (Kogut and Walker, 2001), the analysis also highlights the limits to this process. Specifically, the process of link creation by introduction stops with bridge agents that are essential and connect otherwise disconnected parts of the network. If these individuals earn sufficiently high intermediation rents, then their incentives will be to prevent connections forming between the two parts on either side of the bridge to protect these rents. The stability analysis thus provides a plausible explanation for network connectedness, clustering and the persistence of returns to bridging agents that connect across structural holes, which offers a novel contribution to the literature.

The paper offers insights into how plausible local dynamics in link creation can lead to clustering, as groups of nodes that are already connected to some degree form additional beneficial connections among themselves. 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 underlines how the involvement of additional nearby players with suitable transfers in link creation may help to deal with externalities that tend to push equilibrium outcomes toward over- or under-connected networks, as in Bloch and Jackson (2007). Future research may consider how these insights generalize to, for example, allowing coalitions of more than three players to form connected cliques.

References

Ahuja, Gautam (2000). "Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study." *Administrative Science Quarterly* 45.3 (Sept. 2000), p. 425. DOI: 10.2307/2667105. JSTOR: 2667105.

Ambrus, Attila and **Matt Elliott** (2021). "Investments in Social Ties, Risk Sharing, and Inequality." *The Review of Economic Studies* 88.4 (July 17, 2021), pp. 1624–1664. DOI: 10.1093/restud/rdaa073.

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.

Blume, Lawrence E., **David Easley**, **Jon Kleinberg**, and **Éva Tardos** (2009). "Trading Networks with Price-Setting Agents." *Games and Economic Behavior* 67.1 (Sept. 2009), pp. 36–50. DOI: 10.1016/j.geb.2008.12.002.

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.

Buskens, Vincent and **Arnout van de Rijt** (2008). "Dynamics of Networks If Everyone Strives for Structural Holes." *American Journal of Sociology* 114.2 (Sept. 1, 2008), pp. 371–407. DOI: 10.1086/590674.

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

Chaney, Thomas (2014). "The Network Structure of International Trade." *American Economic Review* 104.11 (Nov. 1, 2014), pp. 3600–3634. DOI: 10.1257/aer.104.11.3600.

Choi, Syngjoo, **Andrea Galeotti**, and **Sanjeev Goyal** (2017). "Trading in Networks: Theory and Experiments." *Journal of the European Economic Association* 15.4 (Aug. 2017), pp. 784–817. DOI: 10.1093/jeea/jvw016.

Coleman, James S. (1988). "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (Jan. 1988), S95–S120. DOI: 10.1086/228943.

Condorelli, Daniele, **Andrea Galeotti**, and **Ludovic Renou** (2017). "Bilateral Trading in Networks." *The Review of Economic Studies* 84.1 (Jan. 2017), pp. 82–105. DOI: 10.1093/restud/rdw034.

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.

Dutta, **Bhaskar**, **Sayantan Ghosal**, and **Debraj Ray** (2005). "Farsighted Network Formation." *Journal of Economic Theory* 122.2 (June 2005), pp. 143–164. DOI: 10.1016/j.jet.2004.05.001.

Easley, **David** and **Jon Kleinberg** (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, July 19, 2010. 745 pp. Google Books: atfCl2agdi8C.

Farboodi, **Maryam**, **Gregor Jarosch**, and **Guido Menzio** (2017). *Intermediation as Rent Extraction*. 24171. National Bureau of Economic Research, Dec. 2017. DOI: 10.3386/w24171.

Farboodi, **Maryam**, **Gregor Jarosch**, and **Robert Shimer** (2017). *The Emergence of Market Structure*. 23234. Cambridge, MA: National Bureau of Economic Research, Mar. 2017. DOI: 10.3386/w23234.

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.

Herings, **P. Jean-Jacques**, **Ana Mauleon**, and **Vincent Vannetelbosch** (2009). "Farsightedly Stable Networks." *Games and Economic Behavior* 67.2 (Nov. 2009), pp. 526–541. DOI: 10.1016/j.geb.2008.12.009.

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

Jackson, **Matthew O.**, **Tomas Rodriguez-Barraquer**, and **Xu Tan** (2012). "Social Capital and Social Quilts: Network Patterns of Favor Exchange." *American Economic Review* 102.5 (Aug. 1, 2012), pp. 1857–1897. DOI: 10.1257/aer.102.5.1857.

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.** 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.

Kogut, **Bruce** and **Gordon Walker** (2001). "The Small World of Germany and the Durability of National Networks." *American Sociological Review* 66.3 (June 2001), p. 317. DOI: 10.2307/3088882. JSTOR: 3088882.

Kossinets, **Gueorgi** and **Duncan J. Watts** (2006). "Empirical Analysis of an Evolving Social Network." *Science* 311.5757 (Jan. 6, 2006), pp. 88–90. DOI: 10.1126/science.1116869.

Manea, **Mihai** (2018). "Intermediation and Resale in Networks." *Journal of Political Economy* 126.3 (June 2018), pp. 1250–1301. DOI: 10.1086/697205.

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.

Mehra, **Ajay**, **Martin Kilduff**, and **Daniel J. Brass** (2001). "The Social Networks of High and Low Self-Monitors: Implications for Workplace Performance." *Administrative Science Quarterly* 46.1 (Mar. 1, 2001), pp. 121–146. DOI: 10.2307/2667127.

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.

Owen-Smith, **Jason** and **Walter W. Powell** (2004). "Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community." *Organization Science* 15.1 (Feb. 1, 2004), pp. 5–21. DOI: 10.1287/orsc.1030.0054.

Podolny, **Joel M.** and **James N. Baron** (1997). "Resources and Relationships: Social Networks and Mobility in the Workplace." *American Sociological Review* 62.5 (Oct. 1997), p. 673. DOI: 10.2307/2657354. JSTOR: 2657354.

Priazhkina, **Sofia** and **Frank H. Page** (2018). "Sharing Market Access in Buyer–Seller Networks." *Journal of Economic Theory* 175 (May 2018), pp. 415–446. DOI: 10.1016/j.jet.2018.01.017.

Renou, **Ludovic** and **Tristan Tomala** (2012). "Mechanism Design and Communication Networks." *Theoretical Economics* 7.3 (Sept. 2012), pp. 489–533. DOI: 10.3982/TE921.

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.

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 bilateral link creation and introduction, which requires a discussion of link costs. For any component in a given network g, Lemma 11 provides a lower bound on the number of links that need to be created bilaterally. The proof is immediate and omitted here.

Lemma 11. *The minimum number of links created bilaterally to create a component of k nodes is $k-1$. The minimum cost to form a component of k nodes with $m \geq k-1$ links is $mc - (m-(k-1))\,\epsilon c$.*

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

Case 1: $c > \frac{1-\delta}{1-\epsilon}$ First, consider the case with $c > \frac{1-\delta}{1-\epsilon}$. I will argue that the star is the efficient configuration to connect k nodes in this case. A star network of k nodes incurs link costs of exactly $(k-1)c$ and generates a net benefit after costs of link formation of:

$$(k-1)\,[1-c] + \frac{(k-1)(k-2)}{2}\delta \tag{4}$$

Now, any other configuration connecting k nodes with $m \geq k-1$ links will generate a net benefit of at most

$$\underbrace{m\,[1-c]}_{\text{direct connections}} + \underbrace{\left[\frac{(k-1)(k-2)}{2} - (m-(k-1))\right]\delta}_{\text{indirect connections}} + \underbrace{[m-(k-1)]\,\epsilon c}_{\text{cost savings from introductions}} \tag{5}$$

The first component in equation 5 represents the net benefits from direct links and the second component reflects the upper bound of the benefits that can be derived from any indirectly connected nodes, assuming that all nodes are at most a distance of two steps from each other. The third component reflects the possible cost savings from using the maximum feasible number of links created through introduction $m-(k-1)$ as per Lemma 11. Subtracting equation 4 from 5 and rearranging yields the maximum possible change in surplus from connecting k nodes with

$m \geq k-1$ links:

$$(m-(k-1))\,[1-c]-(m-(k-1))\,\delta+[m-(k-1)]\,\epsilon c \tag{6}$$

$$=(m-(k-1))\,[1-\delta-(1-\epsilon)c] \tag{7}$$

The first half of this expression is positive as $m \geq k-1$. The second half of the expression is strictly negative as $c > \frac{1-\delta}{1-\epsilon}$. It follows that adding links to the star strictly decreases surplus, unless $m = k-1$. It is thus established that if $c > \frac{1-\delta}{1-\epsilon}$ any efficient network consists of star components and isolated nodes.

Next, we further restrict the set of candidate efficient networks for $c > \frac{1-\delta}{1-\epsilon}$ by establishing that there is either a single star component including all nodes or an empty network.

Assume a candidate network consisting of two star components with $k_1 \geq 1$ and $k_2 \geq 2$ nodes, respectively, each yielding positive surplus. This covers the case of two stars as well as the case of one star and one isolated node. Then the total payoff is:

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

Reconfiguring the nodes into a single star yields:

$$(k_1+k_2-1)\left[1-c+(k_1+k_2-2)\frac{\delta}{2}\right] \tag{9}$$

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

$$[1-c]+(2k_1k_2-2)\frac{\delta}{2} \tag{10}$$

which is strictly positive if each separate star yields positive surplus as $2k_1k_2 > k_1$ and $2k_1k_2 > k_2$. It follows that the surplus generated by the network strictly increases by combining a star component with another star component or with any isolated node. Thus, assuming that a star generates positive surplus, the network with more than one star yields strictly less surplus than a network in which the nodes involved are combined into a single star consisting of n nodes.

The treatment of the case $c > \frac{1-\delta}{1-\epsilon}$ is concluded by comparing payoffs of a single star involving n nodes and the empty network. As the latter derives zero surplus, the star network is the unique efficient network if:

$$(n-1)\,(1-c)+\frac{(n-1)(n-2)}{2}\delta > 0 \tag{11}$$

which reduces to:

$$c < 1 + \frac{(n-2)}{2}\delta \tag{12}$$

Case 2: $\boldsymbol{c \leq \frac{1-\delta}{1-\epsilon}}$ Next consider the case with $c \leq \frac{1-\delta}{1-\epsilon}$. In this case, adding a link by introduction always weakly increases surplus; thus, the star can never be the efficient network.

As per Lemma 11, for any component of k nodes the minimum number of links created bilaterally is $k-1$. As every introduction is profitable, the efficient network maximizes the number of introductions. Given $k-1$ links created bilaterally, the component that maximizes the number of feasible introductions is the fully connected component. This component of size k is formed with $k-1$ links created bilaterally to form a star and all remaining links formed by introduction.

Next, an argument analogous to that for two stars used above shows that in any case where there are more than one fully connected components (or one such component and a singleton node), surplus increases in a single connected component. Thus, if $c \leq \frac{1-\delta}{1-\epsilon}$, the efficient network is either complete or empty. To decide which is the efficient structure, we compare the surplus generated by each. The empty network generates a total surplus of zero. The complete network with n nodes consists of $\frac{n(n-1)}{2}$ links in total, of which at least $n-1$ links are created bilaterally. The surplus in the complete network is then:

$$\underbrace{(n-1)(1-c)}_{\text{bilateral links}} + \underbrace{\frac{(n-1)(n-2)}{2}\left[1-(1-\epsilon)c\right]}_{\text{links from introductions}} > 0 \tag{13}$$

$$\Leftrightarrow \quad 2(1-c) + (n-2)\left[1-(1-\epsilon)c\right] > 0 \tag{14}$$

$$\Leftrightarrow \quad n(1-c) + (n-2)\epsilon c > 0 \tag{15}$$

$$\Leftrightarrow \quad (1-c) + \frac{(n-2)}{n}\epsilon c > 0 \tag{16}$$

$$\Leftrightarrow \quad 1 > \left[1 - \frac{(n-2)}{n}\epsilon\right] c \tag{17}$$

$$\Leftrightarrow \quad \frac{1}{1-\left(1-\frac{2}{n}\right)\epsilon} > c \tag{18}$$

Surplus from the complete network is thus strictly higher than surplus from the empty network if

$$c < \frac{1}{1-\left(1-\frac{2}{n}\right)\epsilon} \tag{19}$$

□

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

This section derives the parameter conditions for the stability of the three network configurations that can be efficient, including the empty network g^0, g^0, the star network g^S, g^0, and the complete network $g^S, g^N \setminus g^S$.

Empty Network In the empty network g^0, g^0, 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:

$$c \geq 1 \tag{20}$$

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

Star Network In the star network g^S, g^0, 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 bilaterally and through introductions:

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

$$c \leq 1 + \frac{n-2}{2}(1+\gamma)\delta \tag{21}$$

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

$$c \geq 1 - (1-\gamma)\delta \tag{22}$$

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

$$c \geq \frac{1-\delta}{1-\epsilon} \tag{23}$$

Complete Network In the complete network $g^S, g^N \setminus g^S$ all possible links are in place. Thus, we only need to consider deviations by link destruction. There are two possible types of links to consider, depending on whether they involve two peripheral nodes of the intermediate star network g^S or only one such node and the center of g^S.

(a) A link involving two peripheral nodes was created by introduction. Destruction of such a link replaces a direct connection created by introduction with an indirect connection that is intermediated by the remaining nodes connected to the two nodes breaking their link. For $n > 3$, there are at least two such intermediaries and the two players destroying the link jointly capture δ. The complete network is thus stable against destruction of this link if:

$$1-(1-\epsilon)\,c \geq \delta \tag{24}$$

$$c \leq \frac{1-\delta}{1-\epsilon} \tag{25}$$

(b) A link involving one peripheral node and the center of the g^S was created bilaterally. Destruction of such a link will disconnect the peripheral node from all other nodes in the network as it removes the only link created bilaterally and by implication all links created by introduction involving this peripheral node. Payoffs for the peripheral node thus go to zero. For the center node, destruction of the link results in a loss of payoffs by $\frac{(1-c)}{2}$. The complete network is thus stable against destruction of this link if:

$$\frac{(1-c)}{2}+\frac{(1-c)}{2}+(n-2)\,\frac{1-(1-\epsilon)\,c}{2} \geq 0 \tag{26}$$

$$2\,(1-c)+(n-2)\,[1-(1-\epsilon)\,c] \geq 0 \tag{27}$$

$$n-nc+(n-2)\,\epsilon c \geq 0 \tag{28}$$

$$c \leq \frac{1}{1-\left(1-\frac{2}{n}\right)\epsilon} \tag{29}$$

A.3 Proof of Proposition 5

Proof. The proof is by contradiction. Assume that network g is myopically pairwise stable with introductions. Further, assume that there are two nodes, i and j, violating the condition in equation 3 such that $\ell_{ij} > 3$ and $d_i(g) + d_j(g) > -\frac{2(1-c)}{\delta(1+\gamma)}$. I will show that in that case a new connection between i and j created bilaterally would be profitable and thus the network cannot be stable.

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

The total change payoff from the connection ij to nodes i and j is thus

$$\Delta\pi_{ij} = (1-c)+\left[d_i(g)+d_j(g)\right]\delta\frac{1+\gamma}{2} \tag{30}$$

$$> (1-c)-(1-c) = 0 \tag{31}$$

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

□

A.4 Proof of Lemma 7

Proof. The introduction of i and j by k creates a new link ij. This proof derives the decomposition of the effect of the new link ij on the joint payoffs of nodes (i, j, and k) denoted by

$$\begin{aligned}\Delta\pi_{ijk} &= \Delta\pi_i + \Delta\pi_j + \Delta\pi_k \\ &= [\pi_i(g+ij) - \pi_i(g)] + \left[\pi_j(g+ij) - \pi_j(g)\right] + [\pi_k(g+ij) - \pi_k(g)]\end{aligned}$$

Recall the payoff function in equation 2, which separates direct connections, introductions, indirect connections, and intermediation rents

$$\pi_i(g) = \underbrace{d_i(g)\frac{1-c}{2}}_{\text{direct connections}} + \underbrace{d_i(g_I)\epsilon\frac{c}{2}}_{\text{introductions}} + \underbrace{\sum_{\substack{j\neq i:\\ \ell_{ij}=2}} \delta\frac{1-I_{\{r_{ij}=1\}}\gamma}{2}}_{\text{indirect connections}} + \underbrace{\sum_{\substack{j\neq i,k\neq i:\\ \ell_{jk}=2\wedge ij\in g\wedge ik\in g}} I_{\{r_{jk}=1\}}\delta\gamma}_{\text{intermediation rents}}$$

We decompose $\Delta\pi_i$ in a similar way and consider each component for each node $p \in \{i, j, k\}$ in turn.

$$\Delta\pi_p = \Delta\pi_p^{\text{Direct}} + \Delta\pi_p^{\text{Introductions}} + \Delta\pi_p^{\text{Indirect}} + \Delta\pi_p^{\text{Rents}}$$

Effect on Direct Connections and Introductions Consider a node $p \in \{i, j, k\}$

$$\begin{aligned}\Delta\pi_p^{\text{Direct}} &= d_p(g+ij)\frac{1-c}{2} - d_p(g)\frac{1-c}{2} \\ &= \left[d_p(g+ij) - d_p(g)\right]\frac{1-c}{2} \\ &= \begin{cases}\frac{1-c}{2} & \text{if } p\in\{i,j\} \\ 0 & \text{if } p \in \{k\}\end{cases}\end{aligned}$$

It then follows that

$$\Delta\pi_{ijk}^{\text{Direct}} = 1 - c$$

Effect on Cost Benefits from Introductions Consider a node $p \in \{i, j, k\}$

$$\begin{aligned}\Delta\pi_p^{\text{Introductions}} =& d_p(g_I + ij)\epsilon\frac{c}{2} - d_p(g_I)\epsilon\frac{c}{2} \\ =& \left[d_p(g + ij) - d_p(g)\right]\epsilon\frac{c}{2} \\ =& \begin{cases} \epsilon\frac{c}{2} & \text{if } p \in \{i, j\} \\ 0 & \text{if } p \in \{k\} \end{cases}\end{aligned}$$

It then follows that

$$\Delta\pi_{ijk}^{\text{Introductions}} = \epsilon c$$

Effect on Indirect Connections First, given that $\ell_{ij}(g) = 2$, the new link ij does not create any new paths in $g + ij$ between nodes that are unconnected in g. Furthermore, at most the new link can reduce the length of the shortest path between any pair of nodes v, w by one step. Thus, either $\ell_{vw}(g + ij) = \ell_{vw}(g)$, or $\ell_{vw}(g + ij) = \ell_{vw}(g) - 1$. Second, for connections between a pair v, w with shortest path of the same length in g and $g + ij$, the new link ij can open an alternative path. If the shortest path between v, w is of length two, then at most one additional path is created. Thus, for v, w with $\ell_{vw}(g + ij) = \ell_{vw}(g) = 2$, we have either $r_{vw}(g + ij) = r_{vw}(g)$, or $r_{vw}(g + ij) = r_{vw}(g) + 1$. Third, for node i any new connections of length two created by link ij will be to nodes that are already connected to j in g, and vice versa for node j. Fourth, in such cases where the length of the shortest path in g was greater than two and a new connection of length two is created, node j will be the unique intermediary for this new connection.

We can then write down the effect on indirect payoffs. We can restrict attention to node $p \in \{i, j\}$. Payoffs from indirect connections of node k are unaffected by the new link ij. Node k is already connected to both i and j via a direct link in g.

$$
\begin{aligned}
\Delta\pi_p^{\text{Indirect}} &= \sum_{\substack{u\neq p:\\ \ell_{pu}(g+ij)=2}} \delta\frac{1-I_{\{r_{pu}(g+ij)=1\}}\gamma}{2} - \sum_{\substack{u\neq p:\\ \ell_{pu}(g)=2}} \delta\frac{1-I_{\{r_{pu}(g)=1\}}\gamma}{2} \\
&= \sum_{\substack{u\neq p:\\ \ell_{pu}(g+ij)=2\\ \wedge\ell_{pu}(g)=3}} \delta\frac{1-I_{\{r_{pu}(g+ij)=1\}}\gamma}{2} + \sum_{\substack{u\neq p:\\ \ell_{pu}(g+ij)=2\\ \wedge\ell_{pu}(g)=2\\ \wedge r_{pu}(g+ij)=2\\ \wedge r_{pu}(g)=1}} \delta\frac{\gamma}{2} - \sum_{\substack{u\neq p:\\ \ell_{pu}(g+ij)=1\\ \wedge\ell_{pu}(g)=2}} \delta\frac{1-I_{\{r_{pu}(g)=1\}}\gamma}{2} \\
&= \sum_{\substack{u\notin\{i,j,k\},\\ v\in\{i,j\}:\\ up\notin g\wedge uv\in g}} I_{\ell_{pu}(g)=3}\delta\frac{1-\gamma}{2} + \sum_{\substack{u\notin\{i,j,k\},\\ v\in\{i,j\}:\\ uv\in g\wedge up\notin g}} I_{\{\ell_{up}(g)=2\wedge r_{up}(g)=1\}}\delta\frac{\gamma}{2} - \delta\left[\frac{1}{2} - \frac{I_{\{r_{ij}(g)=1\}}\gamma}{2}\right] \\
&=\delta \sum_{\substack{u\notin\{i,j,k\},\\ v\in\{i,j\}:\\ up\notin g\wedge uv\in g}} \left[I_{\ell_{up}(g)>2}\frac{1-\gamma}{2} + I_{\{\ell_{up}(g)=2\wedge r_{up}(g)=1\}}\frac{\gamma}{2}\right] - \delta\left[\frac{1}{2} - \frac{I_{\{r_{ij}(g)=1\}}\gamma}{2}\right]
\end{aligned}
$$

Thus, the joint payoff effect from indirect links is

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

Effect on Intermediation Rents Earned The observations from the previous section on indirect connections still apply: The addition of the link ij either decreases the length of a shortest path by one or leaves it unchanged and either increases the number of intermediaries on a given connection of length two by one or leaves it unchanged.

For nodes i and j, the link ij creates new intermediation rents from any new connection of length two that previously was of length greater than two. Such connections involve i and j and one other node connected to exactly one of the two. The new connection will have exactly one path of length two, via the new link ij, and thus will earn a new intermediation rent. For node k, the new link ij does not create new intermediation rents. The new link can, however, result in the loss of intermediation rents for k, in cases where a new alternative route of length two involving the segment ij bypassing k is created. Any such connection involves i and j and a third node that is connected to k as well as to either i and j. In addition, k can lose intermediation rents from the connection between i and j itself, which is replaced by a direct connection of length one.

In terms of the payoff function for node $p \in \{i, j, k\}$:

$$
\begin{aligned}
&\Delta\pi_p^{\text{Rents}} \\
=& \sum_{\substack{u \neq p, w \neq p: \\ pu \in g+ij \\ \wedge pw \in g+ij}} I_{\{\ell_{uw}(g+ij)=2 \wedge r_{uw}(g+ij)=1\}} \delta\gamma - \sum_{\substack{u \neq p, w \neq p: \\ pu \in g \\ \wedge pw \in g}} I_{\ell_{uw}(g)=2 \wedge \{r_{uw}(g)=1\}} \delta\gamma \\
=& \delta \sum_{\substack{u \neq p, w \neq p: \\ pu \in g+ij \\ \wedge pw \in g+ij}} I_{\left\{\substack{\ell_{uw}(g)=3 \wedge \ell_{uw}(g+ij)=2 \\ \wedge r_{uw}(g+ij)=1}\right\}} \gamma - \delta \sum_{\substack{u \neq p, w \neq p: \\ pu \in g \\ \wedge pw \in g}} \left[I_{\left\{\substack{\ell_{uw}(g)=2 \wedge \ell_{uw}(g+ij)=2 \\ \wedge r_{uw}(g)=1 \wedge r_{uw}(g+ij)=2}\right\}} + I_{\left\{\substack{\ell_{uw}(g)=2 \wedge \ell_{uw}(g+ij)=1 \\ \wedge r_{uw}(g)=1}\right\}} \right] \gamma \\
=& \begin{cases} \delta \sum_{\substack{u \neq p, w \neq p: \\ pu \in g+ij \\ \wedge pw \in g+ij}} I_{\left\{\substack{\ell_{uw}(g)=3 \wedge \ell_{uw}(g+ij)=2 \\ \wedge r_{uw}(g+ij)=1}\right\}} \gamma & \text{if } p \in \{i, j\} \\ -\delta \sum_{\substack{u \neq p, w \neq p: \\ pu \in g \\ \wedge pw \in g}} \left[I_{\left\{\substack{\ell_{uw}(g)=2 \wedge \ell_{uw}(g+ij)=2 \\ \wedge r_{uw}(g)=1 \wedge r_{uw}(g+ij)=2}\right\}} + I_{\left\{\substack{\ell_{uw}(g)=2 \wedge \ell_{uw}(g+ij)=1 \\ \wedge r_{uw}(g)=1}\right\}} \right] \gamma & \text{if } p \in \{k\} \end{cases} \\
=& \begin{cases} \delta \sum_{\substack{u \notin \{i,j,k\}, \\ w \in \{i,j\}: \\ pu \in g \wedge uw \notin g}} I_{\{\ell_{uw}(g)>2\}} \gamma & \text{if } p \in \{i, j\} \\ -\delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ pu \in g \wedge uv \in g \\ \wedge uw \notin g}} I_{\{r_{uw}(g)=1\}} \gamma - \delta I_{\{r_{ij}(g)=1\}} \gamma & \text{if } p \in \{k\} \end{cases}
\end{aligned}
$$

It then follows that

$$
\Delta\pi_{ijk}^{\text{Rents}} = \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v, w \in \{i,j\}: \\ vu \in g \wedge us \notin g}} I_{\{\ell_{uw}(g)>2\}} \gamma - \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ ku \in g \wedge uv \in g \\ \wedge uw \notin g}} I_{\{r_{uw}(g)=1\}} \gamma - \delta I_{\{r_{ij}(g)=1\}} \gamma \tag{32}
$$

Summing Across All Components

$$
\begin{aligned}
\Delta\pi_{ijk} =& \Delta\pi_{ijk}^{\text{Direct}} + \Delta\pi_{ijk}^{\text{Introductions}} + \Delta\pi_{ijk}^{\text{Indirect}} + \Delta\pi_{ijk}^{\text{Rents}} \\
=& 1 - c + \epsilon c \\
& + \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ uv \in g \wedge uw \notin g}} \left[I_{\ell_{uw}(g)>2} \frac{1-\gamma}{2} + I_{\{\ell_{uw}(g)=2 \wedge r_{uw}(g)=1\}} \frac{\gamma}{2} \right] - \delta \left[1 - I_{\{r_{ij}(g)=1\}} \gamma \right] \\
& + \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ vu \in g \wedge us \notin g}} I_{\{\ell_{uw}(g)>2\}} \gamma - \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ ku \in g \wedge uv \in g \\ \wedge uw \notin g}} I_{\{r_{uw}(g)=1\}} \gamma - \delta I_{\left\{r_{ij}(g)=1\right\}} \gamma \\
=& 1 - (1-\epsilon)c - \delta \\
& + \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ uv \in g \wedge uw \notin g}} \left[I_{\ell_{uw}(g)>2} \frac{1-\gamma}{2} + I_{\{\ell_{uw}(g)=2 \wedge r_{uw}(g)=1\}} \frac{\gamma}{2} \right] \\
& + \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ vu \in g \wedge us \notin g}} I_{\{\ell_{uw}(g)>2\}} \gamma - \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ ku \in g \wedge uv \in g \\ \wedge uw \notin g}} I_{\{r_{uw}(g)=1\}} \gamma \\
=& 1 - (1-\epsilon)c - \delta \\
& + \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ uv \in g \wedge uw \notin g}} \left[I_{\ell_{uw}(g)>2} \frac{1+\gamma}{2} + I_{\{\ell_{uw}(g)=2 \wedge r_{uw}(g)=1\}} \frac{\gamma}{2} \right] \\
& - \delta \sum_{\substack{u \notin \{i,j,k\}, \\ v \in \{i,j\}, w \in \{i,j\}: \\ uv \in g \wedge uw \notin g \\ \wedge ku \in g}} I_{\{r_{uw}(g)=1\}} \gamma
\end{aligned}
$$

We can now split intermediation rents gained from nodes unconnected to k and from nodes connected to k and sum the latter with intermediation rents lost by node k. The gains from new indirect connections remain unaffected from conditioning on $ku \notin g$ as $u \notin \{i,j,k\}, w \in \{i,j\}$: $\ell I_{\ell_{uw}(g)>2}$ implies $ku \notin g$. Taken together this argument yields the desired term.

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

□

A.5 Proof of Proposition 8

Proof. 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.

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. Furthermore, assume μ is strictly less than $\mu^* = \left\lfloor 2\frac{1-(1-\epsilon)c-\delta}{\delta\gamma} \right\rfloor$.

Substituting the expression in Lemma 7 and dropping the non-negative middle term, the payoff change to ijk is then at least

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

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

□

A.6 Proof of Proposition 9

Proof. The proof is by contradiction. Assume that g is myopically stable with introductions and that the clustering coefficient C_k of node k with degree d_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 N_k. Let g_K be the subnetwork of g formed by the set of nodes N_k and the links that are links between nodes in N_k in g. Note that g_K does not include k and all the links in g_K are links between neighbors of k. Intuitively, the links in g_K are the links forming closed triangles around k, which are required for stability against introductions in the sense of Proposition 8. The degree of node i within g_K is labeled d_i^K and the average degree across nodes in N_k is labeled $\bar{d}^K$.

To show the contradiction, we distinguish two cases based on whether or not the clustering condition of Proposition 9 implies a complete subnetwork g_K:

Case 1: $d_k \leq \lambda^*$**:** If $d_k \leq \lambda^*$, then $\underline{C}(d_k) = 1$; that is, at the lower bound on clustering of Proposition 9 the neighbors of k form a fully connected subnetwork. Thus, $C_k < \underline{C}(d_k)$ implies $\exists i, j \in N_k$ such that $ij \notin g_K$. As $d_k \leq \lambda^*$, we have both $d_i^K < \lambda^* - 1 = \frac{\mu^*}{2}$ and $d_j^K < \lambda^* - 1 = \frac{\mu^*}{2}$, with strict inequality as i and j are not connected to each other. It follows that the total number of neighbors that the pair i and j share with k but not with each other is strictly less than μ^*: $d_i^K + d_j^K < \mu^*$. By Proposition 8 this implies that an introduction of i and j by k is profitable and thereby g is not myopically stable with introductions. This delivers the contradiction.

Case 2: $d_k > \lambda^*$**:** If $d_k > \lambda^*$, then $\underline{C}(d_k) = \left\lfloor \frac{d_k}{\lambda^*} \right\rfloor \frac{(\lambda^*-1)\lambda^*}{(d_k-1)d_k}$; that is, at the lower bound on clustering of Proposition 9 the subnetwork of the neighbors of k is not fully connected. Now, as we assume $C_k < \underline{C}(d_k)$ we have

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

We now show that this upper bound on clustering is equivalent to a limit on the number of links within g_K and thereby on $\bar{d}^K$, the average degree of nodes in N_k. Recall that the total number of possible links in g_K, a network of d_k nodes, is $\frac{(d_k-1)d_k}{2}$. Thus, the upper bound on clustering implies that the number of links in g_K is strictly less than $\frac{(\lambda^*-1)}{(d_k-1)} \times \frac{(d_k-1)d_k}{2} = \frac{(\lambda^*-1)d_k}{2}$ As each link is shared by two nodes, it follows that the average degree $\bar{d}^K < \lambda^* - 1 = \frac{\mu^*}{2}$.

We can now show that there is at least one unconnected pair of nodes whose degrees in g_K

add up to less than μ^*, that is $\min_{ij \notin g_K} \left[d_i^K + d_j^K \right] < \mu^*$.

$$\min_{ij:i,j \in N_k, ij \notin g_K} \left[d_i^K + d_j^K \right] \leq \frac{1}{|\{ij : i, j \in N_k, ij \notin g_K\}|} \sum_{ij:i,j \in N_k, ij \notin g_K} \left[d_i^K + d_j^K \right] \tag{33}$$

$$\leq \frac{1}{|\{ij : i, j \in N_k\}|} \sum_{ij:i,j \in N_k} \left[d_i^K + d_j^K \right] \tag{34}$$

$$= \frac{2}{(d_k - 1)} \sum_{ij:i,j \in N_k} \left[d_i^K + d_j^K \right] \tag{35}$$

$$= \frac{2}{d_k(d_k - 1)} \sum_{i \in N_k} \sum_{j \in N_k : j > i} \left[d_i^K + d_j^K \right] \tag{36}$$

$$= \frac{1}{d_k(d_k - 1)} \sum_{i \in N_k} \sum_{j \in N_k : j \neq i} \left[d_i^K + d_j^K \right] \tag{37}$$

$$= \frac{1}{d_k(d_k - 1)} \left\{ \sum_{i \in N_k} \sum_{j \in N_k : j \neq i} d_i^K + \sum_{j \in N_k} \sum_{i \in N_k : i \neq j} d_j^K \right\} \tag{38}$$

$$= \frac{1}{d_k(d_k - 1)} \left\{ \sum_{i \in N_k} (d_k - 1) d_i^K + \sum_{j \in N_k} (d_k - 1) d_j^K \right\} \tag{39}$$

$$= \frac{1}{d_k} \left\{ \sum_{i \in N_k} d_i^K + \sum_{j \in N_k} d_j^K \right\} \tag{40}$$

$$= 2\bar{d}_k < \mu^* \tag{41}$$

The first step uses the fact that the 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 places less weight on high-degree nodes relative to low-degree nodes.[10] By Proposition 8 this implies that the network is not myopically stable with introductions, delivering the contradiction.

□

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