Network Formation with Local Benefits*

Qingchao Zeng†

September 5, 2025

Abstract

We consider a non-cooperative model of network formation where agents decide on whom to form costly links to. Links are unilaterally formed, and payoff flows one way to the active side. We study discontinuous information flows where agents only receive benefits from other agents that are within a distance of two on the network, i.e. only from their 'friends' and 'friends of friends'. For the static game, we show that the set of strict Nash equilibria encompasses a multiplicity of core-periphery network structures. We further study a noisy best response process to obtain long-run predictions. Doing so, we find that the set of stochastically stable states retains a multiplicity of network structures, many of which are not efficient.

Keywords: Network Formation, One-way flow, Stochastic stability, Efficiency.

JEL: C72; D83; D85

*This paper is the first chapter of the author's Ph.D. thesis at the University of Essex.

†Department of Economics, University of Fribourg, Bd. de Pérolles 90, Fribourg, 1700, Switzerland. Email: qingchao.zeng@unifr.ch.

1 Introduction

The role of networks in various social and economic activities has attracted considerable attention in academic research. Comprehending the impact of networks constitutes a pivotal area of inquiry across diverse contexts, ranging from R&D collaboration to labour markets.[1] Thus, it is important to know which kinds of network configurations will form and what drives the stability and efficiency of networks.

The motivation of this paper is based on the following two observations. First, many social networks exhibit some structures with small diameters.[2] The distance between any two nodes in a network is usually relatively small. Second, friction in the spread of information is inevitable in the real world due to noise, privacy and resistance, etc. This paper introduces a network formation model that effectively captures these characteristics, based on the key assumption that information transmission is constrained within two steps. Specifically, agents obtain information either directly from their immediate connections (referred to as 'friends') or indirectly from the connections of their friends (commonly termed as 'friends of friends').[3] An example is networks of citations, where researchers frequently cite the literature by others.[4] An interesting truth about citation networks is that through the references cited by one paper, one can access to the information about the literature cited by those references. Information beyond 'citations' and 'citations of citations' is not directly seen. Therefore, to get information on other related literature, one has to either reach those 'citations of citations', or reach some new literature not included in those two groups.

In this paper, we set up a non-cooperative model of network formation, where links are unilaterally formed and information flows one way. Under one-way flow, the payoff is only received by the agent who initiates the link.[5] The information carried by each agent is the source of payoffs.

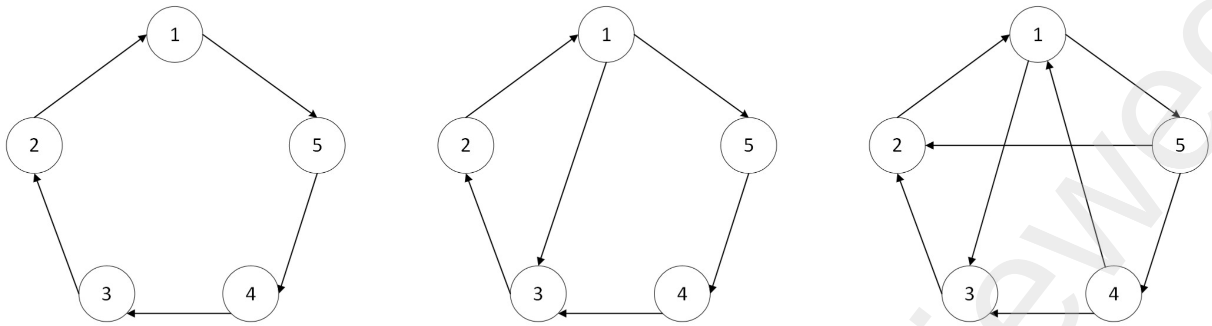
[1]Topics involve R&D networks (Goyal and Moraga-Gonzalez, 2001; Goyal and Joshi, 2003), networks in labour markets (Calvo-Armengol and Jackson, 2004, 2007), public goods in networks (Bramoullé and Kranton, 2007; Allouch, 2015); social coordination (Goyal and Vega-Redondo, 2005; Staudigl and Weidenholzer, 2014 and Cui and Weidenholzer, 2021).

[2]Examples include the small-world phenomenon (Milgram, 1967) and six degrees of separation (Guare, 2016).

[3]This setup is featured as 'truncated connections' in Jackson and Wolinsky [1996] or 'communication threshold' in Hojman and Szeidl [2008].

[4]See e.g. Price [1965] who analyses the growing citation networks and documents that they are scale-free networks.

[5]See e.g. Bala and Goyal, 2000; Billand et al., 2008 and Cui et al., 2013. In contrast, two-way flow assumes that



(a) Wheel (b) Wheel + one additional link (c) Wheel + three additional links

Figure 1: Collapse of the wheel: consider an example of five agents, the wheel depicted in (a) is a strict Nash network in Bala and Goyal's one-way flow model without decay. When agents can only receive information from others within distance two, agent 1 has an incentive to form an additional link to agent 3 to get information from 3 and 2 as shown in (b). Similarly, agent 5 and agent 4 also have incentives to form additional links as shown in (c).

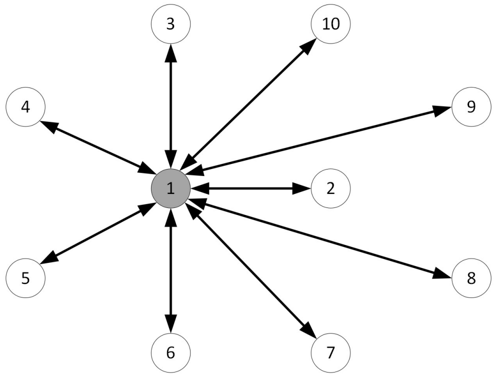
By forming costly links, agents receive payoffs from their friends and friends of friends.

Bala and Goyal [2000] have shown that in the absence of limits on information transmission, the wheel is the unique strict Nash equilibrium. On an intuitive level, when the information flows one way and the linking cost is low enough, the network will have to feature cycles, implying that everybody has to be connected so that there exists a path from each agent to all other agents, while there also exists a path in the reverse direction. But it cannot be insistent where two paths cross in a strict Nash equilibrium because that would imply that some agents are indifferent in their linking choices, leaving the wheel as the only strict Nash equilibrium.

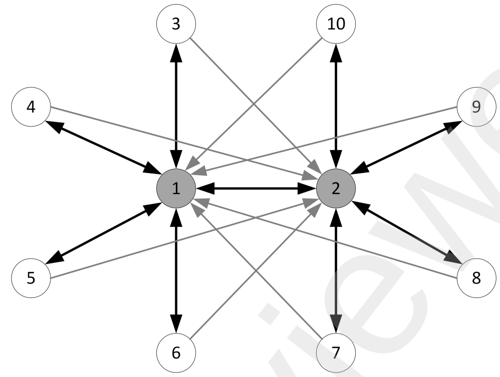
However, in the presence of constraints towards distance, the wheel is not a strict Nash equilibrium since agents have incentives to form additional links to those beyond distance two. Figure 1 illustrates this point. The main result under our key assumption shows that the network formation game has multiple strict Nash network configurations, which we refer to as *core-periphery networks*.[6] In these strict Nash equilibria, there is a small set of core agents who maintain some links to other core agents and periphery agents and a large set of periphery agents who maintain only

both sides of a link receive the payoff (see e.g. Bala and Goyal, 2000; Feri, 2007; Billand et al., 2011 and De Jaegher and Kamphorst, 2015).

[6] Borgatti and Everett [2000] formalize the concept of core-periphery structure in the context of undirected networks. In this paper, our definition and notions of core-periphery networks are closely related to the definition of the directed core-periphery network in Elliott et al. [2020].



(a) CP_1-star



(b) CP_2 Network

Figure 2: Two core-periphery networks with 10 agents

links to core agents. Figure 2 shows two examples of such equilibrium networks. The logic is that due to the constraint imposed on information transmission, agents form links to keep others within a two-step distance. Conversely, each agent has the incentive to minimize the number of links required to fulfil this objective. In a core-periphery network, links formed by core agents ensure that any given agent can keep others within two steps by forming links to all core agents. Besides, the compact size of the core agent set enables agents to link to others with a minimal number of links.

Furthermore, due to the multiplicity of strict Nash equilibria, we study a noisy best response process to characterize stochastically stable states. This approach has various applications in the literature on network formations, particularly in the selection of multi-equilibria.[7] Each agent has a positive probability of receiving opportunities to update their linking strategies in a discrete time. Sometimes they make mistakes and fail to maximize their payoffs. We adopt the methodology developed by Kandori et al. [1993] and Young [1993], allowing us to identify stochastically stable states, i.e. states in support of the invariant distribution of the Markov process as the probability of mistakes vanishes. Our results show that the set of stochastically stable states retains a multiplicity of network structures, which encompasses core-periphery networks. Additionally, we study the number of links that efficient networks have and show that star networks (also referred as CP_1 networks) have the least number of links among all core-periphery networks. Thus, any other

[7] See e.g. Jackson and Watts [2002a], Feri [2007] and Cui et al. [2013] for applications to network formation.

core-periphery networks are inefficient as they are payoff dominated by the star. Thus, with the constraint on the distance of information transmission, we may observe inefficient outcomes. This is in contrast to the case where there is no constraint and the unique strict Nash network, the wheel, is also efficient.

The paper is organized as follows. Section 2 discusses the relation between our paper and the existing literature. In section 3, we describe the details of the setups of the network formation game. In section 4, we present our results on strict Nash equilibria. Section 5 describes the learning dynamics and our analytical results of the long-run predictions. We also have a short discussion on the efficiency of stochastically stable states in section 5. Section 6 concludes. Table 1 provides the list of notions and their definitions in the paper and an appendix contains the proofs of our key results.

2 Literature review

The present paper is closely related to the broad literature on network formation. Jackson and Wolinsky [1996] explore the truncated connection model with some bound D in the cooperative network formation model where forming a link requires mutual consent of both parties. They show that the pairwise stable networks exhibit a property that the maximum distance between any two players is $2D - 1$. Further, Bala and Goyal [2000] presents a non-cooperative network formation model where Nash equilibrium can be used to characterise the stable network architectures. They have broad discussions on models of one-way flow and two-way flow in cases with decay (where there are frictions in information) and without decay (where there is no friction in information). They provide characterizations of Nash equilibrium networks and efficient networks, which exhibit some simple architectures, e.g. the wheel and the star. Based on Bala and Goyal [2000]'s two-way flow model with decay, Hojman and Szeidl [2008] study a model where links have decreasing returns and show that for some parameters, the unique non-empty Nash equilibrium network is the periphery-sponsored star at the presence of communication threshold. Our work differs from these

Table 1: List of Notions and Their Definitions

Notions	Definitions
n	Number of agents, which is larger than or equal to 4.
N	Set of all agents.
g_{ij}	The linking decision of agent i to j.
g_i	A N-tuple of agent i's linking strategy to each agent.
g	A network, i.e. a strategy profile of all agents.
$\mathcal{G}_i$	The strategy profile of agent i.
$\mathcal{G}$	The set of all strategy profiles.
g_{-i}	The network formed by agents other than i.
$g + ij$	The network obtained by adding the link from i to j.
$g - ij$	The network obtained by deleting the link from i to j.
$d(i, j; g)$	The distance from j to i in a given network g.
$N_i^d(g)$	The d-neighbourhood of agent i, i.e. the set of agents who are at distance d to i.
$n_i^d(g)$	The number of agents in i's d-neighbourhood.
d_i^{out}	The out-degree of agent i, i.e. the number of links that agent i actively forms.
d_i^{in}	The in-degree of agent i, i.e. the number of links that agent i passively receives.
C	A component of N.
g^i	The sub-network on component C_i.
c	The cost of forming a link.
$\mathcal{G}^\star$	The set of strict Nash equilibrium networks.
CP_ℓ	A core-periphery network with ℓ core agents.
$C(\ell; g)$	The set of core-agents given a core-periphery network g.
P	The set of periphery agents.
P_i	The set of agent i's periphery agents.
$\mathcal{CP}_\ell$	The set of strict Nash CP_ℓ networks.
$\bar{\ell}$	The maximum number of core agents in any strict Nash core-periphery network.
$\mathcal{CP}_{\bar{\ell}}$	The set of all strict Nash core-periphery networks.
$G^{\star\star}$	An absorbing set.
$\mathcal{G}^{\star\star}$	The set of all absorbing sets.
$\mathcal{G}^{\star\star\star}$	The set of stochastically stable states.
$r(g, g')$	The resistance of transition form network g to g'.
τ_i	A $G_i^{\star\star}$-tree, i.e. a spanning tree rooted in the absorbing set $G_i^{\star\star}$.
T_i	The set of all $G_i^{\star\star}$-trees.
$r(G_i^{\star\star})$	The resistance of a $G_i^{\star\star}$-tree.
$\gamma(G_i^{\star\star})$	The stochastic potential of the absorbing set $G_i^{\star\star}$.
$\boldsymbol{N}$	The number of all absorbing sets.
$W(g)$	The welfare generated by given network g.
$W(ij)$	The contribution of the link i to j to the welfare.

models in one main direction. We introduce constraints on information transmission to Bala and Goyal's one-way flow model without decay and exhibit a different class of strict Nash equilibrium networks – core-periphery networks – which obtain the star.

This paper also adds to the literature on the dynamics of social networks. Watts [2001] presents a dynamic network formation with two-sided links through independent decisions[8] and shows that the star is both stable and efficient for some parameters. Jackson and Watts [2002a] study a stochastic evolution of network formation and find that a stochastically stable network is either pairwise stable or part of a closed cycle. Further, Feri [2007] considers the noisy best response learning in Bala and Goyal's two-way flow model with decay and shows that the periphery-sponsored star is the unique stochastically stable network architecture. Cui et al. [2013] explore the evolutionary version of Bala and Goyal's one-way flow model with decay and find that either the empty network or the wheel is the stochastically stable state. Our work contributes to these results in two respects. First, we demonstrate that core-periphery networks exhibit stochastic stability, highlighting the potential for the emergence of networks other than simple architectures, e.g. the wheel and the star. Second, we showcase the possibility of obtaining inefficient network architectures in the long run.

Our paper also departs from the literature in modelling payoffs. Goyal and Vega-Redondo [2007] study a model of pairwise links where two parties of a link split the payoffs. They show that the star emerges in the absence of capacity constraints on links and the cycle network is stochastically stable when the capacity of links is relatively small to the population. In Galeotti and Goyal [2010], a player's payoff depends on how much information she and her neighbours acquire. They show that the equilibrium networks exhibit 'the Law of the Few' and have a core-periphery structure, i.e. few players in the core acquire information and many players in the periphery acquire no information. In contrast to these studies, we follow Bala and Goyal [2000]' one-way flow model, where all agents carry the same value of information and the payoff goes to the one who initiates the link.

A different branch in the literature analyses models of co-evolution of coordination games and

[8] Two-sided link through independent decisions is that the formation of a link between two agents requires that both sides wish it (see e.g. Goyal and Vega-Redondo, 2005, and Fosco and Mengel, 2011).

network formations, where in addition to their linking choice, agents also have to decide the action played in the coordination game. Jackson and Watts [2002b] study an evolutionary model in social coordination games where the network is bilaterally formed, which prescribes the use of the concept of pairwise stability in Jackson and Wolinsky [1996]. They show that for some parameters the networks in stochastically stable states exhibit fully connected configurations. In Goyal and Vega-Redondo [2005], agents unilaterally decide on whom to link to. They show that the equilibrium network is either empty or complete. Staudigl and Weidenholzer [2014] set up a model with a restricted maximum number of links that each agent can support and show a variety of the equilibrium network structures in the long run. Cui and Weidenholzer [2021] consider a case where an agent is able to receive payoffs from links that other agents form to her. They find that the Nash equilibrium networks do not have to be fully connected and that architectures, where agents use different actions, may sometimes be stochastically stable.

3 Model

We consider a one-way flow model of network formations.[9] There is a population of n agents, denoted by $N = \{1, 2, \cdots, n\}$ with $n \geq 4$. Each agent $i \in N$ decides the set of agents to whom she forms links. A strategy used by agent i is given by a N-tuple $g_i = (g_{i1}, g_{i2}, \cdots, g_{in})$ where $g_{ij} \in \{0, 1\}$ is agent i's linking decision to agent j. We say i links up to j if $g_{ij} = 1$; otherwise, $g_{ij} = 0$. We assume that agents cannot link to themselves, i.e. $g_{ii} = 0, \forall i \in N$. Further, let $\mathcal{G}_i$ be the set of all possible link strategies that agent i can choose. One-way flow model implies that agent i's linking decision to j is independent with j's decision to i, i.e. g_{ji} and g_{ij} are not necessarily equal. A network $g = (g_i)_{i \in N} \in \mathcal{G}$ is the strategy profile of all agents, where $\mathcal{G} = \prod_{i \in N} \mathcal{G}_i$ is the set of all networks. Moreover, let $g_{-i} = g - g_i$ denote the network formed by agents other than i. Further, we denote by $g + ij$ the network obtained by adding the link from agent i to j to network g. Similarly, $g - ij$ denotes the network obtained by deleting the link from i to j in network g.

[9] We follow the notations introduced in Jackson and Wolinsky [1996] and Bala and Goyal [2000].

We say there exists a path from agent j to i if either $g_{ij} = 1$ or there is a set of agents $\{k_1, k_2, \cdots, k_m\}$, such that $g_{ik_1} = g_{k_1k_2} = \cdots = g_{k_mj} = 1$. The distance from j to i, denoted by $d(i, j; g)$, is the number of links of the shortest path from agent j to i.[10] Further, we define agent i's d-neighbourhood as the set of agents who are at distance d to i, denoted by $N_i^d(g) = \{j \in N : d(i, j; g) = d\}$, with $n_i^d(g) = |N_i^d(g)|$ the number of agent i's d-neighbours. We refer to $d_i^{out} := n_i^1(g) = \sum_{j \in N} g_{ij}$ as the out-degree of agent i, i.e. the number of active links that agent i forms. We also denote by $d_i^{in} = \sum_{j \in N} g_{ji}$ the in-degree of agent i, i.e. the number of passive links that agent i receives from others.

A subset $C \subseteq N$ is called a strongly connected component if $\forall i, j \in C$ with $i \neq j$, there exists a path from i to j as well as a path from j t i, and there is no strict superset, i.e. $C \subset C' \subseteq N$ for which this is true. A network g is strongly connected if it has a unique strongly connected component. Let g^i be the sub-network within agents in the strongly connected component C_i.

We now define the payoffs in our network formation game. As in Bala and Goyal [2000], each agent receives information from others by forming costly links and benefits from doing so. Without loss of generality, the value of information that each agent has is homogeneously normalized to one. The cost of each link is $c > 0$. We assume that the cost is only incurred by the party that initiates the link. We deviate from Bala and Goyal [2000] by studying the case where the distance that information can travel is constrained. Let D be the constraint. This implies that agents can receive information from others who are within a distance of D. As motivated in the introduction, we assume that agents can only observe their neighbours and the neighbours of neighbours, i.e. we focus on the case where $D = 2$.[11]

An agent's payoff is calculated as the sum of benefits derived from the information she receives from others, minus the total cost incurred from forming links. More formally, given a network

[10] In a directed network g, $d(i, j; g)$ and $d(j, i; g)$ can be different. We say $d(i, j; g) = \infty$ if j is not linked by i neither directly nor indirectly.

[11] This is a special case of the communication threshold in Hojman and Szeidl [2008], which in contrast to our model focuses on two-way flow. The existence of limits in communication is also observed in the one-way flow case, which their model cannot characterize.

$g = (g_i)_{i \in N}$, agent i's payoff is given by

$$U_i(g_i, g_{-i}) = 1 + n_i^1(g) + n_i^2(g) - c \cdot \sum_{j \in N} g_{ij} \tag{1}$$

where the constant captures the value of information of agent i self.

4 Strict Nash Networks

In the first step, we characterize some important properties of strict Nash networks. We follow the definition of strict Nash networks in Bala and Goyal [2000], which is formalized as Definition 4.1.

Definition 4.1 (Strict Nash Networks)**.** *A network $g = (g_i)_{i \in N}$ is a strict Nash network if and only if $U_i(g_i, g_{-i}) > U_i(g_i', g_{-i})$ for all $g_i' \in \mathcal{G}_i$ and $i \in N$.*

We denote by $\mathcal{G}^\star$ the set of all permissible strict Nash networks. The first two technical lemmas establish some useful properties of the natures of out-degrees and in-degrees in a strict Nash network.

Lemma 1. *For any non-empty strict Nash network $g \in \mathcal{G}^\star$, $d_i^{out} \geq 1$ for any $i \in N$.*

Lemma 1 provides that every agent supports at least one active link. The intuition is that if supporting an active link is profitable for some agents, then it must be the case that it is profitable for every agent. It is trivial for the case where $c < 1$, forming a link to another agent yields at least $1 - c > 0$. Note that in this case, any strict Nash network has to be non-empty. The next lemma establishes a similar insight regarding the in-degrees of agents.

Lemma 2. *For any non-empty strict Nash network $g \in \mathcal{G}^\star$, $d_i^{in} \geq 1$ for any $i \in N$.*

Lemma 2 implies that every agent receives at least one passive link. For the case $c < 1$, this result is trivial because if an agent i receives no passive links, another agent j will get an additional payoff $1 - c$ by forming a link to i. Consider the other case $c > 1$, the intuition is more complicated. The main idea in this case is based on the observation that agents without passive links turn out to be

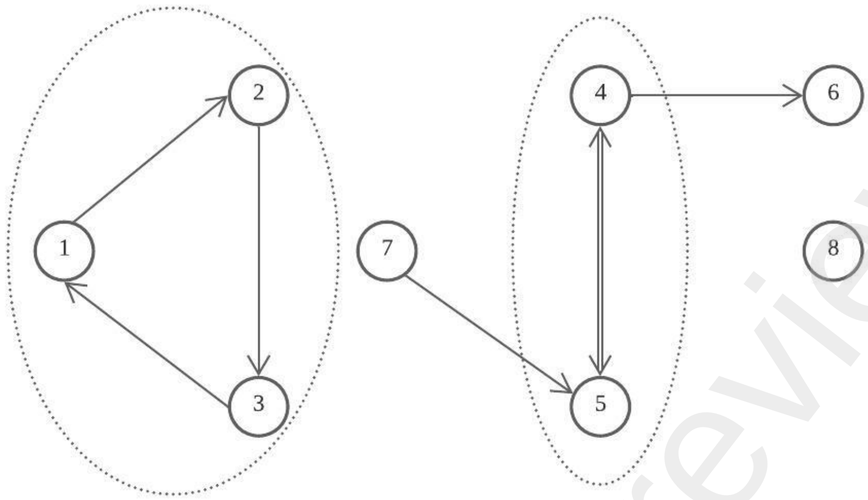


Figure 3: A network which is not strongly connected

either indifferent between whom to form links to or have profitable deviation. This is incompatible with a strict Nash network.

We now move towards studying the implications of the two observations for the general structures of strict Nash networks. Recall the definition of strongly connected networks. Strong connectedness implies that there exists a unique strongly connected component in the network, which requires that for any two agents i and j, there is a path from i to j and a path from j to i. Any network that is not strongly connected, has to consist of multiple strongly connected components. These strongly connected components could be isolated from each other, or a path exists from one component to another, but not vice versa. The network depicted in Figure 3 illustrates such properties. In this network, there are five strongly connected components in the network: $\{1, 2, 3\}$, $\{4, 5\}$, $\{6\}$, $\{7\}$ and $\{8\}$. Component $\{8\}$ is totally isolated from other components, while components $\{6\}$ and $\{4, 5\}$ are not strongly connected since there is no path from agent 6 to either agent 4 or agent 5.

Lemma 3. *Any non-empty strict Nash network is strongly connected.*

Lemma 3 is trivial for the case $c < 1$. If a strict Nash network g is not strongly connected, then there exist two agents i and j such that there is no path from i to j. Therefore, there is a profitable deviation for agent i to form a link to j. For $c > 1$, the lemma follows from combining the previous two lemmas and pointing out that in any not strongly connected component some agents will necessarily have an incentive to link to agents in other components.

While we are able to obtain the previous results for the general case where c may be larger than one, for the following analysis we have to restrict ourselves to $c < 1$. Revisit networks depicted in Fig 2. In the star network shown in Fig 2(a), the agent in the centre links to everyone else whilst receiving links from everyone. One can check that the star is a strict Nash network as the agent in the centre and agents in the periphery all give a unique best response.

In fact, there also exists another class of networks that possess these properties. We refer to these networks as core-periphery networks.[12] Less formally, a core-periphery network consists of two sets of agents: core and periphery. Each agent in the core (termed as *core agent*) links to every agent in the core and also links to a subset of the other agents in the periphery. Each agent in the periphery (termed as *periphery agent*) links to all core agents and forms no link to any other periphery agent. The more formal definition is given as follows.

Definition 4.2. *A network g is called a core-periphery network, denoted by CP_ℓ if*

1) each agent is either a core agent or a periphery agent, i.e. $N = C(\ell; g) \cup P(g)$ *and* $C(\ell; g) \cap P(g) = \emptyset$*, where* $C(\ell; g) = \{1, 2, \cdots, \ell\}$ *is the set of core agents and* $P(g) = \{\ell+1, \ell+2, \cdots, n\}$ *is the set of periphery agents;*

2) core agents link to each other directly, i.e. $g_{ij} = 1, \forall i, j \in C(\ell; g)$ *with* $i \neq j$*;*

3) each core agent i links to a subset of periphery agents, i.e. $P_i(g) = \{j \in P(g) : g_{ij} = 1\}$*;*

4) each periphery agent is linked by a single core agent, i.e. $\bigcap_{i \in C(\ell;g)} P_i(g) = \emptyset$ *and* $P(g) = \bigcup_{i \in C(\ell;g)} P_i(g)$*;*

5) each periphery agent links to all core agents, i.e. $g_{ij} = 1, \forall i \in P(g), j \in C(\ell; g)$*;*

6) there is no link between periphery agents, i.e. $g_{jk} = g_{kj} = 0, \forall j, k \in P(g)$.

The network shown in Fig 2(b) depicts such a core-periphery network with two core agents.

[12] Core-periphery networks are also featured previously in the literature (see in Borgatti and Everett, 2000 and Elliott et al., 2020), but with slightly different definitions in our paper.

Agents 1 and 2 are core agents and all other agents link up to them. The other eight agents are periphery agents, each of whom is linked by either agent 1 or agent 2.

We classify the core-periphery networks by the number of core agents so that CP_ℓ denotes a core-periphery network with ℓ core agents. Note that for a given number of core agents ℓ, there is a multiplicity of different CP_ℓ networks, varying in the identities of core agents and their peripheries with similar structures. The network in Fig 2(b) is a CP_2 network and the star is the special case of CP_1 network. One can also check that the CP_2 network depicted in Fig 2(b) is a strict Nash network since core and periphery agents all give a unique best response.

Having defined core-periphery networks, we proceed to prove that the set of strict Nash networks includes CP_ℓ networks with a certain condition. The following proposition exhibits our main results on strict Nash networks.

Proposition 4.1. *Any CP_ℓ network g with $|P_i(g)| \geq 3, \forall i \in C(\ell; g)$ is a strict Nash equilibrium.*

The intuition of Proposition 4.1 is as follows. First, core agents link up to the other core agents and their respective peripheries such that they can observe everyone. Adding any other link yields no additional payoffs, and deleting any link means a reduction in payoffs. Further, periphery agents observe everybody by linking to all core agents. Thus, they have no incentives to form additional links. If they delete some links, they will lose access to the periphery agents of these core agents, leading to a decrease in payoffs. More importantly, the condition on the number of periphery agents ensures that the best response of each agent is unique, i.e. there is no alternative strategy that yields the same payoffs as in a CP_ℓ network. Figure 4 (a) depicts a CP_2 network where core agent 1 has only two periphery agents. In this network, the periphery agent 3 is indifferent between linking to agent 1 and agent 4 (see Figure 4b). Therefore, the CP_2 network in Figure 4 (a) is not a strict Nash equilibrium.

Proposition 4.1 shows that any core-periphery network in which each core agent links up to at least three periphery agents is a strict Nash network. Now, we denote by $\mathcal{CP}_\ell$ the set of strict Nash

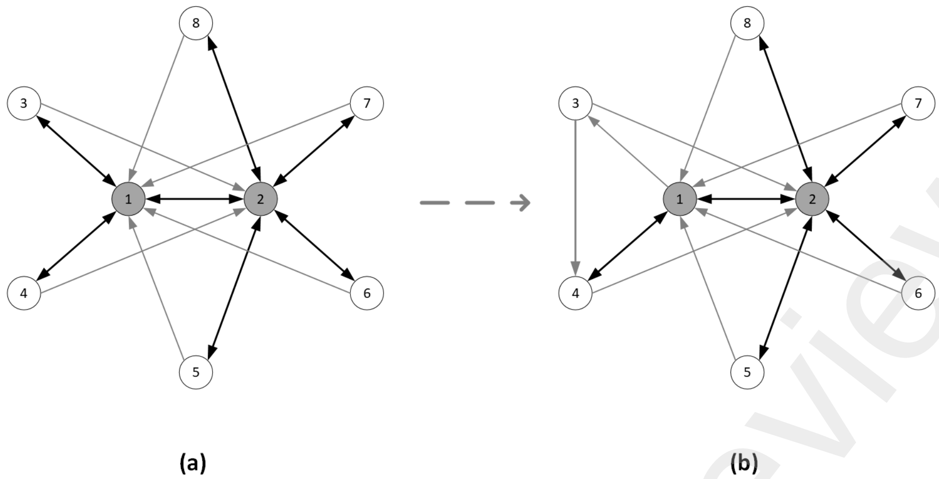


Figure 4: A core-periphery network where the core agent 1 has only two periphery agents.

core-periphery networks with ℓ core agents. More formally, $\mathcal{CP}_\ell$ is defined by

$$\mathcal{CP}_\ell := \{g : g \text{ is a } CP_\ell \text{ network and } |P_i(g)| \geq 3, \text{for all } i \in C(\ell; g)\}.$$

Note that the condition on the size of each set of periphery agents imposes an upper bound on the number of core agents in a strict Nash core-periphery network. To be more precise, in a strict Nash CP_ℓ network, each core agent has to have at least three periphery agents, therefore the number of periphery agents is at least three times more than ℓ. Thus, given the number of all agents n, the number of core agents can never exceed $\lfloor \frac{n}{4} \rfloor := \bar{\ell}$. If a core-periphery network has more than L core agents, then it is not a strict Nash network. We denote by $\mathcal{CP}_{\bar{\ell}}$ the set of all strict Nash CP_ℓ networks in which the number of core agents is less or equal to $\bar{\ell}$. More formally, the set $\mathcal{CP}_{\bar{\ell}}$ at a given n is defined as

$$\mathcal{CP}_{\bar{\ell}} := \bigcup_{1 \leq \ell \leq \bar{\ell}} \mathcal{CP}_\ell.$$

Proposition 4.1 implies that there is a multiplicity of strict Nash network configurations. This result differs from Bala and Goyal [2000] and Hojman and Szeidl [2008]. In Bala and Goyal [2000]'s one-way flow model without decay, the wheel is the unique strict Nash network for some parameters within certain ranges. With local benefits in our model, the wheel is not permissible as agents can never observe others who are more than distance two far away and thus have incentives to form additional links to those who are not observed by them. In Hojman and Szeidl's two-way flow model, the non-empty strict Nash network is unique and exhibits a structure of either periphery-sponsored stars or extended stars. Our model presents that the star, denoted by CP_1, is the unique strict Nash core-periphery network for any $n < 8$. For any $n \geq 8$, Proposition 4.1 shows that there are multiple strict Nash networks which present similar structures, i.e. the core-periphery networks (see also Figure 2 for an illustration).

Further, in the one-way flow case, the strict Nash networks have to feature cycles, implying that each pair of two agents has to be connected by two paths in both directions. In the absence of constraints on the distance of information transmission, the circumference of these cycles is not limited but it requires there exist no crossing cycles. However, in the presence of constraints, the circumference of the cycles is limited to the constraint but it allows the existence of crossing cycles.

5 Stochastically Stable States

Since there are multiple strict equilibria, we are interested in which kinds of network architectures are more likely to be selected in the long run. Previous literature (see e.g. Jackson and Watts 2002a, Feri 2007 and Cui et al. 2013) has established that the best response dynamics with random noise may select a subset of them. For this reason, we consider a best response learning dynamics due to Kandori et al. [1993] and Young [1993]. An agent is randomly selected to renew her strategy at each period t in discrete time, i.e. $t = 0, 1, 2, \cdots$. The selected agent chooses a best response to the strategy profile of other agents at the previous period $t-1$, i.e.

$$g_i(t) \in arg \max_{g_i \in \mathcal{G}_i} U_i(g_i, g_{-i}(t-1))$$

where $g_i(t)$ refers to agent i's strategy at period t, and $g_{-i}(t-1)$ means the strategy profile of other agents except i at period $t-1$. In the case that there are multiple best responses, agents randomly choose one with equal probability.

Given the equation above, the new network configuration in period t only depends on the network in the previous period $t-1$. Technically, this revision process can be defined as a Markov chain on the strategy space $\mathcal{G} \equiv \mathcal{G}_1 \times \mathcal{G}_2 \times \cdots \mathcal{G}_n$. Each network g is a state in this space. An absorbing set is defined as a minimum subset of $\mathcal{G}$ with the property that the dynamics can never leave it once reached. We denote by $G^{\star\star}$ an absorbing set, and $\mathcal{G}^{\star\star}$ denotes the set of all absorbing sets.

First, we provide an important property of $\mathcal{G}^{\star\star}$.

Proposition 5.1. *All networks in $\mathcal{CP}_{\bar{\ell}}$ are absorbing. Each of them forms a singleton absorbing set $G^{\star\star} \in \mathcal{G}^{\star\star}$.*

Proposition 5.1 implies that all strict Nash core-periphery networks are absorbing. The result derives from the fact that in a strict Nash equilibrium, no agent is indifferent between multiple strategies. Thus, all agents are playing their unique best response. As a consequence, they will remain at their current strategy whenever they receive the opportunity to revise. Thus, the revision dynamics can never leave a strict Nash equilibrium without any mistakes, implying that all networks in $\mathcal{CP}_{\bar{\ell}}$ are singleton absorbing sets.

Note that Proposition 5.1 does not provide a full characterisation of all absorbing sets. That is, other absorbing sets may exist, such as strict Nash networks different from core-periphery networks, and a collection of Nash networks among which the dynamics could circulate but never leave.[13] Let a denote the number of all absorbing sets given the population n.

Now, we proceed to characterise the selection among multiple absorbing sets, by adopting the standard techniques developed by Kandori et al. [1993] and Young [1993]. Consider that agents might fail to choose the optimal strategy during the revision process, which we call a mistake. The probability of agents making mistakes is positive, denoted by $\epsilon > 0$. We assume that the agent who makes a mistake chooses randomly among all strategies. Given the positive ϵ, the revision

[13] Despite our best efforts, we have not identified such strict Nash equilibria or such cycles. Therefore. we have not been able to rule out the existence of such absorbing sets.

process is ergodic and aperiodic. The Markov process is therefore irreducible and aperiodic, which means it has a unique stationary distribution $\mu(\epsilon)$. As ϵ goes to zero, $\mu(\epsilon)$ converges to a limited distribution $\mu^\star$, i.e. $\lim_{\epsilon\to 0}\mu(\epsilon) = \mu^\star$. A network g is called stochastically stable if $\mu^\star(g) > 0$. The set of stochastically stable states is defined as $\mathcal{G}^{\star\star\star} \equiv \{g \in \mathcal{G} : \mu^\star(g) > 0\}$.

The following algorithm introduced by Freidlin and Wentzell [1998] and Foster and Young [1990] is used to identify the set of stochastically stable states. Consider two states in different absorbing sets, $g \in G_i^{\star\star}$ and $g' \in G_j^{\star\star}$. Denote by $r(g, g') > 0$ the resistance of transition from g to g', which is the minimum number of mistakes required for this transition. Further, a $G_i^{\star\star}$-tree is defined as a spanning tree rooted in $G_i^{\star\star}$, such that there is a unique path from each other absorbing set to $G_i^{\star\star}$. Denote by τ_i a $G_i^{\star\star}$-tree and T_i denotes the set of all τ_i. The resistance of a $G_i^{\star\star}$-tree is defined as the sum of resistances of its edges, i.e. $r(\tau_i) = \sum_{(g,g')\in\tau_i} r(g', g)$. The stochastic potential of the absorbing set $G_i^{\star\star}$ is defined as the minimum resistance among all τ_i, i.e. $\gamma(G_i^{\star\star}) = arg \min_{\tau_i\in T_i} r(\tau_i)$. Finally, a state in the absorbing set $G_i^{\star\star}$ is stochastically stable if $G_i^{\star\star}$ has the minimum stochastic potential, i.e. $\gamma(G_i^{\star\star}) = \min_{G_j^{\star\star}\in\mathcal{G}^{\star\star}} \gamma(G_j^{\star\star})$. With this technique, we are able to identify a class of stochastically stable states by analysing the relative robustness of absorbing states to mistakes.

Before showing our main results, we exhibit two examples of transitions between the two core-periphery networks, which will play a key role in our analysis.

Example 1. Consider the CP_2 network in Figure 5 (a). Assume that agent 1 makes a mistake and forms additional links to agent 2's periphery agents as shown in Figure 5 (b). Following this, given the revision opportunity, agent 2 will delete the links to agents 7, 8, 9 and 10 as shown in Figure 5 (c). In the next steps, any periphery agent receiving revision opportunities will consequently delete the link to agent 2 as the network, see Figure 5 (d). Thus, with one mistake, we have reached another absorbing state which is a CP_1 network.

Example 2. Now consider a CP_1 network in Figure 6 (a). Assume that agent 2 makes a mistake and forms additional links to agents 7, 8, 9 and 10 as illustrated by Figure 6 (b). In the next step, agent 1 receives the opportunity to revise and find it optimal to delete the links to agents 7, 8, 9, and

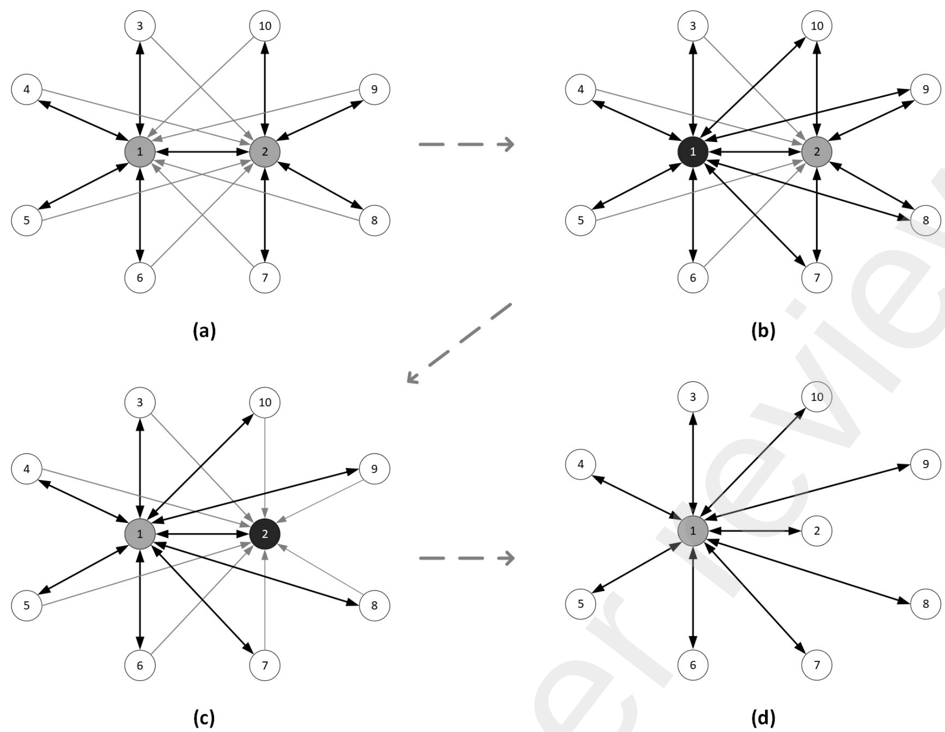


Figure 5: The transition from a CP_2 network to a CP_1 star. The grey circles are core agents, and the white circles are periphery agents. The black circles are the agents who are revising their strategies.

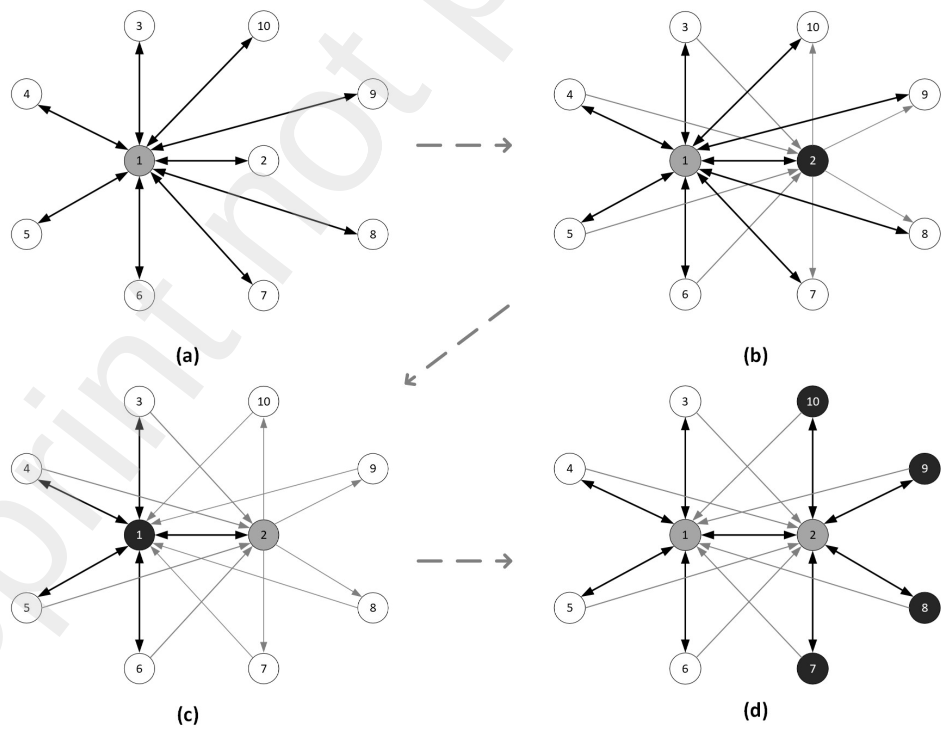


Figure 6: The transition from a CP_1 star to a CP_2 network.

10 as shown in Figure 6 (c). Following this, agents 7, 8, 9 and 10, receiving opportunities to revise will form a link to agent 2, see Figure 6 (d). Thus, with one mistake the dynamics has reached a CP_2 network where agents 1 and 2 are the two core agents.

To predict which kinds of network configurations are stochastically stable, we construct a sequence of absorbing sets where the transition between any two adjacent sets requires one mistake such as the two examples illustrated above. To do so, we first establish that the transition from any absorbing set to a CP_1 network requires one mistake. Second, we show that the transition from a CP_ℓ network to a $CP_{\ell+1}$ network with ℓ common core agents is able at the cost of one mistake. By doing so, we argue that the stochastic potential of each absorbing set is equal to $\boldsymbol{N} - 1$, which is the minimum stochastic potential. Thus, all absorbing sets characterised in Proposition 5.1 are stochastically stable. The following proposition establishes this result.

Proposition 5.2. $\mathcal{CP}_{\bar{\ell}} \subset \mathcal{G}^{\star\star\star}$.

The result that core-periphery networks are stochastically stable is significantly different from Feri [2007], who predicts the periphery-sponsored star for the two-way flow model with decay and Cui et al. [2013], who predicts either the wheel or the empty network for the one-way flow model with decay. Our model implies that there are multiple networks being stochastically stable.

In the last step, we are interested in the welfare properties of stochastically stable states. Welfare of a network g is defined as the sum of payoffs of individuals, i.e. $W(g) = \sum_{i \in N} U_i(g)$. A network g is said to be efficient if and only if $W(g) \geq W(g')$ for all $g' \in \mathcal{G}$. We derive one important property that efficient networks have to fulfil. The lemma proposes that the maximum number of links in any efficient network has to be at most $2 \cdot (n-1)$.

Lemma 4. *A network g is not efficient if the number of links in g exceeds $2 \cdot (n-1)$.*

This lemma follows from the following observation. Note that the welfare of a CP_1 network is given as

$$W(CP_1) = \underbrace{1 + (n-1) - c \cdot (n-1)}_{\text{payoff of agent in core}} + (n-1) \cdot \underbrace{[1 + (n-1) - c]}_{\text{payoff of agent in periphery}}$$
$$= n^2 - 2c \cdot (n-1)$$

Without considering linking costs, n^2 is the highest benefit that can be yielded by a network, where each agent receives benefits from all agents. Thus, if the number of links in a network g is larger than $2 \cdot (n-1)$, the linking costs in g are larger than the linking cost in a CP_1 network, implying that g is not efficient since $W(g) < W(CP_1)$.

Lemma 4 implies that the set of stochastically stable states contains networks that are not efficient. To see this, consider a CP_ℓ network in $\mathcal{CP}_{\bar{\ell}}$. By the definition of core-periphery networks, the number of links between core agents is $2 \cdot (\ell - 1)$. The number of links from core agents to periphery agents is $n - \ell$ as each periphery agent is linked by only one core agent. Further, the number of links from periphery agents to core agents is $(n - \ell) \cdot \ell$ as each periphery agent links up to all core agents. We thus have that the number of links in a CP_ℓ network is

$$f(\ell, n) = 2 \cdot (\ell - 1) + n - \ell + (n - \ell) \cdot \ell = -\ell^2 + (n+1) \cdot \ell + n - 2.$$

This function $f(\ell, n)$ is increasing provided that $\ell < \frac{n+1}{2}$. Note that any CP_ℓ network that is a strict Nash equilibrium and further stochastically stable state requires that each core agent forms links to at least three periphery agents. Thus given the number of agents n, the number of core agents fulfils that $\ell \leq \lfloor \frac{n}{4} \rfloor < \frac{n+1}{2}$. Hence, the number of links in any $g \in \mathcal{CP}_{\bar{\ell}}$ increases with the number of core agents. Thus we have that $W(CP_1) > W(CP_2) > \cdots > W(CP_{\bar{\ell}})$.[14] Therefore, the set of stochastically stable states contains network configurations that are not efficient.

The implication is that when benefits are global, as in Bala and Goyal [2000], the wheel network is efficient since the wheel exhibits a structure where every agent can use one link to get access to

[14]Whilst we have been able to use numerical calculations to show that CP_1 networks are efficient for $n = 5$ and 6, we have not been able to provide a general result showing that this for arbitrary n.

all other agents. In contrast with local benefits, to observe other agents who are located two steps away, forming additional links is necessary. This results in an increase in the number of links in the network, causing a reduction in welfare. Therefore, network structures may arise in the long run which are dominated by others in terms of welfare.

6 Conclusion

In this paper, we explored a non-cooperative framework for modelling the process of network formation, where agents unilaterally form costly links. The payoffs accrue to agents who initiate the links. We focus on investigating how constraints on information transmission affect stable networks in the long run.

In contrast to the conventional results of either the wheel or the periphery-sponsored star being the unique strict Nash network or stochastically stable, we reveal that core-periphery networks are strict Nash equilibria when agents can only receive information from their neighbours and the neighbours of their neighbours. This finding sheds light on the diverse range of equilibrium structures that can emerge in the context of network formation.

Additionally, because of the multiplicity of strict Nash networks, we study the selection between multiple equilibria in a perturbed best response learning dynamics, to find which kind of networks are stochastically stable. We show that the set of stochastically stable states encompasses multiple network configurations that exhibit the structure of the core-periphery networks. Surprisingly, our analysis reveals that the set of stochastically stable states includes network configurations that are inefficient from a welfare perspective.

Acknowledgments:

I wish to thank my supervisor, Simon Weidenholzer, for his constant help. I am also indebted to Friederike Mengel, Ennio Bilancini, Albin Erlanson, Aditya Kuvalekar, Fernando Vega-Redondo, and to the participants at the 10th CMGTA Conference, the 2024 RES Annual Conference, and

SUNBELT 2025 for their helpful comments.

Financial Support:

I acknowledge financial support from the University of Essex Social Sciences Doctoral Scholarship and the Swiss National Science Foundation, No.10000259.

References

Allouch, N. [2015], 'On the private provision of public goods on networks', Journal of Economic Theory **157**, 527–552.

Bala, V. and Goyal, S. [2000], 'A noncooperative model of network formation', Econometrica **68**(5), 1181–1229.

Billand, P., Bravard, C. and Sarangi, S. [2008], 'Existence of nash networks in one-way flow models', Economic Theory **37**(3), 491–507.

Billand, P., Bravard, C. and Sarangi, S. [2011], 'Strict nash networks and partner heterogeneity', International Journal of Game Theory **40**(3), 515–525.

Borgatti, S. P. and Everett, M. G. [2000], 'Models of core/periphery structures', Social Networks **21**(4), 375–395.

Bramoullé, Y. and Kranton, R. [2007], 'Public goods in networks', Journal of Economic Theory **135**(1), 478–494.

Calvo-Armengol, A. and Jackson, M. O. [2004], 'The effects of social networks on employment and inequality', American Economic Review **94**(3), 426–454.

Calvó-Armengol, A. and Jackson, M. O. [2007], 'Networks in labor markets: Wage and employment dynamics and inequality', Journal of Economic Theory **132**(1), 27–46.

Cui, Z., Wang, S., Zhang, J. and Zu, L. [2013], 'Stochastic stability in one-way flow networks', Mathematical Social Sciences **66**(3), 410–421.

Cui, Z. and Weidenholzer, S. [2021], 'Lock-in through passive connections', Journal of Economic Theory **192**, 105187.

De Jaegher, K. and Kamphorst, J. [2015], 'Minimal two-way flow networks with small decay', Journal of Economic Behavior & Organization **109**, 217–239.

Elliott, A., Chiu, A., Bazzi, M., Reinert, G. and Cucuringu, M. [2020], 'Core–periphery structure in directed networks', Proceedings of the Royal Society A **476**(2241), 20190783.

Feri, F. [2007], 'Stochastic stability in networks with decay', Journal of Economic Theory **135**(1), 442–457.

Fosco, C. and Mengel, F. [2011], 'Cooperation through imitation and exclusion in networks', Journal of Economic Dynamics and Control **35**(5), 641–658.

Foster, D. and Young, P. [1990], 'Stochastic evolutionary game dynamics', Theoretical Population Biology **38**(2), 219–232.

Freidlin, M. I. and Wentzell, A. D. [1998], Random perturbations, Springer.

Galeotti, A. and Goyal, S. [2010], 'The law of the few', American Economic Review **100**(4), 1468–92.

Goyal, S. and Joshi, S. [2003], 'Networks of collaboration in oligopoly', Games and Economic Behavior **43**(1), 57–85.

Goyal, S. and Moraga-Gonzalez, J. L. [2001], 'R& d networks', Rand Journal of Economics **32**(4), 686–707.

Goyal, S. and Vega-Redondo, F. [2005], 'Network formation and social coordination', Games and Economic Behavior **50**(2), 178–207.

Goyal, S. and Vega-Redondo, F. [2007], ‘Structural holes in social networks’, Journal of Economic Theory **137**(1), 460–492.

Guare, J. [2016], Six degrees of separation, in ‘The Contemporary Monologue: Men’, Routledge, pp. 89–93.

Hojman, D. A. and Szeidl, A. [2008], ‘Core and periphery in networks’, Journal of Economic Theory **139**(1), 295–309.

Jackson, M. O. and Watts, A. [2002a], ‘The evolution of social and economic networks’, Journal of Economic Theory **106**(2), 265–295.

Jackson, M. O. and Watts, A. [2002b], ‘On the formation of interaction networks in social coordination games’, Games and Economic Behavior **41**(2), 265–291.

Jackson, M. O. and Wolinsky, A. [1996], ‘A strategic model of social and economic networks’, Journal of Economic Theory **71**(1), 44–74.

Kandori, M., Mailath, G. J. and Rob, R. [1993], ‘Learning, mutation, and long run equilibria in games’, Econometrica **61**(1), 29–56.

Milgram, S. [1967], ‘The small world problem’, Psychology Today **2**(1), 60–67.

Price, D. J. D. S. [1965], ‘Networks of scientific papers.’, Science **149**(3683), 510–515.

Staudigl, M. and Weidenholzer, S. [2014], ‘Constrained interactions and social coordination’, Journal of Economic Theory **152**, 41–63.

Watts, A. [2001], ‘A dynamic model of network formation’, Games and Economic Behavior **34**(2), 331–341.

Young, H. P. [1993], ‘The evolution of conventions’, Econometrica **61**(1), 57–84.

A Appendix

A.1 Proofs in Section 4

Proof of Lemma 1. Suppose g is a non-empty strict Nash network. Thus, there exists one agent i who maintains at least one link. Suppose now there exists one agent j who supports no links. Consider the payoffs received by agents i and j. If $U_i(g) > 1$, agent j is strictly better off by replicating i's links; if $U_i(g) < 1$, agent i is better off by deleting all her links; if $U_i(g) = 1$, agent j is indifferent between maintaining no links and replicating i's links. All cases contradict g being a strict Nash network. This implies that there exists no agent who supports no links. Thus, every agent supports at least one link in a non-empty strict Nash network. □

Proof of Lemma 2. Given a non-empty strict Nash network g, note that there can exist at most one agent without any incoming links. To see this, assume that there exist two agents without incoming links, denoted by i and j. If $U_i(g) < U_j(g)$, i is strictly better off by replicating j's links. Analogously, if $U_i(g) > U_j(g)$, j is strictly better off by replicating i's links. Otherwise, if $U_i(g) = U_j(g)$, either agent i or j is indifferent between maintaining her current links and replicating the other's links. All cases contradict g being a strict Nash network. Thus, there cannot exist more than one agent without any incoming links.

In the next step, we show that in fact, there cannot exist a single agent without any incoming links. To see this, assume an agent i has no incoming links. First, consider the case where there is an agent k, such that $U_k(g) > U_i(g)$. Note that payoffs of all other agents are independent of i's linking strategy. So by deleting her current links and forming the links that k supports, agent i can assure herself the same payoff as k gets, contradicting the assumption that g is a strict Nash equilibrium.

Second, we consider the case where there exists an agent k whose payoff is equal to i's, i.e. $U_k(g) = U_i(g)$. In the first sub-case $g_i \neq g_k$, i is indifferent between replicating k's links and maintaining her current links, contradicting g being a strict Nash network. Then consider the second sub-case $g_i = g_k$. According to Lemma 1, k has at least one incoming link. Suppose that agent ℓ is the agent who forms a link to k. Since i and k have the same linking strategies, ℓ is indifferent between linking to k and to i, which also contradicts g being a strict Nash equilibrium. Thus, it is impossible to have an agent k whose payoff is equal to i's.

Third, we consider the case where agent i's payoff is the highest among all agents, i.e. $U_i(g) > U_k(g), \forall k \in N$. In the first sub-case, there are some agents from whom the distance to agent i is larger than two, i.e. $\exists j \in N, d(i,j;g) > 2$. As a result of the restriction on information transition, agent i's payoff is independent of j's linking strategy. Agent j is strictly better off by replicating i's links, which contradicts g being a strict Nash equilibrium. In the second sub-case, the furthest distance from any agent to i is two. There exists an agent ℓ with $d(i,\ell;g) = 2$. Sort i's $1-$ neighbours by the number of links they form. Without loss of generality, rename them $i_1, i_2, \cdots, i_{n_i^1}$, where the number of their links weakly increases with subscripts, i.e. $n_{i_1}^1 \leq n_{i_2}^1 \leq \cdots \leq n_{i_{n_i^1}}^1$. We can rewrite i's payoff function in the following way:

$$U_i(g_i, g_{-i}) = 1 + \sum_{x=1}^{n_x^1}(1 + n_{i_x}^1 - c) \tag{2}$$

Analogously, sort ℓ's $1-$ neighbours by the numbers of links they form and without loss generality, rename them $\ell_1, \ell_2, \cdots, \ell_{n_\ell^1}$, where the number of their links weakly increases with subscripts, i.e. $n_{\ell_1}^1 \leq n_{\ell_2}^1 \leq \cdots \leq n_{\ell_{n_\ell^1}}^1$. Agent ℓ's payoff function can be written as:

$$U_l(g_\ell, g_{-\ell}) = 1 + \sum_{x=1}^{n_\ell^1}(1 + n_{\ell_x}^1 - c) \tag{3}$$

There are two sub-subcases:

i) The number of ℓ's links is at least as large as i's, i.e. $n_\ell^1 \geq n_i^1$. Since $U_i(g) > U_\ell(g)$, the largest number of links that agents in $N_i^1(g)$ form must be larger than the smallest number of links that agents in $N_\ell^1(g)$ form, i.e. $n_{i_{n_i^1}}^1 > n_{\ell_1}^1$. [15] Then, agent ℓ is strictly better off by deleting the link

[15] Consider an inequality $a_1 + a_2 + \cdots + a_n > b_1 + b_2 + \cdots + b_{n'}$, with $a_1 \leq a_2 \leq \cdots \leq a_n$, $b_1 \leq b_2 \leq \cdots \leq b_{n'}$ and $n \leq n'$. Since $n \cdot a_n > a_1 + a_2 + \cdots + a_n > b_1 + b_2 + \cdots + b_{n'} > n' \cdot b_1 > n \cdot b_1$, it must be true that $a_n > b_1$ must be true.

to ℓ_1 and replicating i's link to $i_{n_i^1}$, which contradicts g being a strict Nash equilibrium. [16]

ii) Agent ℓ forms less links than i does, i.e. $n_\ell^1 < n_i^1$. Again there are two sub-cases.

a) If $n^1_{\ell_{n_\ell^1}} \geq n^1_{i_1}$, then i is strictly better off by deleting the link to i_1 and replicating ℓ's link to $\ell_{n_\ell^1}$. [17]

Now consider the second case b).

b) If $n^1_{\ell_{n_\ell^1}} < n^1_{i_1}$, then ℓ is better off by deleting the link to ℓ_1 and replicating i's link to $i_{n_i^1}$.

Thus, if there exists a single agent i such that $U_i(g) > U_k(g), \forall k \in N$, it follows that the distance from any agent to agent i is less than one, i.e. $d(i,k;g) \leq 1, \forall k \neq i$. In other words, agent i links up to all other agents, i.e. $k \in N_i^1, \forall k \neq i$. According to Lemma 1, note that every agent must have at least one outgoing link, and by assumption, there exists no agent linking to i. It follows that $\exists m, n \in N_i^1, m \neq n$, such that m links up to n, or vice versa. This contradicts g being a strict Nash network since i is strictly better off by deleting the link to n.

Thus, it follows that it is impossible to have a single agent without incoming links. Consequently, every agent has at least one incoming link. □

Proof of Lemma 3. The proof proceeds by contradiction. Consider a non-empty strict Nash network g that is not strongly connected. Assume that there exists a non-empty network $g \in \mathcal{G}$ which is a strict Nash network but not strongly connected. So there are multiple strongly connected components $\{C_1, C_2, \cdots, C_m\}$ with $m \geq 2$. Without loss of generality, we consider the case where $m = 2$. [18]. Let g^1 and g^2 be the two sets of strategies used by agents in C_1 and C_2, respectively. Note that g consists of several strongly connected components. Without loss of generality, consider any two strongly connected components C_1 and C_2. Lemma 1 and Lemma 2 imply that both C_1 and C_2 contain multiple agents. Then there are two cases:

i) C_1 and C_2 are separated. Consider two agents $i, m \in C_1$ and two agents $j, k \in C_2$. We

[16] Even though agents i and ℓ may have some common neighbours, we can always find an agent i_x who is not in $N_\ell^1(g)$ and forms more links than ℓ_1. To see this, consider the inequality $a_1 + a_2 + \cdots + a_n > b_1 + b_2 + \cdots + b_{n'}$. Agents i and ℓ have common neighbours, implying that there exists at least one pair of a_i and b_j, such that $a_i = b_j$. We can eliminate a_i and b_j on both sides and still apply the property above.

[17] Analogously, even if i and ℓ have common neighbours, following an argument similar to the one in the footnote 16, we can always find an agent ℓ_x that $n^1_{\ell_x} \geq n^1_{i_1}$.

[18] For any $m \geq 3$, we can start with any two of the components and iteratively apply the same logic as with $m = 2$.

assume that m forms a link to i, i.e. $i \in N_m^1(g)$ and k forms a link to j, i.e. $j \in N_k^1(g)$. Since g is a strict Nash network, the payoff of m's link to i is positive and no larger than $1 + n_i^1 - c$. [19] Therefore, it must be true that $1+n_i^1-c > 0$. Then, consider agent k in C_2. Since k does not receive benefit from any agent in C_1, she is strictly better off by forming a link to i since $1 + n_i^1 - c > 0$. It contradicts the assumption that g is a strict Nash network. Now consider the link from k to j. By the same logic, m is strictly better off by forming a link to agent j.

ii) There exist paths from agents in C_1 to agents in C_2, but not vice versa. Following a similar argument in the case above, an agent in C_2 is strictly better off by forming a link to an agent in C_1.

Therefore, C_1 and C_2 are strongly connected and contained in one strongly connected component. Consequently, all strongly connected components are strongly connected. The network g is strongly connected. □

Proof of Proposition 4.1. Consider a core-periphery network CP_ℓ in which $|P_i| \geq 3$ for any $i \in C(\ell; g)$. A core agent i's payoff is given by

$$U_i(g_i, g_{-i}) = N - c \cdot \sum_{j \in N} g_{ij} = N - c \cdot (\ell + |P_i|)$$

Notice that agent i

i) has no incentives to form more links. Since choosing g_i already allows i to receive benefits from every other agent, adding more links increases i's cost, but her benefit remains the same.

ii) has no incentives to delete any link. If i deletes one link to her periphery j, she would reduce her cost by c but lose the benefit from j. Since linking costs $c < 1$, i's payoff would decrease by $1 - c$. If i deletes one link to another core agent k, she would at least lose the benefits from k's periphery agents. Agent i's payoff would decrease by at least $|P_k| - c$. In both cases, player i is worse off by deleting a link.

[19] It is possible that m links to agents in $N_i^1(g)$ directly or via other paths

Now, consider a periphery agent j, her payoff is given by

$$U_j(g_j, g_{-j}) = N - c \cdot \sum_{k \in N} g_{jk} = N - c \cdot \ell$$

we argue in the following that agent j

i) has no incentives to delete any link to core agents. If j deletes one link to a core agent i, she would lose the benefit from i's periphery agents. The payoff of j would decrease by $|P_i| - c$. Note that $|P_i| \geq 3$ and $c < 1$. Agent j is worse off by deleting a link to the core agent.

ii) has no incentives to form any link to other periphery agents. Since by linking to all core agents, j receives benefits from every agent, adding more links only increases j's cost.

Moreover, we consider a case where no agent has incentives to replace any one of her links with another. To see this, replacing one link can be divided into two steps: deleting a link and forming a new link. Note that every agent has already been linked to all core agents. An agent is unable to choose to form a new link to any other core agent since she has already linked to all core agents. Thus, there are two sub-cases for a core agent i to discuss.

a) Deleting a link to a core agent k and forming a new link to a periphery agent j. The first step reduces i's payoff by at least $|P_k| - c$. The second step can increase i's benefit by at most $1 - c$.[20] Overall, i's payoff is reduced by $|P_k| - 1$. Since $|P_k| \geq 3$, agent i is worse off by doing so.

b) Deleting a link to a periphery agent and forming a new link to another periphery agent. The first step reduces i's payoff by $1 - c$. The second step reduces i's payoff by c since she can observe every other periphery agent via other core agents. Thus, i's payoff decreases by $1 - c + c = 1$.

In both sub-cases, i is worse off. Thus, i doesn't have any incentives to replace any of her links. Note that periphery agents form no links to other periphery agents. The discussion of link replace-

[20] If j is k's periphery agent, forming the link to j yields $1 - c$. If j is the periphery agent of another core agent, this link only costs c since she already receives the benefit from j via another core agent.

ment of a periphery agent can focus only on the sub-case a).

To conclude, neither core agents nor periphery agents have incentives to change their links. □

A.2 Proofs in Section 5

Proof of Proposition 5.1. Consider a network $g \in \mathcal{CP}_{\bar{\ell}}$. Note that all networks in $\mathcal{CP}_{\bar{\ell}}$ are strict Nash networks. Any agent i revising her strategy will remain at her current strategy since $U_i(g_i, g_{-i}) > U_i(g'_i, g_{-i})$ for all g'_i. Thus, without any mistakes, the network will remain unchanged. □

Proof of Proposition 5.2.

First, consider the transition from any state $g \in \mathcal{CP}_{\bar{\ell}}$ to a state $g' \in \mathcal{CP}_1$ with agent i as the unique core agent. One mistake is sufficient for this transition, i.e. $r(g, g') = 1$. To see this, assume that agent i makes a mistake and forms links to all other agents. Now, consider other agents who get the chance to revise. Forming a single link to agent i is a best response for them. Consequently, the dynamics reaches the CP_1 network with i as the unique core agent.

Then, consider the transition from $g \in \mathcal{CP}_1$ to $g' \in \mathcal{CP}_2$ where the core agent in g is still a core agent in g'. One mistake is sufficient for this transition. To see this, consider the case where a periphery agent j makes a mistake and forms three extra links to other periphery agents of core agent i. Now, given the revision opportunity, the core agent i will find it optimal to delete links to those periphery agents of j. Following this, give the revision opportunity to other periphery agents. Their best response is forming one additional link to agent j to get access to the periphery agents of j. Consequently, the dynamics reaches a CP_2 network with i and j as two core agents.

Now, consider two states $g \in \mathcal{CP}_\ell$ and $g' \in \mathcal{CP}_{\ell+1}$ where they have ℓ common core agent, i.e. $C(\ell; g) \subset C(\ell + 1; g')$. The resistance of the transition from g to g' is also one. Following a similar argument as above, a periphery agent j makes a mistake and links to three periphery agents. Consider core agents who link to the three periphery agents. They are indifferent or have a profitable deviation by deleting the links to these periphery agents and forming a link to j. Following this, consider other core agents and periphery agents. Their best response is to form a link to j.

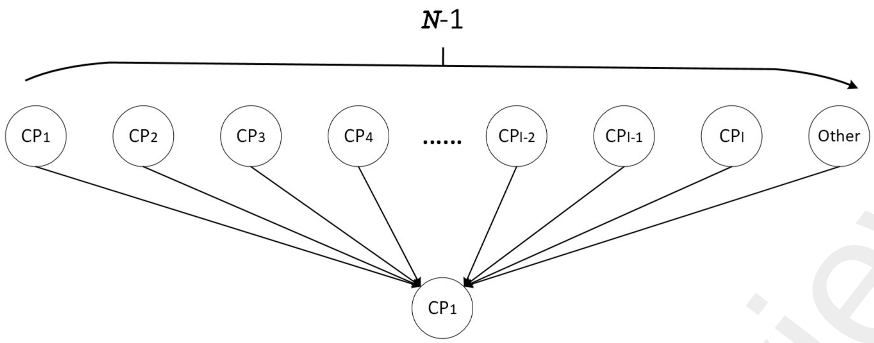


Figure 7: The structure of a $G_i^{\star\star}$-tree formed by $g \in \mathcal{CP}_1$.

Consequently, the dynamics reaches a state in $\mathcal{CP}_{\ell+1}$ with j as the $\ell + 1$ core agent. The cost of this transition is one, i.e. $r(g, g') = 1$.

Now, consider the transition from a state g in any $\mathcal{CP}_\ell$ to g' in any $\mathcal{CP}_{\ell'}$. Note that g and g' are different in two dimensions. First, the numbers of core agents are different, i.e. $\ell \neq \ell'$. Second, the identities of core agents vary, i.e. $C(\ell; g) \neq C(\ell'; g')$. To identify the resistance of this transition, we conduct a path of states $\{s_0, s_1, s_2, \cdots, s_{\ell'-1}, s_{\ell'}\}$ with $s_0 = g$ and $s_{\ell'} = g'$, which are characterized by the following properties: $s_k \in \mathcal{CP}_k$ such that $C(k-1; s_{k-1}) \subset C(k; s_k) \subseteq C(\ell'; g')$, and $P_j(s_i) = P_j(g')$ for each j in the core, for all $k = 2, 3, \cdots, \ell'$.

(i) $s_1 \in \mathcal{CP}_1$ such that $C(1; s_1) \subseteq C(\ell'; g')$, i.e. the core agent in s_1 is also a core agent in g';

(ii) $s_k \in \mathcal{CP}_k$ such that $C(k-1; s_{k-1}) \subset C(k; s_k) \subseteq C(\ell'; g')$, and $P_j(s_i) = P_j(g')$ for each j in the core, for all $k = 2, 3, \cdots, \ell'$.

As argued above, the resistance of the transition from s_i to s_{i+1} is one. Thus, the resistance of this path is equal to the sum of resistances, i.e. $r(g, g') = \sum_{i=1}^{\ell'} r(s_{i-1}, s_i) = \ell'$.

Now we move on to characterise the set of stochastically stable states. Note that the number of all absorbing sets is $\boldsymbol{N}$. Recall that a $G_i^{\star\star}$-tree consists of edges that connect every absorbing set. Thus, the number of edges of every permissible $G_i^{\star\star}$-tree is $\boldsymbol{N} - 1$. Note that the stochastic potential of an absorbing set $G_i^{\star\star}$ is defined as the minimum sum of resistances of edges of the $G_i^{\star\star}$-tree. The minimum is obtained when the resistance of every edge of the $G_i^{\star\star}$-tree is one.

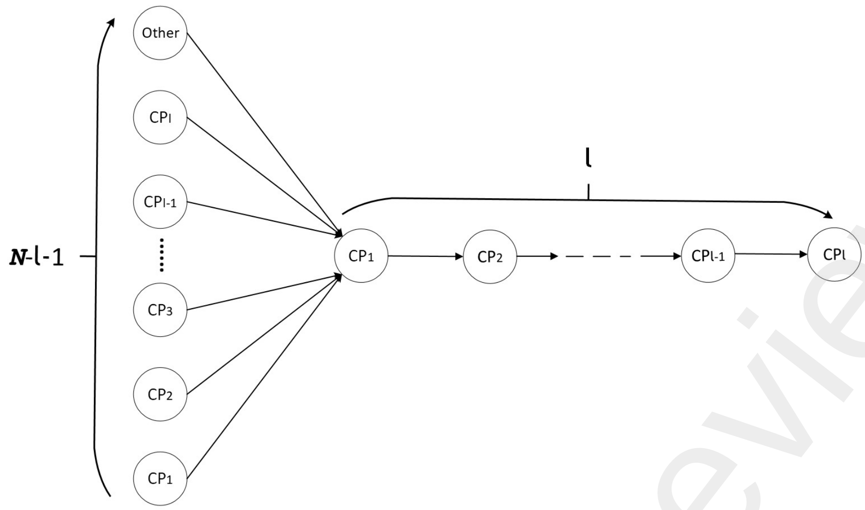


Figure 8: An sketch structure of a $G_j^{\star\star}$-tree with any $g \in \mathcal{CP}_\ell$.

Following this, we move forward to calculate the stochastic potential of each absorbing set characterised by Proposition 5.1. First, consider the stochastic potential of absorbing set $G_i^{\star\star}$ formed by a $g \in \mathcal{CP}_1$. Consider the $G_i^{\star\star}$-tree where the transition from each absorbing set to $G_i^{\star\star}$ is direct as the structure depicted in Figure 7. Note that as we argued above, the resistance from any state to a CP_1 network is one. The stochastic potential of this $G_i^{\star\star}$-tree is given by the sum of resistances of all edges, i.e. $\gamma(G_i^{\star\star}) = \boldsymbol{N} - 1$.

Next, consider the absorbing sets $G_\ell^{\star\star}$ formed by a state $g \in \mathcal{CP}_\ell$. We first conduct a sequence of absorbing sets $\{G_1^{\star\star}, G_2^{\star\star}, \cdots, G_{\ell-1}^{\star\star}, G_\ell^{\star\star}\}$ with $G_j^{\star\star}$ formed by a state s_i in $\mathcal{CP}_j$, $j = 1, 2, \cdots, \ell$, where s_i holds the same property as above. Then, for those absorbing sets not included in this sequence, consider the transition from them to $G_1^{\star\star}$ directly. Figure 8 depicts such a $G_\ell^{\star\star}$-tree. According to the above construction, the resistance of transition from $G_i^{\star\star}$ to $G_{i+1}^{\star\star}$ is one, for any $i = 1, 2, \cdots, \ell - 1$. Further, from other $\boldsymbol{N} - \ell - 1$ absorbing sets, one mistake is sufficient from each absorbing set to $G_1^{\star\star}$. Thus, the sum of resistance of this $G_j^{\star\star}$-tree is $\ell + \boldsymbol{N} - \ell - 1$, i.e. the stochastic potential of this $G_j^{\star\star}$-tree is $\gamma(G_j^{\star\star}) = \boldsymbol{N} - 1$.

Since $\boldsymbol{N} - 1$ is the minimum stochastic potential, for any absorbing state $g \in \mathcal{CP}_{\bar{\ell}}$ is thus stochastically stable according to Kandori et al. [1993] and Young [1993]. □