

## Dynamic Network Formation with Incomplete Information

Yangbo Song · Mihaela van der Schaar

Received: / Accepted:

**Abstract** How do networks form and what is their ultimate topology? Most of the literature that addresses these questions assumes complete information: agents know in advance the value of linking even with agents they have never met and with whom they have had no previous interaction (direct or indirect). This paper addresses the same questions under the much more natural assumption of incomplete information: agents do not know in advance – but must learn – the value of linking. We show that incomplete information has profound implications for the formation process and the ultimate topology. Under complete information, the network topologies that form and are stable typically consist of agents of relatively high value only. Under incomplete information, a much wider collection of network topologies can emerge and be stable. Moreover, even with the same topology, the locations of agents can be very different: an agent can achieve a central position purely as the result of chance rather than as the result of merit. All of this can occur even in settings where agents eventually learn everything so that information, although initially incomplete, eventually becomes complete. The ultimate network topology depends significantly on the formation history, which is natural and true in practice, and incomplete information makes this phenomenon more prevalent.

**Keywords** Network formation · Incomplete information · Dynamic network formation · Link formation · Formation history · Externalities

**JEL Classification** A14 · C72 · D62 · D83 · D85

---

We thank two anonymous referees for useful comments. We are also grateful to William Zame, Ichiro Obara, Moritz Meyer-ter-Vehn, Sanjeev Goyal, Simpson Zhang, Onur Atan and a number of seminar audiences for suggestions which have improved the paper. This work was supported by the National Science Foundation (NSF) under Grant CCF-1218136.

---

Yangbo Song  
Department of Economics, UCLA, Los Angeles, CA 90095, USA  
Tel.: +1-424-298-0463  
E-mail: darcy07@ucla.edu

Mihaela van der Schaar  
Department of Electrical Engineering, UCLA, Los Angeles, CA 90095, USA  
Tel.: +1-310-825-5843  
E-mail: mihaela@ee.ucla.edu

## 1 Introduction

How do social and economic networks form and what is their ultimate shape (topology)? This important question is addressed in a substantial literature which begins with Jackson and Wolinsky (1996) and continues with Bala and Goyal (2000), Jackson and Watts (2002), Ballester et al (2006), and others. The central conclusion of this literature is that only special shapes of networks can occur and persist. However, this literature makes the strong assumption of *complete information*: agents know *in advance* the value of linking to any other agents – even agents they have never met and with whom they have had no previous interaction, direct or indirect. (Most of the literature on dynamic network formation, for instance Bala and Goyal (2000) and Watts (2001), also assumes that agents have complete information about the entire history of link formation at every moment in time.) The present paper addresses the same questions under what seems to us to be the much more natural assumption of *incomplete information*: agents do not know in advance – but must *learn* – the values of linking to agents they have never met<sup>1</sup>. As is usual in environments of incomplete information, agents begin only with *beliefs* about the values of linking to other agents, make choices on the basis of their beliefs, and update their beliefs (learn the true values) on the basis of their experience (history).

We show that the assumption of incomplete information has profound implications for the process of network formation, the shape of the networks that ultimately form and persist, and the location of various agents within the network. Much of the literature that assumes complete information shows that the only networks that can form and persist have a star or core-periphery form, with agents of relatively high value in the core. By contrast, when information is incomplete, we show that a much larger variety of networks and network shapes can form and persist: frequently a strict superset of the set of networks that can form and persist when information is complete. This phenomenon occurs because, due to incomplete information, some unattractive (“low-value”) agent may get linked and even secure a high connectivity degree in an early stage of the formation process, and then would remain connected thereafter since it offers other agents indirect access to numerous “high-value” agents. Under complete information, this “connection by mistake” never happens. However, these connections by mistake which happen only in the incomplete information setting may be often socially valuable: the ultimate network that forms and persists when information is initially incomplete may often yield higher social welfare than any network that can form and persist when information is initially complete. We stress that all of this can occur even in settings where agents eventually learn *everything* so that information, although initially incomplete, eventually becomes complete. Incompleteness of information may eventually disappear but its influence may persist forever.

To make these points we adopt precisely the same framework as in Watts (2001) except that we assume that information is initially incomplete. As usual, agents begin with common prior beliefs about the types of other agents but learn these types over time by forming links (which might later be broken) in a process where they are randomly selected to take action. Moreover, even when the network shapes that form are the same or similar as in the complete information case, the locations of agents within the network can be very different. For instance, when information is incomplete, it is possible for a star network with a “low-value” agent in the center to form and persist indefinitely; thus, an agent can achieve a central position purely as the result of chance rather than as the result of merit. Perhaps even more strikingly, when information is incomplete, a connected network can form and persist even if, when information were complete, *no links* would ever form so that the final form would be a totally disconnected network.

However, the most important consequence of incomplete information is not that a larger variety of network shapes (topologies) might emerge, but that the particular shape that *does* emerge depends on the history of link formation and of link formation opportunities. For instance, when information is incomplete, agents  $i, j$  might choose to form a link because each *expects* the value of the link to exceed the cost of forming it. Having formed the link, the agents may learn that their expectations were wrong and so might wish to sever it. Nevertheless, before the agents have the opportunity to sever the link, each of them may have formed *other* links, so that the *indirect* value of the link between  $i$  and  $j$  – the value of the connection to *other* agents – may be sufficiently large that they prefer to maintain the link between them after all; as a result, even when all information is eventually revealed, the link persists<sup>2</sup>. However, whether these other links have formed will depend not only on the values of those links but also on the random opportunities presented to form them.

<sup>1</sup> We emphasize that we study learning *during the network formation process*, rather than learning in an exogenously given and fixed network structure. For a study on the latter, see for example Acemoglu et al (2011) .

<sup>2</sup> Indirect linking provides a positive externality when information is complete as well, but the effect is completely different:  $i, j$  know exactly the value of the (potential) link between them, so they will never form a link which either may regret forming.

## 2 Literature Review

The conclusions about network formation and stability that can be found in the existing literature depend both on the process of formation and the notion of stability. Our paper is most closely related to Watts (2001), which is an extension of Jackson and Wolinsky (1996). The focus of Watts (2001) is to analyze the formation of networks in a dynamic framework, where agents are homogeneous (hence information is complete in the strongest form), link formation is *undirected* and requires bilateral consent, and the pair of agents selected to update their potential link follows a given stochastic process. This model predicts that the network topology that ultimately form and become stable must be either *empty* or *connected*. We adopt exactly the same network formation game and the same notion of pairwise stability (as in Jackson and Wolinsky (1996) as well), but we assume instead that agents are of different types and thus connecting to them results in *different (heterogeneous)* payoffs; moreover, there is *incomplete information*: an agent does *not* know the types of agents that he has never connected with, but he is able to form *beliefs* based on which he chooses the optimal action. Networks which result under this, more realistic, assumption are strikingly different from those obtained in the model assuming homogeneous agents and complete information. On one hand, connectedness is no longer a key property of stable non-empty networks – the formation process converges to connected networks in some range of parameters and to non-empty networks with singleton agents in other ranges. On the other hand, even if the network stays empty forever under complete information, a non-empty, even connected network may emerge and be stable with positive probability under incomplete information.

Jackson and Watts (2002) study the same dynamic network formation process and applies the same notion of stable network as above, but their focus is to characterize the set of stable networks when there is a tremble in link formation, i.e. there is a small probability that a link is automatically deleted or added after the agents take their action. In contrast, our paper emphasizes the strategic interaction among agents along the network formation process, and focuses on how agents' belief affects their optimal behavior and thus the network topology.

The renowned paper by Bala and Goyal (2000) is close to our paper in motivation, as they also analyze the dynamics in network formation under the assumption of homogeneous agents. However, their model is based on *directed* link formation, such that the action of one individual to link to another neither requires the second individual's consent nor incurs any cost to the second individual. In other words, new links can be formed unilaterally. Moreover, instead of allowing only one pair of agents to meet and interact in one period as in Watts (2001), they assume that all agents move simultaneously in each period and play some Nash equilibrium. Such differences in modelling lead to fundamentally different theoretical results and applications.

In the literature studying how agent heterogeneity affects network formation, such as Haller and Sarangi (2003), Galeotti (2006), Galeotti et al (2006), Galeotti and Goyal (2010) and Zhang and van der Schaar (2012)Zhang and van der Schaar (2013), even though the theoretical frameworks vary from one to another, *complete information* is still a common assumption, and the predictions in the above papers are often restricted to a few, specific, types of network topologies such as stars, wheels or core-periphery networks with "high-value" or low-cost agents enjoying higher connectivity than others. We differ from these works in three aspects. First, as mentioned above, agents have *no precise knowledge* about their exact payoffs due to *incomplete information*; instead, they choose an optimal action according to their *beliefs* about the payoffs that they will obtain from connecting to others. Secondly, we show that the interaction of incomplete information and agent heterogeneity produces a *much wider* range of network topologies, which includes stars, wheels, core-periphery networks, etc. Finally, the topology that emerges and becomes stable strongly depends on the *formation history*: an agent may exhibit a high degree of connectivity in equilibrium not necessarily because he is of a special (high) type but also because initially he was fortunate to obtain sufficiently many links by chance, which in turn attracted others to form and maintain links with him due to the large indirect benefits that he can offer. Therefore, unlike most existing literature that only emphasizes *what* topologies can be formed, we argue that *how* a certain topology comes into being is equally important.

Among the empirical literature on network formation games, works such as Falk and Kosfeld (2003), Corbae and Duffy (2008), Goeree et al (2009) and Rong and Houser (2012) have conducted experimental studies on the types of emerging topologies. The experimental results indicate that (1) typical equilibrium network topologies predicted by the existing theoretical analysis are *not* always consistent with the empirical observations; especially, stars are formed only in *a fraction* of the total number of experiments conducted (see Corbae and Duffy (2008), Rong and Houser (2012)) and such fractions, under some treatments such as a two-way flow of payoffs, are rather low (see Falk and Kosfeld (2003)); (2) even in the experiments where equilibrium network topologies do emerge with high frequency, such topologies are *developed* rather than *born* (see Goeree et al (2009)), which we believe suggests a dynamic network formation process of a sufficiently long duration as a more appropriate environment for stable networks to emerge, compared with a static one. Moreover, the study of networks which actually get formed in large social communities, as presented for instance by Mele (2010) and Leung (2013), shows that in

environments where agents are heterogeneous and withhold certain private information in their payoffs from links, numerous phenomena which are not predicted by the existing theoretical literature (such as multiple components and clustering of agents with different attributes) can happen. We do not claim to explain all these phenomena in this paper since the way to model network formation and stability makes a big difference to the conclusions and not every model is appropriate in every circumstance. However, we believe that incorporating incomplete information in the dynamic network formation game represents a first and important step towards understanding why several previously seemingly irregular network topologies can emerge and remain stable in practice.

The rest of the paper is organized as follows. Section 3 introduces the model. Section 4 analyzes the model in detail and interprets the results. Section 5 discusses an alternative approach in modeling. Section 6 concludes and introduces relevant future research topics.

### 3 Model

#### 3.1 Networks with Incomplete Information

##### 3.1.1 Networks and the Agents' Types

Let  $I = \{1, 2, \dots, N\}$  denote a group of  $N$  agents. Agents are characterized by their private type  $k \in X$ , where  $X$  is a finite set of types<sup>3</sup>. The probability distribution of types on  $X$  is  $H$ . In the actual network there are  $N$  agents  $- 1, 2, \dots, N$ , whose types are drawn independently from  $X$  according to distribution  $H$ . Each agent  $i$  knows its private type  $k_i \in X$ . Let  $\kappa = \{k_i\}_{i=1}^N$  denote the type vector of the agents.

A network is denoted by  $\mathbf{g} \subset \{ij : i, j \in I, i \neq j\}$ , and the sub-network of  $\mathbf{g}$  on  $I' \subset I$ , denoted  $\mathbf{g}_{sub}(I')$ , is defined as a subset of  $\mathbf{g}$  such that  $ij \in \mathbf{g}_{sub}(I')$  if and only if  $i, j \in I'$  and  $ij \in \mathbf{g}$ .  $ij$  is called a link between agents  $i$  and  $j$ . We assume throughout that links are *undirected*, in the sense that we do not specify whether link  $ij$  points from  $i$  to  $j$  or vice versa. A network  $\mathbf{g}$  is *empty* if  $\mathbf{g} = \emptyset$ .

We say that agents  $i$  and  $j$  are *connected*, denoted  $i \stackrel{\mathbf{g}}{\leftrightarrow} j$ , if there exist  $j_1, j_2, \dots, j_n$  for some  $n$  such that  $ij_1, j_1j_2, \dots, j_nj \in \mathbf{g}$ . Let  $d_{ij}$  denote the *distance*, or the smallest number of links between  $i$  and  $j$ . If  $i$  and  $j$  are not connected, define  $d_{ij} := \infty$ . An agent  $i$  in a network is a *singleton* if  $ij \notin \mathbf{g}$  for any  $j \neq i$ .

Let  $N(\mathbf{g}) = \{i | \exists j \text{ s.t. } ij \in \mathbf{g}\}$ . A *component* of network  $\mathbf{g}$  is a maximally connected sub-network, i.e. a set  $C \subset \mathbf{g}$  such that for all  $i \in N(C)$  and  $j \in N(C)$ ,  $i \neq j$ , we have  $i \stackrel{C}{\leftrightarrow} j$ , and for any  $i \in N(C)$  and  $j \in N(\mathbf{g})$ ,  $ij \in \mathbf{g}$  implies that  $ij \in C$ . Let  $C_i$  denote the component that contains link  $ij$  for some  $j \neq i$ . Unless otherwise specified, in the remaining parts of the paper we use the word "component" to refer to any *non-empty* component.

A network  $\mathbf{g}$  is said to be *empty* if  $\mathbf{g} = \emptyset$ , and *connected* if  $\mathbf{g}$  has only one component which is itself.  $\mathbf{g}$  is *minimal* if for any component  $C \subset \mathbf{g}$  and any link  $ij \in C$ , the absence of  $ij$  would disconnect at least one pair of formerly connected agents.  $\mathbf{g}$  is *minimally connected* if it is minimal and connected.

##### 3.1.2 Payoff Structure

We assume non-local externalities in payoffs: once agents  $i$  and  $j$  form a link,  $i$  not only obtains payoffs from his immediate neighbor  $j$ , but also from the agents that he is indirectly connected to via that particular link. The payoff to forming a direct link is type dependent and given by the function  $f : X \rightarrow \mathbb{R}^{++}$ ;  $f(k)$  is the direct payoff to any agent who forms a link to an agent of type  $k$ . Agents also obtain utilities from indirect links (discounted by distance) and pay costs for maintaining links. Specifically, an agent  $i$ 's payoff in the network  $\mathbf{g}$  in a given period is given by

$$u_i(k_{-i}, \mathbf{g}) = u_i(k_{-i}, C_i) := \sum_{\substack{C_j \\ j \stackrel{C_j}{\leftrightarrow} i}} \delta^{d_{ij}-1} f(k_j) - \sum_{j: ij \in C_i} c$$

$f(k_j) > 0$  denotes the payoff to an agent  $i$  from the link to an agent  $j$ , whose value depends on  $j$ 's type  $k_j$ .  $\delta \in (0, 1)$  denotes a common decay factor, such that the payoff of  $i$  from  $j$  with a distance of  $d_{ij}$  is  $\delta^{d_{ij}-1} f(k_j)$ .  $c > 0$  is the cost of maintaining a link, which is assumed to be bilateral and homogeneous across agents. The assumption of  $c$  being homogeneous across agents is without loss of generality and is made merely to avoid redundant analysis, as the incomplete information and heterogeneity in agents' payoffs have been reflected by the potentially different types of agents. All of our major results can be obtained with slight technical changes in an environment where the cost is heterogeneous (even when the cost is also private information) across agents.

<sup>3</sup> A general measure space of types could be accommodated easily but with added technical complications.

Let  $\mathbb{E}[f(x)] = \int_X f(x)dH(x)$  denote the expected benefit from a link to a single agent, under the prior type distribution. As mentioned before, the only assumption we require on  $X$  (and functions  $H$  and  $f$ ) is that this expected payoff is well-defined.

### 3.2 Dynamic Network Formation Game

The dynamic game is the same as in Watts (2001), except that information is incomplete. Time is discrete and the horizon is infinite:  $t = 0, 1, 2, \dots$ . The game is played as follows: agents start with an empty network  $\emptyset$  in period 0. In each following period, a pair of agents  $(i, j)$  is randomly selected to update the link between them. As long as each pair of agents is selected with positive probability, the specific probability distribution for the selection process does not affect our main results.

The two agents then play a simultaneous move game, where either agent can choose to sever the link between them if there is one, and if there is not, whether to agree to form a link with the other agent. Let  $a_{ij} = 1$  denote the action that  $i$  agrees to form a link with  $j$  (if there is no existing link) or not to sever the link (if there is an existing one), and  $a_{ij} = 0$  otherwise. A link is formed or maintained after bilateral consent (i.e.  $a_{ij} = a_{ji} = 1$ ). The agents are assumed to be myopic and not forward looking, i.e. they only care about their *current* payoffs when choosing an action.

With incomplete information, agents maximize their *expected* payoffs, rather than actual payoffs, when deciding the optimal action. Therefore, the *belief* of an agent on the types of the other agents plays a crucial role in shaping his behavioral patterns. We assume the following simple updating rule for the agents:

- 1. If two agents were ever connected, they know each other's type.
- 2. Otherwise, their belief on each other's type remains at the prior.

The first part of this updating rule is a simple representation of the realistic situation that people will (sooner or later) find out the true value they receive via connection, and this specification does not affect any technical part of analysis or the implication that follows. The second part of the updating rule can be regarded as a straight forward way of modelling agents' constrained interpretation of the past formation history. In a lot of realistic situations, agents can only observe a very small part or even none of the past formation history due to many different constraints; this is especially true in social networks and business networks where a large number of agents' actions are kept private. In addition, agents may not be able to perform complicated and precise update based on their observations. As a result, when agent  $i$  meets  $j$ ,  $i$  may have some idea about how well  $j$  is connected to others, but due to the above constraints, high connectivity alone does not imply high quality/value, and vice versa. In fact, our analysis shows exactly this point, i.e. a low quality agent can assume a central position in the network just by chance. As a result, when agents do not know a lot about past history and neither can infer a lot about someone else's value from her connectivity, the prior seems a justifiable belief to hold. A similar assumption appears in McBride (2006), which is referred to as imperfect monitoring and describes agents' inability to update according to all other agents' strategies in a static network formation game. In the next section, we will highlight the role of this updating rule in the network formation process.

A plausible alternative updating rule is to enable agents to perform complete Bayesian update according to the entire formation history, which results in very complicated belief formation. We will discuss this alternative updating rule in Section 5. Of course, many other assumptions can be made about how agents update their beliefs. The two types of updating rules we consider in this paper are two extreme assumptions about observations and beliefs which readily illustrate the general difference that incomplete information makes in the network formation process.

## 4 Analysis

In this section, we analyze the nature of the network formation process and show a clear contrast between the existing results under complete information and our results under incomplete information. We begin by defining the solution concept for the two-player game in each period, and the notion of a stable network.

### 4.1 Stable Equilibrium and Stable Network

#### 4.1.1 Strategy

In this section, we formally define a strategy in the network formation game. We first provide a standard and conventional definition, and then introduce a way of simplification for the subsequent analysis.

First, we define a strategy in a standard way. Let  $(i_\tau, j_\tau)$  denote the pair of agents selected in period  $\tau$ , and let  $\mathbf{g}_\tau$  denote the network formed in period  $\tau$ . In the network formation process, the following two conditions must be satisfied:

$$\begin{aligned} \mathbf{g}_0 &= \emptyset \\ \mathbf{g}_{\tau+1} &\in \{\mathbf{g}_\tau + i_{\tau+1}j_{\tau+1}, \mathbf{g}_\tau - i_{\tau+1}j_{\tau+1}\}. \end{aligned}$$

The first condition refers to the initially empty network; the second reflects the fact that  $i_{\tau+1}j_{\tau+1}$  is the only link that can be potentially changed in period  $\tau + 1$ . Denote  $\sigma_t = \{(i_\tau, j_\tau), \mathbf{g}_\tau\}_{\tau=1}^t$  as the formation history up to period  $t$ , and let  $\Sigma_t$  denote the set of all possible formation histories up to period  $t$ , with the initial condition  $\Sigma_0 = \emptyset$ . Let  $\Sigma = \cup_{t=0}^\infty \Sigma_t$ , a (pure) strategy of agent  $i$  is then a mapping  $s_i : X \times \Sigma \times (I - i) \rightarrow \{0, 1\}$ , where  $X$  is the set of possible types and  $I - i$  is the set of agents other than  $i$ . For agent  $i$ , given the strategy profile of other agents  $s_{-i}$ , a *best response* is then a strategy  $s_i$  such that given any  $k_i \in X$ ,  $\sigma \in \Sigma$  and  $j \in I - i$ ,  $s_i$  selects an action that maximizes  $i$ 's (current) expected payoff according to her belief.

The complete set of strategies for an agent is enormous: a different action can be chosen based on each different history, and in infinite periods of time there will be infinitely many possible histories. However, we argue below that it is sufficient to consider a particular subset of strategies, and these strategies admit a particularly simple description.

First, note that the current-period payoff of an agent does not depend on her own type. In any period, given any strategy profile of other agents, if some action is strictly optimal for agent  $i$  of type  $k_i$ , then it must also be strictly optimal for agent  $i$  of any other type. (Except in non-generic cases where an agent is indifferent between the two available actions, any best response – any candidate for equilibrium – must be independent of an agent's own type.)

Agents are assumed to be myopic; they choose an action to maximize current expected payoff according to belief. Now for agent  $i$ , a weakly best response for  $i$  is always choose 1, if her expected payoff according to her belief is non-negative and 0, otherwise. We assume agents link when indifferent. (Again, except for non-generic cases where  $i$  is indifferent between her two available actions, this is in fact the strict best response.) Here, the sufficient information for  $i$  to determine her optimal action is her belief on the type vector, and the component structure of the other agent selected.

Hence, we do not need to consider all strategies, only those that are ever best responses (optimal) and these admit a simple alternative description as mappings from beliefs and component structures to actions. Such strategies are straightforward to characterize. Formally, let  $B_i : \Sigma \rightarrow \Delta(X^N)$  be the belief updating function of agent  $i$ , with the constraint that  $B_i$  always assigns probability 1 on  $i$ 's true type and  $B_i(\emptyset)$  is equal to the prior belief, and let  $C_j : \Sigma \rightarrow G$  ( $G$  being the set of all possible networks) be the mapping from the formation history to the resulting component containing  $j$ . Then let  $Y_j = \{(B_i(\sigma), C_j(\sigma)) : \sigma \in \Sigma\}$ . We can describe a strategy of agent  $i$  as a mapping  $s_i : X \times Y_i \times (I - i) \rightarrow \{0, 1\}$ <sup>4</sup>. With this characterization, we can break down the solution of equilibria with sequential rationality to the much more straight forward inspection of equilibria in the 2-person link formation game in each period.

#### 4.1.2 Stable Equilibrium

Now we define our solution concept, which we refer to as the *stable equilibrium (SE)*. With a slight abuse of notation, in the subsequent analysis we use  $B_i$  to denote the *realized* belief vector of agent  $i$  based on her belief updating function, and  $C_i$  to denote the component containing  $i$ .

**Definition 1.** A strategy profile  $s$  is a *stable equilibrium (SE)* if for any agent  $j \neq i$ ,  $s_i(k_i, (B_i, C_j), j) = 1$  if and only if

$$\mathbb{E}[u_i(k_{-i}, (C_i \cup C_j) + ij) | B_i] \geq u_i(k_{-i}, C_i).$$

We would like to emphasize here that even though the definition of SE does not involve an explicit expression of best response, it essentially represents bilateral best response in the one-period linking game. Note that the above strategy profile does not depend on the agents' types, i.e. for agent  $j$ , given  $B_j$ , she takes the same action no matter what her type is. Hence for agent  $i$ , her payoff from a link with agent  $j$  only depends on her belief vector  $B_i$  and the component  $C_j$ . In other words,  $i$ 's expected payoff is  $\mathbb{E}[u_i(k_{-i}, (C_i \cup C_j) + ij) | B_i]$  if a link is formed, and  $u_i(k_{-i}, C_i)$  otherwise. A weakly best response for  $i$  is then to choose action 1 if  $\mathbb{E}[u_i(k_{-i}, (C_i \cup C_j) + ij) | B_i] \geq u_i(k_{-i}, C_i)$  and 0 otherwise, which is exactly as prescribed by the above strategy. Hence the above strategy maximizes  $i$ 's current expected payoff given  $s_j$ , and vice versa.

<sup>4</sup> As mentioned before, in generic cases agent  $i$ 's best response will not depend on her own type. Nevertheless, we still keep the type of  $i$  as an argument in  $i$ 's strategy since it is part of  $i$ 's knowledge after all.

In words, a SE is a Nash equilibrium where any agent would choose to agree to form a link (if there is no existing link) or choose not to sever the link (if there is an existing link) as long as the expected payoff from the link is non-negative according to her current belief. It is *stable* because the prescribed strategy is robust to small probabilistic changes in the counter party's strategy when the above inequality is strict. More specifically, if some agent  $j$  other than  $i$  changes her action with a sufficiently small probability  $\varepsilon$ ,  $i$ 's best response would not change. In particular, it excludes the "pessimistic" or null equilibria in which no link formation occurs even though each agent has a non-negative expected payoff from the potential link. In other words, agents choose to link as long as they are at least indifferent. This equilibrium notion is similar to pairwise stability in Jackson and Wolinsky (1996), but for the setting of incomplete information; thus, the conditions that characterize "stability" are based on expected rather than realized payoffs.

The following lemma shows the existence and uniqueness of such an equilibrium.

**Lemma 1.** *SE exists and is unique.*

*Proof.* Consider any period. Assume that agents  $i$  and  $j$  are selected in that period. Given the belief vectors  $B_i$  and  $B_j$ , and the components  $C_i$  and  $C_j$ ,  $i$ 's (similarly,  $j$ 's) strategy satisfying the condition in the definition of a SE can be expressed as

$$s_i(k_i, (B_i, C_j), j) = \begin{cases} 1, & \text{if } \mathbb{E}[u_i(k_{-i}, (C_i \cup C_j) + ij) | B_i] \geq u_i(k_{-i}, C_i) \\ 0, & \text{otherwise} \end{cases}$$

Since the terms on both sides of the above inequality are well-defined and agent  $i$ 's action is binary, agent  $i$ 's optimal choice in this period is unique. Since the period and agents assumed are arbitrary, SE is unique. (Given that agents link when indifferent, the argument is that agents follow a dominant strategy.)  $\square$

Lemma 1 together with the above robustness property of a SE ensures that the outcome of the formation process is unique and robust to small perturbation (or tremble) in agents' action. It is also useful to note that Lemma 1 holds for any belief updating function.

#### 4.1.3 Stable Network

From now on, we assume that agents play the SE in every period. Let  $\gamma := \{(i_\tau, j_\tau)\}_{\tau=1}^t$  ( $t \geq 1$ ) denote a *selection path* up to time  $t$ , or the set of selected pairs of agents, ordered from 1 to  $t$ . Since the equilibrium in each period exists and is unique, the realization of the entire network formation process can be fully characterized by  $\kappa$ , the type vector of agents and  $\gamma_\infty$ , the realization of the random selection process. We use  $\Gamma$  to denote a formation process, and in particular denote  $\Gamma_C$  and  $\Gamma_{IC}$  as network formation processes under complete and incomplete information correspondingly.

Let  $\mathbf{g}_C(\gamma)$  denote the unique network formed after period  $t$ , following a selection path of  $\gamma$ , under complete information, and  $\mathbf{g}_{IC}(\gamma)$  the network under incomplete information. Let  $B_C(\gamma) = \{B_{i,C}(\gamma)\}_{i=1}^N$  denote the associated belief vector after period  $t$  under complete information (where  $B_{i,C}(\gamma)$  is always equal to the degenerate belief on the true type vector  $k$ , for every  $i$ ), and  $B_{IC}(\gamma)$  the belief vector under incomplete information. By Lemma 1 we know that both  $\mathbf{g}_C(\gamma)$  ( $\mathbf{g}_{IC}(\gamma)$ ) and  $B_C(\gamma)$  ( $B_{IC}(\gamma)$ ) are well-defined. We say that:

- 1. A network  $\mathbf{g}$  can emerge under complete information (incomplete information) if there exists a selection path  $\gamma$  for some  $t$ , such that  $\mathbf{g} = \mathbf{g}_C(\gamma)$  ( $= \mathbf{g}_{IC}(\gamma)$ ).
- 2. A network with the associated belief vector,  $(\mathbf{g}, B)$  ( $B = \{B_i\}_{i=1}^N$ ), is a *stable network* under complete information (incomplete information) if no link is formed or broken given any subsequent selection path in  $\Gamma_C$  ( $\Gamma_{IC}$ ). If  $B$  refers to the belief vector under complete information (where  $B_i$  is the degenerate belief on the true type vector for all  $i$ ), then we just say that  $\mathbf{g}$  is a stable network.
- 3.  $\Gamma_C$  ( $\Gamma_{IC}$ ) can converge to  $\mathbf{g}$  if there exists a selection path  $\gamma$  for some  $t$ , such that  $\mathbf{g} = \mathbf{g}_C(\gamma)$  ( $= \mathbf{g}_{IC}(\gamma)$ ) and  $(\mathbf{g}_C(\gamma), B_C(\gamma))$  ( $(\mathbf{g}_{IC}(\gamma), B_{IC}(\gamma))$ ) is stable.

#### 4.2 Information Revelation

Before discussing the differences between the network topologies that can emerge and be stable under complete information and incomplete information, we first inspect how long incomplete information can persist. We say that *information is complete* (following a given history) when every agent's belief (following that history) is the degenerate belief on the true type vector  $\kappa$ , i.e.  $\text{Prob}(\kappa | B_i) = 1$  for any  $i$ , and that *information is incomplete* otherwise.

Recall that agents update their beliefs based on the simple updating rule: if two agents have ever been connected, they know each other's type; otherwise, their beliefs on each other's type remain at the prior. In other words, for agent  $i$ , the following is true for any realized belief  $B_i$ : the argument in  $B_i$  for the type of any other agent  $j$  is the degenerate belief on  $j$ 's true type  $k_j$  if  $i$  and  $j$  have ever been connected, and is equal to the prior belief otherwise. Hence, from the definition of SE, we know that if  $\mathbb{E}[f(x)] < c$ , no link will ever form from the very beginning; if  $\mathbb{E}[f(x)] \geq c$ , then every agent is willing to form a link with any other agent that she has not been connected to before. This observation enables us to determine the property of *information revelation* in the network formation process, which is noted in the following proposition.

**Proposition 1.** *For any  $\kappa$ :*

- 1. *If  $\mathbb{E}[f(x)] < c$ , information never becomes complete: agents' beliefs remain forever at the prior.*
- 2. *If  $\mathbb{E}[f(x)] \geq c$ , information becomes complete within finitely many periods almost surely, and information is complete in any stable network.*

*Proof.* If  $\mathbb{E}[f(x)] < c$ : by inspecting the SE we know that for any pair of agents selected, no link would be formed in any period. Therefore, no agent ever learns the type of any other agent, and the beliefs would stay at the prior.

If  $\mathbb{E}[f(x)] \geq c$ : we first show that information becomes complete within finitely many periods almost surely. It suffices to show that any two agents are connected for at least one period within finitely many periods almost surely. Since  $\mathbb{E}[f(x)] \geq c$ , by the definition of SE it further suffices to show that any two agents are selected at least once within finitely many periods almost surely. Consider any two agents  $i$  and  $j$ ; the probability of the event that they are not selected in one period is  $1 - \frac{2}{N(N-1)} < 1$ , and thus the probability of this event occurring for infinitely many periods is 0.

Next, we show that information must be complete in any stable network. If  $\mathbf{g}$  is connected, then clearly there is complete information. If  $\mathbf{g}$  is unconnected and information is not complete, then there must exist two unconnected agents such that their beliefs on each other's type remain at the prior. When they are selected they would form a link, which implies that  $(\mathbf{g}, B)$  is not stable, a contradiction.  $\square$

When the expected benefit from a link under the prior weakly exceeds the link formation cost, everyone has the incentive to form links with others whose types are unknown, and thus eventually learns the true type vector with probability 1 over time. Indeed, as agents are always willing to form links with strangers, after sufficiently many periods the probability of pairs of agents who never connected (i.e. pairs of agents that have never met and hence have never learnt about each other) would be arbitrarily small. In other words, only in an early stage of the formation process can incomplete information make a difference and affect the ultimate network topology, as compared with complete information.

#### 4.3 Contrast between Complete and Incomplete Information

Even though Proposition 1 may leave the impression that incomplete information is not crucial for the formation process as it only takes effect in the short run, we will emphasize in the following analysis that such short-term influence is actually persistent over time.

Let  $G_C(\kappa)$  denote the set of networks that can emerge under complete information given  $\kappa$ , and  $G_{IC}(\kappa)$  that under incomplete information. Similarly, let  $G_C^*(\kappa)$  denote the set of networks that can emerge and be stable under complete information given  $\kappa$ , and  $G_{IC}^*(\kappa)$  that under incomplete information.

**Theorem 1.** *For any  $\kappa$ :*

- 1. *If  $\mathbb{E}[f(x)] < c$ , then  $G_{IC}(\kappa) = G_{IC}^*(\kappa) = \{\emptyset\}$ .*
- 2. *If  $\mathbb{E}[f(x)] \geq c$ , then  $G_{IC}(\kappa) \supset G_C(\kappa)$ , and  $G_{IC}^*(\kappa) \supset G_C^*(\kappa)$ .*

*Proof.* If  $\mathbb{E}[f(x)] < c$ : as in the proof of Proposition 1, no link would ever form and thus the only network that can emerge is the empty network. This proves 1.

To prove 2, assume  $\mathbb{E}[f(x)] \geq c$  and consider  $\mathbf{g} \in G_C(\kappa)$ .

If  $\mathbf{g}$  is empty: since  $\mathbf{g} \in G_C(\kappa)$ , there must exist two agents  $i, j$  such that  $f(k_i) < c$  or  $f(k_j) < c$ . Consider the selection path  $\gamma_2 = ((i, j), (i, j))$  under incomplete information. It is clear that a link would form between  $i$  and  $j$  in period 1, but the link would then be severed in period 2, and thus  $\mathbf{g} = \mathbf{g}(\gamma_2)$ , which implies that  $\mathbf{g} \in G_{IC}(\kappa)$ .

If  $\mathbf{g}$  is non-empty: consider any selection path  $\gamma_t^C$  such that  $\mathbf{g}$  emerges for the first time in period  $t$ . By the definition of  $G_C(\kappa)$ , we know that such  $\gamma_t^C$  exists. We construct a different selection path under which  $\mathbf{g}$  forms when information is incomplete (but the true type vector is  $\kappa$ ). Let  $\gamma_t^{IC}$  be a selection path constructed from  $\gamma_t^C$  such that the pairs of agents in  $\gamma_t^{IC}$  between whom there is no existing link, but a new link is not formed either,

are deleted. Note that the two selection paths may take different number of time periods:  $\tau \leq t$  by the above construction.

Consider the formation process under incomplete information given  $\gamma_\tau^C$ . First, it is clear that for any agent  $i$  with  $f(k_i) < c$ , no link will be formed between  $i$  and any other agent under complete information. Hence, if a link is formed between  $i$  and  $j$  under complete information, it must be the case that  $f(k_i)$  and  $f(k_j)$  are both weakly higher than  $c$ ; then according to the simple updating rule, the same link will also be formed under incomplete information regardless of whether  $i$  and  $j$  know each other's type. Also, it is clear that the decision of severing a link by any agent is the same under complete information and incomplete information, since such a decision is based on the realized payoff. Therefore, the formation process yields the same link formation and severance results under complete information given  $\gamma_\tau^C$  and under incomplete information given  $\gamma_\tau^C$ . Hence given  $\gamma_\tau^C$ ,  $\mathbf{g}$  emerges for the first time in period  $\tau$  under incomplete information, which implies that  $\mathbf{g} \in G_{IC}(\kappa)$ . Therefore  $G_C(\kappa) \subset G_{IC}(\kappa)$ . This proves the first part of 2.

To prove the second part of 2 (stability), we first observe that, by the above argument, we already know that any network that can emerge under complete information can also emerge under incomplete information. Thus it suffices to show that, for any network  $\mathbf{g} \in G_C^*(\kappa)$ , there exists a subsequent selection path under incomplete information after  $\mathbf{g}$ 's first appearance that would make  $\mathbf{g}$  stable. We prove the result by construction.

If  $\mathbf{g}$  is empty: consider the selection path  $\gamma_{\frac{N(N-1)}{2}}$  such that every pair of agents is selected exactly twice consecutively. Since by assumption  $\mathbf{g} \in G_C^*(\kappa)$ , we know that for every pair of agents a link would first form and then be severed in the next period. In period  $\frac{N(N-1)}{2} + 1$ , information is complete and the empty network becomes stable.

If  $\mathbf{g}$  is non-empty: denoting the number of components and singleton agents in  $\mathbf{g}$  as  $q(\mathbf{g})$ , under incomplete information let the subsequent selection path after  $\mathbf{g}$ 's first appearance be such that in the first  $q(\mathbf{g})(q(\mathbf{g}) - 1)$  periods, two agents from different components are selected exactly twice consecutively and every two components (or singleton agents) are involved. By assumption  $\mathbf{g} \in G_C^*(\kappa)$ , which means that, should two agents from different components know each other's type, no link would be formed between them. Therefore, under incomplete information, when a pair of agents from different components is selected for the second time, either there is no existing link between them and no link would be formed, or an existing link would be severed. In either case, the agents know each other's type as well as the types of agents in the counter party's component. Therefore, after  $q(\mathbf{g})(q(\mathbf{g}) - 1)$  periods, every agent knows  $\kappa$  and essentially there is no incomplete information. Again by the assumption that  $\mathbf{g} \in G_C^*(\kappa)$ , we can conclude that  $\mathbf{g}$  is such that no link would be formed or severed in any later period given any selection path. Thus  $\mathbf{g} \in G_{IC}^*(\kappa)$ , which completes the proof.  $\square$

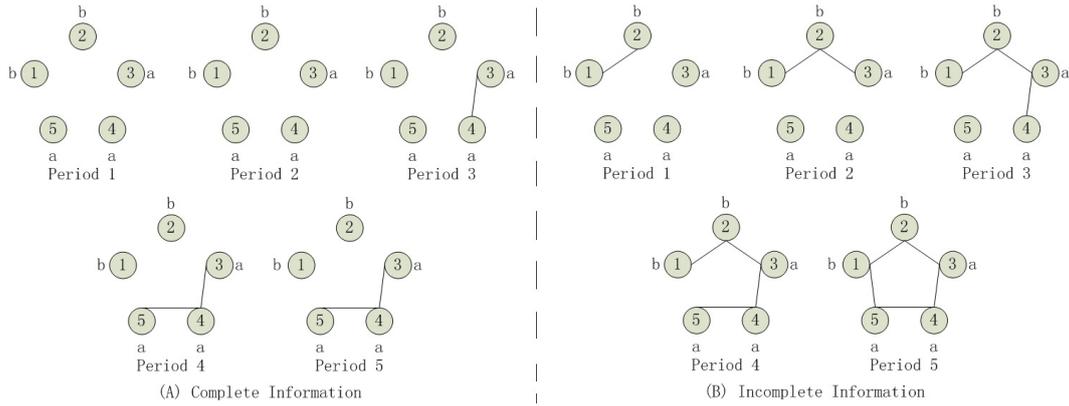
Theorem 1 states that, when expected benefits are sufficiently high, if some network can emerge (and be stable) under complete information, then it can also emerge (and be stable) under incomplete information, but the reverse may not be true. Intuitively, if a link could be formed under complete information, then given high expected benefits and the simple updating rule, it can also be formed under incomplete information, whether or not the relevant agents know each other's type. Note that the reverse is not necessarily true: even if a link could be formed under incomplete information, it may not form under complete information because the expected payoffs are sufficiently higher than the realized payoffs, i.e. in the incomplete information setting, were the agents to know each other's type beforehand, the link may never be formed. In other words, under incomplete information, high expected payoffs can initialize link formation such that even though agents would "regret" the links they formed after knowing each other's type. However, if more links have formed before they are selected again to update their initial links, then due to increasing returns to link formation the positive externalities would in turn ensure that the initial links are maintained. The following example illustrates this point.

**Example 1.** Assume the following:  $N = 5$ ,  $X = \{a, b\}$ , and the other parameters are such that  $f(b) < c < f(a)$ ,  $\mathbb{E}[f(x)] \geq c$ ,  $(1 + \delta - \delta^2 - \delta^3)f(b) \geq c$ , and  $(1 - \delta^3)f(a) < c$ . In addition, we assume that the realized type distribution is consistent with the agents' prior belief, i.e. the actual number of type  $a$  agents  $= N \times \text{Prob}(k_i = a)$ . For illustrative purpose, we consider the case that  $k_1 = k_2 = b$ ,  $k_3 = k_4 = k_5 = a$ , and  $\text{Prob}(k_i = a) = 0.6$ .

Consider the following selection path:  $((1, 2), (2, 3), (3, 4), (4, 5), (1, 5))$ . Under complete information, the network is never connected, as in Figure 1(A). Under incomplete information, the SE can be explicitly computed in each period. For example, in period 1, agent 1's expected payoff from the link with agent 2 is  $\mathbb{E}[f(x)] \geq c$  and vice versa, and thus the link is formed. The formation process is shown in Figure 1(B).

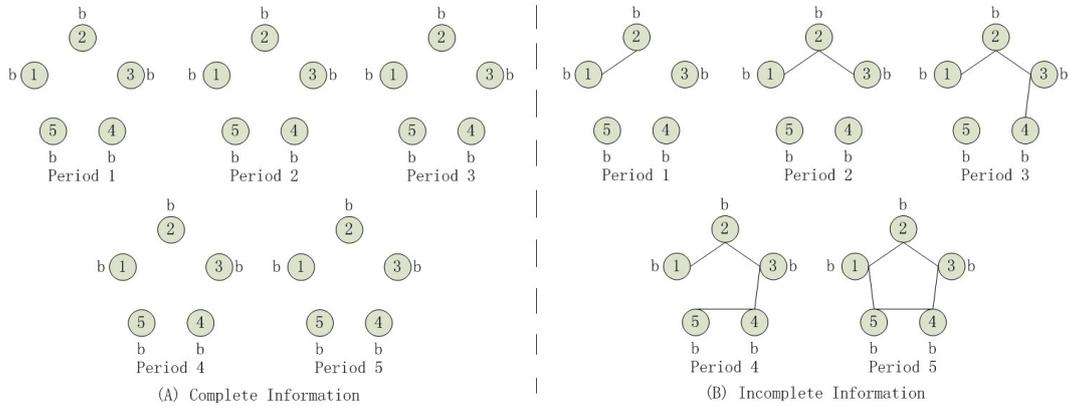
According to the assumptions on parameters, one can then easily show that the network formed in period 5 is stable. Note that agent 1 prefers to maintain the link to agent 2 even though agent 2 is of a low type because that shortens agent 1's path to agent 3.

Of course, though the agents' beliefs are always consistent with each other and with the prior distribution of types, they may not be always consistent with the realized type vector; when they are not, the difference between complete and incomplete information is even larger. For instance, assume that  $k_i = b$  for all  $i$  and consider the same



**Fig. 1** Not Connected under Complete Info. vs. Connected under Incomplete Info.

selection path as above. Under complete information, the network stays empty; under incomplete information, the formation process still converges to a wheel network.



**Fig. 2** Empty under Complete Info. vs. Connected under Incomplete Info.

An important factor that enables networks that are never formed under complete information to form and be stable under incomplete information is that agents do not sever undesirable links immediately. In the model, this is an event that happens with significantly high probability, since the probability that the same pair of agents is selected twice consecutively is rather low. It becomes even lower in a larger society. In numerous works on network formation, such as Jackson and Wolinsky (1996), Watts (2001) and Dutta et al. (2005), the random selection of a potential link to update is a common modelling assumption. In practical situations, this event can be understood as follows: it takes time for people to make up their minds to disconnect with someone they do not like. Before deciding to sever a link with a person, that person may have built new connection with others, which will change the value of linking with her after all. The reason for such delay in decision may be some legal, geographic or technological barrier. For instance, if a link represents a binary contract, one cannot easily terminate the contractual relationship before the expiration date. Alternatively, it might be the case that knowing the true value of linking with others is a time-consuming process. It usually takes a certain amount of communication and interaction before people/entities really know the value of their connection, especially in social networks.

Another feature of incomplete information is the *history dependence* of the formation process, in the sense that the ultimate network topology depends greatly on the selection path. As a result, even if a type is more valuable or preferable than another, under incomplete information it is not necessary that an agent of that type ends up with a higher connectivity degree. Consider the following example: assume the same parameter values as in Example 1, and consider a group of agents consisting of 4 type *a* agents and 5 type *b* agents. There exists a selection path such that: under complete information, the formation process converges to a star network with only type *a* agents (Figure 3(A)); under incomplete information, the formation process converges to a “hub-and-spokes” network (Figure 3(B))<sup>5</sup>.

<sup>5</sup> One such selection path is ((1,2), (1,3), (2,6), (3,7), (6,7), (1,5), (5,9), (6,9), (1,4), (4,8), (7,8), (8,9), (2,5), (3,5), (4,5)).

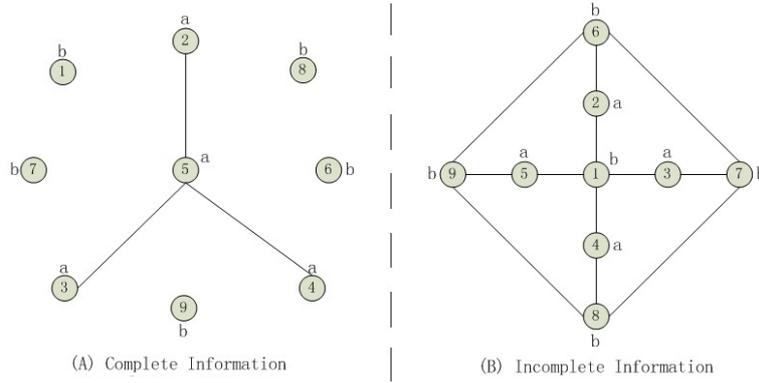


Fig. 3 Different Connectivity Degree Distributions

Under complete information, the center of the star network has to be a type *a* agent since no type *b* agent ever gets linked with anyone else. In fact, regardless of the selection path, no type *b* agent can ever get linked. By contrast, under incomplete information, it first becomes possible for two type *b* agents to form a link; and then, as it turns out in this particular topology, each type *b* agent’s distance with the type *a* agent is sufficiently small. Even though the type *a* agent has a low connectivity degree, the other agents do not find a new link with the type *a* agent attractive, because it does not offer sufficient indirect benefits. Hence, the agent with the more valuable type – type *a* – ends up with the lowest connectivity degree in the network. This is in stark contrast with the existing results in the literature (for instance the property of “law of the few” in Galeotti and Goyal (2010)), which often show that a more valuable agent is better connected. As mentioned in the literature review, violation of such theoretical predictions has been documented in a number of empirical and experimental studies. From this perspective, our result can be regarded as a micro-foundation for the prevalent phenomenon that an agent may obtain a central position in a network by chance instead of merit.

Furthermore, the event that some agent of “low value” ends up with relatively high connectivity under incomplete information is not a rare event. In Figure 4, we show the simulation result of a network formation process under incomplete information with 15 “high-type” (in the sense that  $(1 - \delta)f(x) < c < f(x)$ ) agents and 1 “low-type” (in the sense that  $f(x) < c$ ) agent<sup>6</sup>. The figure shows the “low-type” agent’s rank in terms of connectivity. Under complete information, the “low-type” agent would have never been connected; under incomplete information, even though the probability that the “low-type” agent obtains the highest connectivity is rather low, the probability that she ranks among the top half (at or above 8th) is more than 0.2.

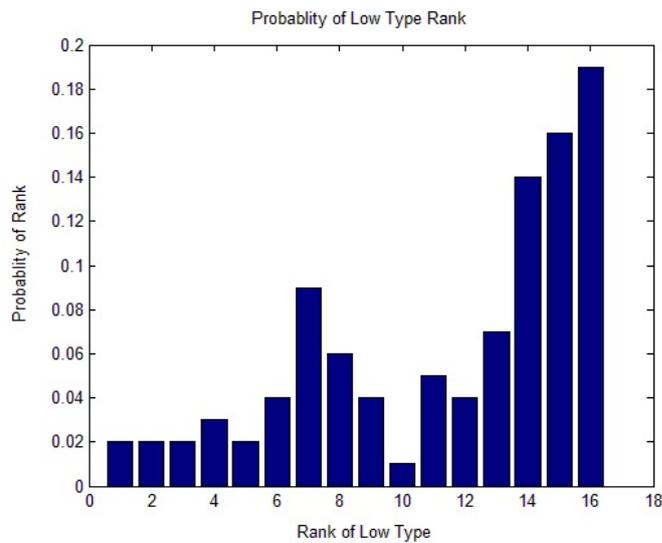
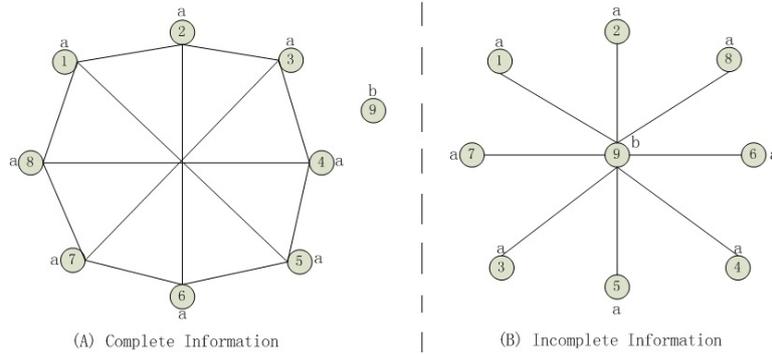


Fig. 4 Simulations: Rank of “Low-Type” Agent

<sup>6</sup> We set the “high-type” agent’s value to be  $f(x) = 6$  and the “low-type” agent’s value to be  $f(x) = 4$ . In addition, we assume that  $c = 5$  and  $\delta = 0.6$ . In each simulation, we let the formation process run for 2500 periods. We use 100 simulations to find the average result.

The above examples also highlight and clarifies the point made by Theorem 1: incomplete information generates a superset of *networks*, not a superset of *links*, as compared with complete information. In other words, *new and different* networks can be formed under incomplete information, rather than a mere addition of links to networks formed under complete information. Indeed, the network in Figure 3(A) has 3 links and that in Figure 3(B) has 12, but they share *no links in common*.

Moreover, even when  $\mathbb{E}[f(x)] \geq c$ , incomplete information does *not* imply that more links are always formed in the stable network. For instance, let  $X = \{a, b\}$ , and consider a group of 8 type *a* agents (indexed 1, 2, ..., 8) and 1 type *b* agent (indexed 9). The payoffs are  $f(b) < c$ ,  $f(a) \geq c$ ,  $(1 - \delta)f(a) < c \leq (1 - \delta^2)f(a)$  and  $\mathbb{E}[f(x)] \geq c$ . Let the selection path be as follows: first, select 9 once with each of 1, 2, ..., 8. Then select  $((1, 2), (2, 3), \dots, (7, 8), (8, 1))$ . Finally, select  $((1, 5), (2, 6), (3, 7), (4, 8))$ . The resulting stable network is shown in Figure 5 below: under complete information, the network has 12 links, while under incomplete information it has only 8.



**Fig. 5** More Links under Complete Information

The above examples also show that, unlike the literature on network formation with complete information, which proves that in each model only one or two types of network topologies can be stable (see Bala and Goyal (2000), Galeotti et al (2006) and Galeotti and Goyal (2010)), under incomplete information many *more types* of networks can emerge and be stable. Even when compared with models that allow for more possibilities in network types (see Jackson and Wolinsky (1996) and Watts (2001)), incomplete information again brings about a wider range of stable network topologies. The next theorem formalizes this statement; but first we need to recall some familiar definitions of network structures.

- 1.  $\mathbf{g}$  is *complete* if  $ij \in \mathbf{g} \forall i, j$  such that  $i \neq j$ .
- 2.  $\mathbf{g}$  is a *star* network if there exists  $i \in I$  such that  $ij \in \mathbf{g} \forall j \neq i, j \in I$  and  $i'j \notin \mathbf{g} \forall i', j \neq i$ .
- 3.  $\mathbf{g}$  is a *core-periphery* network if there exists non-empty  $I' \subsetneq I$ , such that  $ij \in \mathbf{g} \forall i, j \in I', i \neq j$ , and that  $\forall j' \in I \setminus I', ij' \in \mathbf{g}$  for some  $i \in I'$  and  $jj' \notin \mathbf{g} \forall j \neq i$ . Note that a star network is a special case of a core-periphery network.
- 4.  $\mathbf{g}$  is a *tree* network if there exists a partition of  $I, I_1, \dots, I_n$ , such that (1)  $\#(I_1) = 1$ ; (2)  $\forall n' = 2, \dots, n$ , each agent in  $I_{n'}$  has one and only one link with some agent in  $I_{n'-1}$ ; (3) no other link exists.<sup>7</sup>
- 5.  $\mathbf{g}$  is a *wheel* network if there exists a bijection  $\pi : I \rightarrow I$  such that  $\mathbf{g} = \{\pi^{-1}(1)\pi^{-1}(2), \pi^{-1}(2)\pi^{-1}(3), \dots, \pi^{-1}(N-1)\pi^{-1}(N)\}$ .

**Theorem 2.** Assume that  $\mathbb{E}[f(x)] \geq c$ . Fix a type vector  $\kappa$  and a network  $\mathbf{g}$ . If  $\mathbf{g}$  is stable when information is complete and belongs to any one of the following categories:

- 1. *Empty network*;
- 2. *Minimally connected network* (i.e. *tree network*, including *star network*);
- 3. *Fully connected network*;
- 4. *Core-periphery network*;
- 5. *Wheel network*.

then  $\mathbf{g}$  can emerge and be stable when information is incomplete, i.e.  $\mathbf{g} \in G_{IC}^*(\kappa)$ .

Note that we have assumed only that  $\mathbf{g}$  is stable when information is complete, not that it can emerge when information is complete.

<sup>7</sup> Essentially, a tree network is equivalent to a minimally connected network, and a star network is a special case of a tree network.

*Proof.* For 1, see the proof of Theorem 1. For 2 - 4, since  $\mathbf{g}$  is connected and by assumption  $\mathbf{g}$  is stable under complete information, it suffices to show that when  $\mathbf{g}$  belongs to any of the categories there exists a selection path such that  $\mathbf{g}$  can emerge in the formation process. We discuss case by case and prove them by construction below.

2: Let  $L$  be the total number of links in  $\mathbf{g}$ . Let the selection path be such that the pair of agents for each link in  $\mathbf{g}$  is selected once and only once in the first  $L$  periods. Since  $\mathbb{E}[f(x)] \geq c$ , we know that each link will be formed, and thus  $\mathbf{g}$  emerges in period  $L$ .

3: Since  $\mathbf{g}$  is fully connected and stable, we know that for any two agents  $i$  and  $j$ ,  $(1 - \delta)f(k_i) \geq c$ . Therefore regardless of the selection path  $\mathbf{g}$  would emerge.

4: Let the selection path be such that: first each periphery agent is selected once and only once with their corresponding core agent, then every two core agents are selected once and only once before any other pair of agents is selected. Since  $\mathbb{E}[f(x)] \geq c$  and  $\mathbf{g}$  is stable under complete information, we know that each link will be formed, and thus  $\mathbf{g}$  emerges after the last pair of core agents is selected.

5: Let the selection path be such that the pair of agents for each link in  $\mathbf{g}$  is selected once and only once in the first  $N - 1$  periods. Since  $\mathbb{E}[f(x)] \geq c$  and  $\mathbf{g}$  is stable under complete information, we know that each link will be formed, and thus  $\mathbf{g}$  emerges in period  $N - 1$ .  $\square$

Theorem 2 explicitly characterizes types of connected networks that can emerge and be stable under incomplete information, and most typical networks in both the literature and empirical studies are included. However, note that there may be some stable networks that cannot emerge under complete or incomplete information – e.g. a network  $\mathbf{g}$  with a subset of links  $\mathbf{g}'$  such that (1) within  $\mathbf{g}'$ , the benefit from any one link cannot cover the maintenance cost without the existence of the other links and (2)  $\mathbf{g} \setminus \mathbf{g}'$  is still connected. Such network topologies may never be formed since only one pair of agents is selected in each period and the agents are myopic.

#### 4.4 Characterizing Topological Differences

In the previous analysis, we have seen that even with the same selection path, very different networks can emerge and be stable under incomplete information; in this section, we formalize a way of describing such topological differences, and characterize the corresponding conditions under which these differences are achieved.

To obtain a clear characterization result on the topological differences, we first categorize the agents based on their types. We say that  $i$  is a **low-value** agent if  $f(k_i) < c$ , i.e. a link with this agent is not worthwhile anyway; a **medium-value** agent if  $(1 - \delta)f(k_i) < c \leq f(k_i)$ , i.e. a link with this agent would be beneficial if there exists no indirect path; and a **high-value** agent if  $(1 - \delta)f(k_i) \geq c$ , i.e. a link would still be beneficial even if there already exists an indirect path with two links (the shortest indirect path). Let  $n_l$ ,  $n_m$  and  $n_h$  denote the number of agents in the corresponding category.

We will consider selection paths for which the formation process converges under both complete and incomplete information. The following lemma establishes that such paths exist.

**Lemma 2.** *For every  $\kappa$ , there exists a formation path such that the formation process leads to a stable network under both complete and incomplete information.*

*Proof.* Consider the following selection path:

- 1. Fix a high-value or medium-value agent  $i^*$ . In the first  $N - 1$  periods, select  $i^*$  and every other agent exactly once.
- 2. In the following  $N - 1$  periods, select  $i^*$  and every other agent exactly once again.
- 3. In the following  $\frac{n_h(n_h - 1)}{2}$  periods, select every pair of high-value agents.
- 4. In the following  $n_l(n_l - 1)$  periods, select every pair of low-value agents twice consecutively.

If  $n_m + n_h \geq 1$ : under complete information: after step 2, a star with  $i^*$  as the center and all other high-value or medium-value agents as the periphery would be formed. After step 3, there will be a link between every pair of high-value agents. The network formed after step 3 is stable. Under incomplete information: after step 1, a star with  $i^*$  as the center and all other agents as the periphery would be formed. After step 2, every link between  $i^*$  and a low-type agent will be severed, and a star with  $i^*$  as the center and all other high-value or medium-value agents as the periphery would be formed. After step 3, there will be a link between every pair of high-value agents. In step 4, no link will be formed since information has been complete after step 1. Hence, the network formed after step 3 is stable.

If  $n_m + n_h = 0$ : under complete information, it is clear that no link ever forms. Under incomplete information, during step 4 a link would be formed and then severed between every pair of low-value agents, after which information would be complete. Therefore, under both complete and incomplete information, the empty network after step 4 is stable.  $\square$

We say that two networks are *identical* if they are both empty or have the same links, and *entirely different* if at least one of them is non-empty and they share no link in common. To say that two formation processes can converge to identical networks (or entirely different networks), we mean that there exists a selection path for both formation processes such that identical (or entirely different) networks emerge and be stable.

The following proposition states the topological differences in terms of the resulting stable network topology, between a formation process under complete information  $\Gamma_C$  and one under incomplete information  $\Gamma_{IC}$ .

**Proposition 2.** *For every  $\kappa$ , the following properties hold:*

- 1. If  $\mathbb{E}[f(x)] < c$ , then  $\Gamma_C$  and  $\Gamma_{IC}$  converge to identical networks with probability 1 if  $n_m + n_h \leq 1$ , and they converge to entirely different networks with probability 1 if  $n_m + n_h > 1$ .
- 2. If  $\mathbb{E}[f(x)] \geq c$ , then  $\Gamma_C$  and  $\Gamma_{IC}$  converge to identical networks with positive probability for any values of other parameters, and:
  - a. If  $n_h \geq 2$  or  $n_l = 0$ , then  $\Gamma_C$  and  $\Gamma_{IC}$  never converge to entirely different networks.
  - b. If  $n_h < 2$  and  $n_l > 0$ , then  $\Gamma_C$  and  $\Gamma_{IC}$  converge to entirely different networks with positive probability if (1)  $n_m + n_h$  is sufficiently large, or (2)  $n_m + n_h \geq 2$ ,  $\delta$  is sufficiently close to 1 and  $n_l$  is sufficiently large.

*Proof.* 1: We already know from Theorem 1 that if  $\mathbb{E}[f(x)] < c$ , the network stays empty under incomplete information for any  $\kappa$  and  $\gamma_\infty$ . If  $n_m + n_h \leq 1$ , clearly the network stays empty under complete information for any  $\kappa$  and  $\gamma_\infty$ , otherwise at least one pair of high-value or medium-value agents will be linked with probability 1.

2: The claim that  $\Gamma_C$  and  $\Gamma_{IC}$  converge to identical networks with positive probability is a direct result from Theorem 1.

If  $n_h \geq 2$ , in any stable network under complete and incomplete information, any pair of high-value agents must be linked. Thus  $\Gamma_C$  and  $\Gamma_{IC}$  never converge to entirely different networks. If  $n_l = 0$ , then the formation processes under complete and incomplete information would be the same, so again  $\Gamma_C$  and  $\Gamma_{IC}$  never converge to entirely different networks.

If  $n_h < 2$  and  $n_l > 0$ , first consider the following selection path when  $n_m + n_h \geq \frac{c - \max_{i \text{ is low-value}} f(k_i)}{\delta c} + 1$ :

- 1. Fix an agent  $j^* \in \arg \max_{i \text{ is low-value}} f(k_i)$ . In the first  $n_m + n_h$  periods, select  $j^*$  and every medium-value or high-value agent.
- 2. In the following  $n_l - 1$  periods, select  $j^*$  and every other low-value agent twice consecutively.
- 3. In the remaining periods, let the selection path be the same as in the proof of Lemma 2.

Under complete information, as in the proof of Lemma 2, the formation process would converge to a network only consisting of links between medium-value or high-value agents. Under incomplete information, after step 1, a star with  $j^*$  as the center and all the medium-value or high-value agents as the periphery would be formed. In step 2, a link would be formed and then severed between  $j^*$  and every other low-value agent. After that, information becomes complete and no low-value agent except  $j^*$  would ever be linked. For every medium-value or high-value agent, since  $n_m + n_h \geq \frac{c - \max_{i \text{ is low-value}} f(k_i)}{\delta c} + 1$ , the benefit from the link with  $j^*$  is at least  $\max_{i \text{ is low-value}} f(k_i) + \delta(n_m + n_h - 1)c \geq c$ , which implies that the agent has incentive to maintain the link. In addition, as  $n_h < 2$ , no link would be formed between any pair of medium-value or high-value agents, and thus the network is stable. This last fact also shows that there are no common links between the networks converged to under complete and incomplete information, and thus  $\Gamma_C$  and  $\Gamma_{IC}$  converge to entirely different networks.

Secondly, consider the following selection path when  $n_m + n_h \geq 2$ ,  $\delta$  is sufficiently close to 1 and  $n_l \geq n_m + n_h - 1$ :

- 1. In the first period, select a low-value agent and a medium-value or high-value agent; in the second period, select a second medium-value or high-value agent and the previous low-value agent; in the third period, select a second low-value agent and the previous medium-value or high-value agent;  $\dots$ ; in the  $2(n_m + n_h - 1)$ th period, select the last medium-value or high-value agent and the previous low-value agent.
- 2. In the following  $n_l - (n_m + n_h - 1)$  periods, select a medium-value or high-value agent and every remaining low-value agent.
- 3. In the remaining periods, let the selection path be the same as in the proof of Lemma 2.

Under complete information, as in the proof of Lemma 2, the formation process would converge to a network only consisting of links between medium-value or high-value agents. Under incomplete information, after step 1, a line network only consisting of links between a low-value agent and a medium-value or high-value agent is formed. After step 2, information becomes complete, and as  $\delta$  is sufficiently close to 1, the network is stable (note that  $\delta$  being sufficiently close to 1 is consistent with the condition  $n_h < 2$ ). Therefore there are no common links between the networks converged to under complete and incomplete information, and thus  $\Gamma_C$  and  $\Gamma_{IC}$  converge to entirely different networks.  $\square$

Case 2(b) in the above proposition is of particular interest, because apart from establishing the property that  $\Gamma_C$  and  $\Gamma_{IC}$  may converge to entirely different networks, it also provides insight on the particular types of network topologies that can result in such a difference. For instance, when  $n_m + n_h$  is sufficiently large, a *star* network with a low-value agent can emerge and be stable under incomplete information, which immediately implies that there are no common links with any stable network under complete information.

The above implication points to various applications. One particular case of this scenario occurring is when the selection process exhibits “preferential attachment”, i.e. agents with higher connectivity degree are more likely to be selected, as in several well-documented networks such as that of movie actors and the world wide web (see Barabasi and Albert (1999)). In this case, when a low-value agent gets “lucky” and obtains a high connectivity degree initially, it is more and more likely over time that agents with higher quality would link to this low-value agent, instead of linking between themselves.

On the other hand, when  $n_l$  is sufficiently large, a *line* network (i.e. a tree network with only one agent in each subset in the partition of  $I$ ) under incomplete information, where low-value agents and medium-value or high-value agents are linked alternately, will ensure that  $\Gamma_C$  and  $\Gamma_{IC}$  converge to entirely different networks. Similar scenarios, for example tree networks with only a few but lengthy branches, are more likely to emerge in relatively sparse and sometimes anonymous communities where agents only have the opportunity to make a limited number of links with unknown others, for example technical and biological networks (see Fricke et al (2013)). In other words, disassortativity – in this case, the tendency for agents to have small and similar connectivity degrees – is another source of significant differences between complete and incomplete information.

The above two types of phenomena both result from the interaction of incomplete information, characteristics of the selection process, and agents’ myopia.

Figures 6 and 7 provide simulation results on expected (or average) difference, both in its absolute value and as a fraction of the total number of links, between networks under complete and incomplete information<sup>8</sup>. We define difference between two networks as the number of non-common links in the networks. More precisely, the difference between networks  $\mathbf{g}_1$  and  $\mathbf{g}_2$  is given by  $|(\mathbf{g}_1 \setminus \mathbf{g}_2) \cup (\mathbf{g}_2 \setminus \mathbf{g}_1)|$ , and the total number of links is given by  $|\mathbf{g}_1 \cup \mathbf{g}_2|$ . The results indicate that there is a significant difference between networks that emerge and are stable under complete and incomplete information; this difference tends to be larger when the total number of agents  $N$  increases, and when there are more low-value agents in the group.

Figure 8 provides simulation results on the number of stable networks<sup>9</sup> under complete and incomplete information. By comparing the difference between the curves in each graph, we can see clear monotonicity: the more likely an agent’s type is low and the less likely an agent’s type is high, the more likely the formation process converges to different networks under complete and incomplete information. This is because low type agents do not connect under complete information and they may under incomplete information, and high type agents link to each other under both scenarios.

#### 4.5 Social Welfare

An alternative and very important way of comparing complete and incomplete information is to evaluate the upper bound in social welfare in the two cases. Formally, let  $W_C(\kappa)$  and  $W_{IC}(\kappa)$  be the maximum social welfare that can be achieved by a network in  $G_C^*(\kappa)$  and  $G_{IC}^*(\kappa)$  respectively. By Theorem 1, it is clear that under incomplete information, if  $\mathbb{E}[f(x)] \geq c$ , then more networks are possible and hence, the welfare upper bound under incomplete information is weakly higher than that under complete information; the following results provide sharper information.

**Lemma 3.** *Under both complete and incomplete information, if some stable network  $\mathbf{g}_1$  is a proper superset of some other stable network  $\mathbf{g}_2$ , then every agent’s payoff is weakly higher in  $\mathbf{g}_1$  than in  $\mathbf{g}_2$ . As a result,  $\mathbf{g}_1$  yields a weakly higher social welfare than  $\mathbf{g}_2$ .*

*Proof.* Note that the social welfare is the sum of each agent’s payoff. For agents having the same links in  $\mathbf{g}_1$  and  $\mathbf{g}_2$ , it is clear that they are weakly better off in  $\mathbf{g}_1$ .

Now consider an agent  $i$  whose links in  $\mathbf{g}_1$  is a proper superset of those in  $\mathbf{g}_2$ . Let  $ij_1, \dots, ij_m$  denote  $i$ ’s links in  $\mathbf{g}_1$  but not in  $\mathbf{g}_2$ . Suppose that  $u_i(k_{-i}, \mathbf{g}_1) < u_i(k_{-i}, \mathbf{g}_2)$ . It implies that there must be a permutation of  $ij_1, \dots, ij_m$ ,

<sup>8</sup> In the simulation, we assume that the payoffs from a high-value, medium-value and low-value agent are 15, 10 and 4 respectively. We assume that  $c = 5$  and  $\delta = 0.6$ . The probabilities for an agent to be of high-value, medium-value and low-value are  $(1/3, 1/3, 1/3)$  for uniform distribution,  $(4/7, 2/7, 1/7)$  for high-type environment,  $(1/7, 2/7, 4/7)$  for low-type environment and  $(1/7, 4/7, 2/7)$  for medium-type environment. In each simulation, we let the formation process run for 2500 periods. We use 100 simulations to find the average difference.

<sup>9</sup> We use the network that forms at the end of the simulation as a proxy for a stable network. In counting the number of networks we keep the agents anonymous, i.e. they only differ by their types, so that the result reflects the number of distinct network topologies rather than the number of permutations.

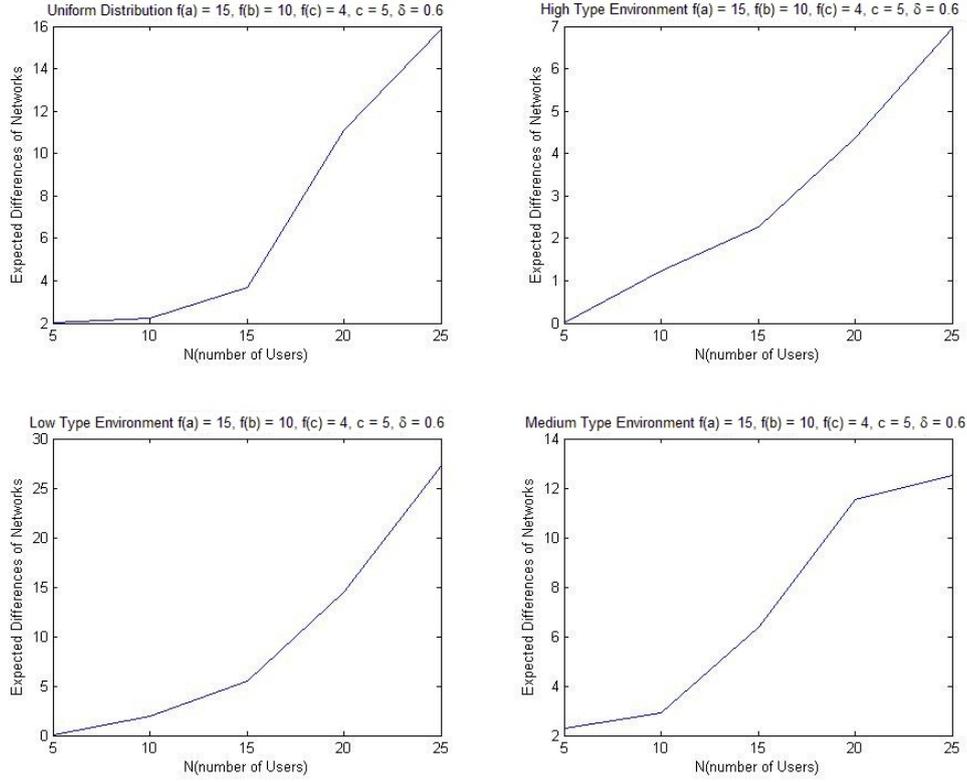


Fig. 6 Simulations: Expected Difference

denoted  $ij'_1, \dots, ij'_{m'}$ , such that for some  $m' \in \{1, \dots, m\}$ ,  $u_i(k_{-i}, \mathbf{g}_1 - ij'_1 - \dots - ij'_{m'}) > u_i(k_{-i}, \mathbf{g}_1 - ij'_1 - \dots - ij'_{m'-1})$ .

Denote  $L_i$  as an arbitrary proper subset of  $i$ 's links (including the empty set), and observe that for any  $\mathbf{g}$  and any of  $i$ 's link  $ij$ ,  $u_i(k_{-i}, \mathbf{g} - ij) - u_i(k_{-i}, \mathbf{g}) \geq u_i(k_{-i}, \mathbf{g} \setminus L_i - ij) - u_i(k_{-i}, \mathbf{g} \setminus L_i)$ . Therefore,  $u_i(k_{-i}, \mathbf{g}_1 - ij'_{m'}) - u_i(k_{-i}, \mathbf{g}_1) \geq u_i(k_{-i}, \mathbf{g}_1 - ij'_1 - \dots - ij'_{m'}) - u_i(k_{-i}, \mathbf{g}_1 - ij'_1 - \dots - ij'_{m'-1}) > 0$ , which implies that in  $\mathbf{g}_1$ , severing  $ij'_{m'}$  would strictly increase  $i$ 's payoff. But this is a contradiction with the assumption of stability, and thus it must be the case that  $u_i(k_{-i}, \mathbf{g}_1) \geq u_i(k_{-i}, \mathbf{g}_2)$ . Therefore, we can conclude that  $\mathbf{g}_1$  yields a weakly higher social welfare than  $\mathbf{g}_2$ .  $\square$

This lemma determines the social welfare relation between two stable networks when one contains the other. The following partial characterization can then be shown.

**Proposition 3.** *For any  $\kappa$ , the following properties hold:*

- 1. If  $\mathbb{E}[f(x)] < c$ , then  $W_C(\kappa) = W_{IC}(\kappa) = 0$  if  $n_m + n_h \leq 1$ , and  $W_C(\kappa) > W_{IC}(\kappa) = 0$  otherwise.
- 2. If  $\mathbb{E}[f(x)] \geq c$ , then  $W_C(\kappa) \leq W_{IC}(\kappa)$ , and:
  - a. If  $n_l = 0$ , then  $W_C(\kappa) = W_{IC}(\kappa)$ .
  - b. If  $n_l > 0$  and  $n_m + n_h = 1$ , then  $W_C(\kappa) < W_{IC}(\kappa)$  if there exists a stable wheel network among a subset of the agents.
  - c. If  $n_l > 0$  and  $n_m + n_h > 1$ , then  $W_C(\kappa) < W_{IC}(\kappa)$  if  $\delta$  is sufficiently close to 1.

*Proof.* 1: We already know from Theorem 1 that if  $\mathbb{E}[f(x)] < c$ , the network stays empty under incomplete information, yielding  $W_{IC}(\kappa) = 0$ . Therefore,  $W_C(\kappa) > W_{IC}(\kappa)$  if and only if there is some non-empty network in  $G_C^*(\kappa)$ , which is equivalent to the condition  $n_m + n_h > 1$ .

2: The claim that  $W_C(\kappa) \leq W_{IC}(\kappa)$  is a direct result from Theorem 1.

If  $n_l = 0$ ,  $G_C^*(\kappa)$  and  $G_{IC}^*(\kappa)$  are identical, and thus  $W_C(\kappa) = W_{IC}(\kappa)$ .

If  $n_l > 0$  and  $n_m + n_h = 1$ , under complete information the network stays empty, yielding a social welfare of 0. By Lemma 3, we know that in the stable wheel network, every agent's payoff is at least 0. In addition, since  $n_m + n_h = 1$  the assumption that such a network is stable implies that the total number of agents is at least 5, and that the medium-value or high-value agent must be non-singleton in this network. Thus, the two low-value

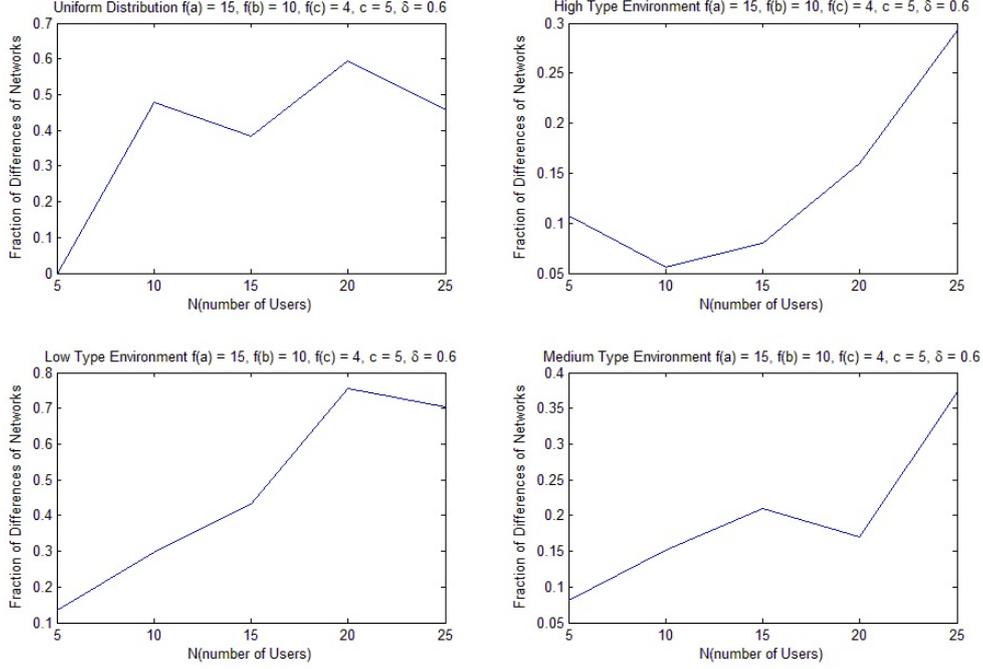


Fig. 7 Simulations: Fraction of Expected Difference

agents who link with the medium-value or high-value agent must have a strictly positive payoff, which means that the social welfare is strictly positive. Finally, it is easy to see that this network cannot be formed under complete information. Thus  $W_C(\kappa) < W_{IC}(\kappa)$ .

If  $n_l > 0$  and  $n_m + n_h > 1$ , consider the network  $\mathbf{g} \in G_C^*(\kappa)$  which yields the highest social welfare (this network must exist, since there are only finitely many networks in  $G_C^*(\kappa)$ ). Let  $\delta$  be sufficiently close to 1 such that  $\mathbf{g}$  is minimal. Thus, there exist medium-value or high-value agents  $i$  and  $j$  (in fact, when  $\delta$  is very close to 1, there is no high-value agent) such that  $ij$  is the only link  $i$  has in  $\mathbf{g}$ . Note that since  $\mathbf{g}$  is minimal,  $\mathbf{g}$  is also stable for any larger  $\delta$ .

Consider a selection path under complete information in which  $\mathbf{g}$  emerges, such that no link is formed or severed after  $ij$  is formed, and no low-value agent is selected before  $ij$  is formed. Under incomplete information, consider the following variation of this selection path: before the period in which  $ij$  is formed, insert two periods: in the first period, select some low-value agent  $i'$  and  $i$ ; in the second period, select  $i'$  and  $j$ . Since  $\mathbb{E}[f(x)] \geq c$ , we know that  $ii'$  and  $i'j$  would both be formed. As  $\delta$  gets sufficiently close to 1, the payoffs of the medium-value or high-value agents would strictly increase due to the connection to  $i'$ . Therefore  $W_C(\kappa) < W_{IC}(\kappa)$ .  $\square$

Just as in Proposition 2, this result points to particular network topologies (2(b) and 2(c)) that result in a clear welfare comparison. In the presence of low-value agents, when there is only one medium-value or high-value agent, the empty network is the only stable network that can emerge under complete information; under incomplete information, for any other network to be stable, the network must exhibit a “wheel-like” feature, i.e. apart from the medium-value or high-value agent, every agent must have at least two links. Once such a network is stable, it can be immediately shown that it yields a strictly positive social welfare. When there are more than one medium-value or high-value agents, as  $\delta$  gets sufficiently close to 1 the network that yields the highest social welfare must be minimal; then under incomplete information, there always exists a way to “insert” a low-value agent between two medium-value or high-value agents, which brings almost no change to the payoffs of the medium-value or high-value agents (since  $\delta$  is close to 1) but generates a strictly positive payoff for the low-value agent. Therefore social welfare is strictly improved.

Figure 9 generated from simulation shows the expected (or average) social welfare achieved under complete and incomplete information, in various environments. The interpretation of this figure is two-fold. On one hand, the numerical value of the difference in social welfare is mostly an artifact of the particular simulations, since it is rather sensitive to how much the values of different types of agents differ from the cost, as well as how large the discount factor  $\delta$  is. On the other hand, the significant “difference in difference” in terms of social welfare across various environments is a general and robust phenomenon. In an environment with more low type agents, the

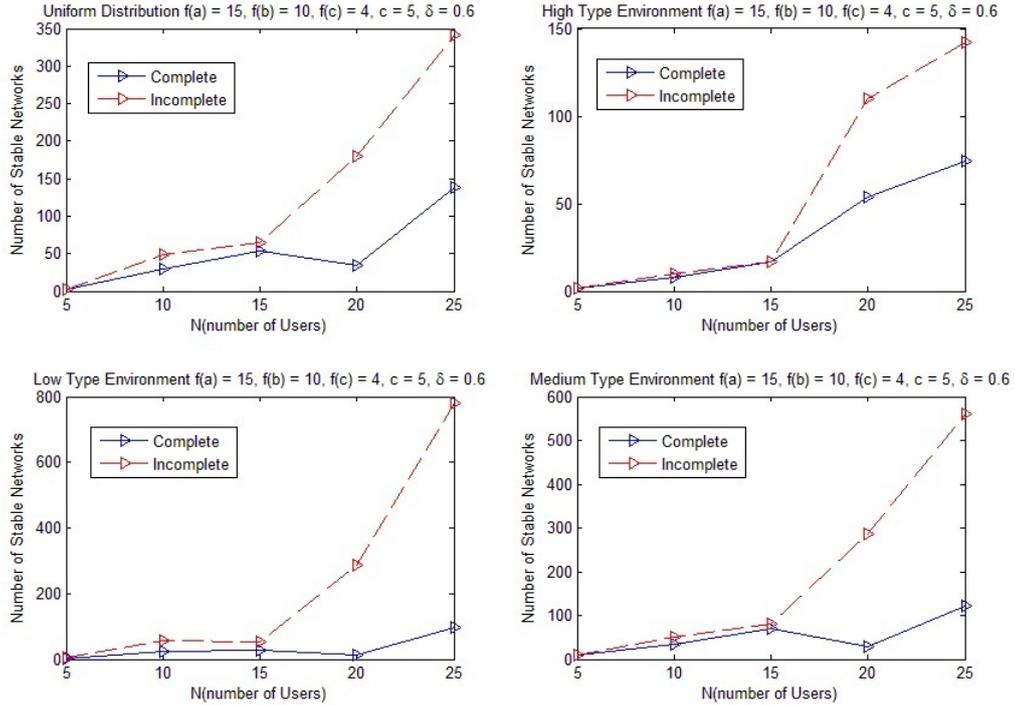


Fig. 8 Simulations: Number of Stable Networks

difference in social welfare between complete and incomplete information is larger than than in an environment with less low type agents.

## 5 Bayesian Learning

The results we have derived so far are based on the simple updating rule, which assumes that every agent's posterior belief on another agent's type is binary: either it is the degenerate belief on the true type, or the prior. Note again that such an updating rule implicitly assumes that agents can only observe their own formation history. If the agents adopt a different updating rule, which reflects either more or less available information, the formation process can exhibit a much different pattern. We discuss one such alternative in detail, which we call *Bayesian learning by formation history*.

We assume that agents can observe the entire formation history, i.e. the pair of agents selected and the resulting network structure each period, in addition to knowing the types of agents connected to themselves. However, if a link is severed, they do not observe the identity of the agent that severs the link. They then apply Bayesian updating in forming posterior beliefs. In the literature, for instance Jackson and Wolinsky (1996), Bala and Goyal (2000) and Watts (2001), since there is no uncertainty on payoffs and agents are myopic, it does not matter whether the entire formation history is observed. However, with incomplete information, what agents can observe and how they update information accordingly are crucial for shaping the ultimate network topology. The following result highlights the key difference between the simple learning rule and this alternative.

**Theorem 3.** *Under Bayesian learning by formation history, assume that in the prior type distribution, the probability of an agent being low-value is positive. Then when there are sufficiently many low-value agents, there is always a positive probability that the formation process converges to an empty network, and that information remains incomplete forever.*

*Proof.* We prove the result by construction. Consider any agent  $i$ , and consider the selection path  $\gamma_i$  that  $i$  is selected twice consecutively with a low-value agent, then selected twice consecutively with another low-value agent, and so on. We know that initially a link forms and then breaks each time  $i$  is selected with a different low-value agent. Let  $p'_m$  be the posterior probability that  $i$  is a medium-value or high-value agent after the link between  $i$  and the  $m$ th low-value agent breaks. We know that by Bayesian updating,  $p'_{m+1} = \frac{p'_m(1-p'_0)}{1-p'_m p'_0}$  with the initial condition that

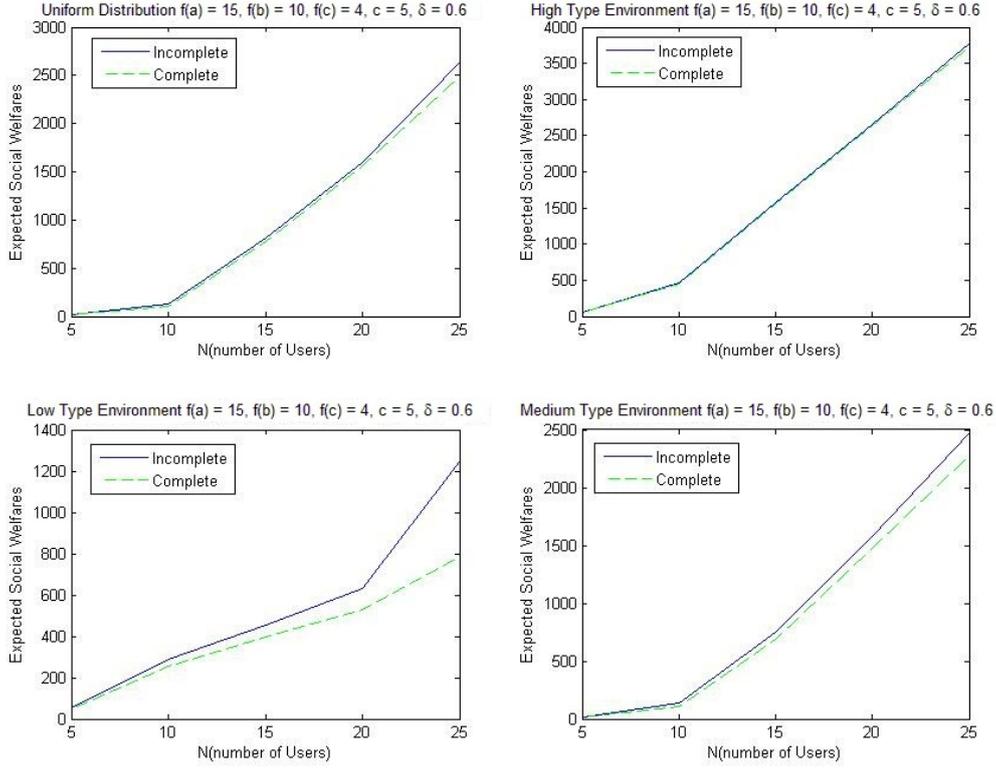


Fig. 9 Simulations: Expected Social Welfare

$p'_0$  is equal to the prior probability that an agent is medium-value or high-value. By assumption, we know that  $p'_0 < 1$ . Therefore  $\frac{p'_{m+1}}{p'_m} = \frac{1-p'_0}{1-p'_m p'_0} \leq \frac{1-p'_0}{1-p_0'^2} < 1$ , and thus there exists a sufficiently large  $n_l$  such that, following the above described selection path  $\gamma_t$ , for any agent  $j'$  who has not been selected with  $i$  before,  $\mathbb{E}[f(k_i)|B_{j'}(\gamma)] < c$ .

After  $\gamma_t$ , when  $i$  is selected with any other agent  $j$ , no link can ever form: if  $j$  has not been selected with  $i$  before,  $j$  is not willing to form a link with  $i$  since  $j$  believes  $i$  to be low-value with a high probability; if  $j$  has been selected with  $i$  before,  $i$  is not willing to form a link with  $j$  since  $i$  already knows that  $j$  is low-value. This process can be replicated for every agent  $i$ ; as a result, no link will be formed between any two agents, and the formation process converges to an empty network. Finally, since not every pair of agents has been connected before the formation process converges, information remains incomplete forever.  $\square$

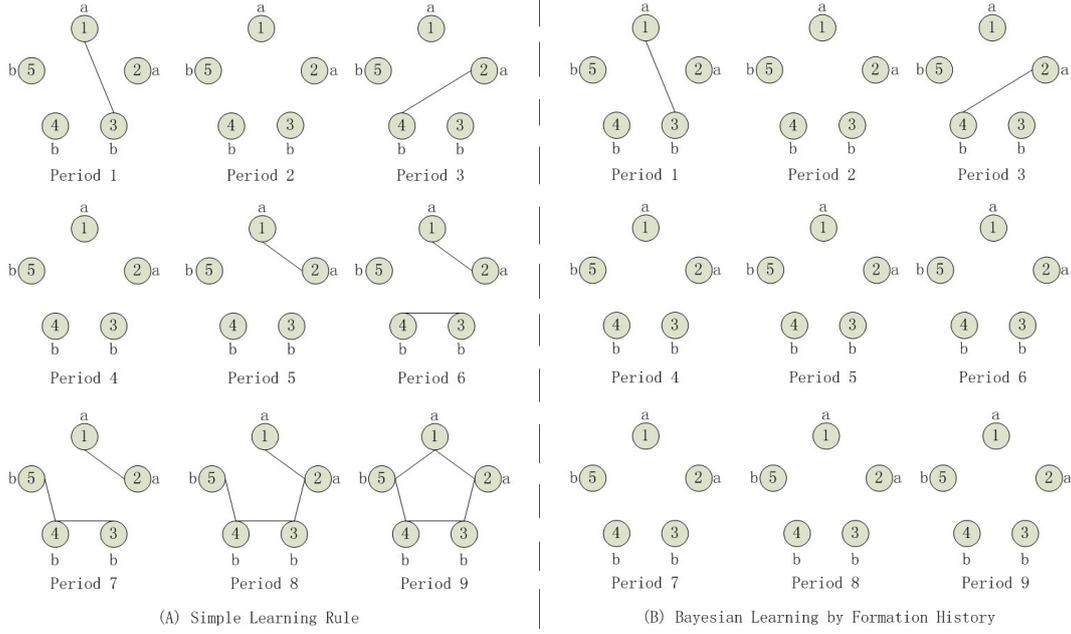
A major implication here is that Bayesian learning by formation history makes it possible for the posterior probability of an agent being of high type to fall close to 0. As a result, even if making a link with some agent  $i$  is incentivized with the simple learning rule, it may no longer be the case under Bayesian learning by formation history, given some particular selection path that would drag posterior beliefs towards  $i$  being low-value. The following example illustrates this difference.

**Example 2.** Assume the following:  $N = 5$ ,  $X = \{a, b\}$ ,  $k_i = b \forall i = 1, \dots, 5$ , and the other parameters are such that  $f(b) < c$ ,  $\mathbb{E}[f(x)] \geq c$ ,  $(1 + \delta - \delta^2 - \delta^3)f(b) \geq c$ . Let  $p = h(a)$ , and assume that  $\frac{p(1-p)}{1-p^2} f(a) + \frac{p(1-p) + (1-p)^2}{1-p^2} f(b) < c$ . Consider the following selection path in period 1-9:  $((1, 3), (1, 3), (2, 4), (2, 4), (1, 2), (3, 4), (4, 5), (2, 3), (1, 5))$ .

Under the simple learning rule, the formation process is shown in Figure 10(A). The network formed in period 9 is stable; under Bayesian learning by formation history, the formation process is shown in Figure 10(B). The network formed in period 4 is stable.

Here with the simple learning rule, agents hold the prior belief each time they are selected with another agent with an unknown type, and thus the given selection path induces a connected network at last. Yet with Bayesian learning by formation history, each agent updates from their observation to conclude that others are low-value with a sufficiently large probability, and thus are unwilling to make any link.

One implication of the above theorem and example is that more learning can sometimes be “bad”, i.e. it may lead to inefficient outcomes. Despite the specific differences brought about by an alternative updating rule, our



**Fig. 10** Connected under Simple Updating Rule vs. Empty under Bayesian Learning by Formation History

general results still hold under a range of parameters. It can be shown that if any typical network as depicted in Theorem 2 can emerge and be stable under complete information, then it can under incomplete information and Bayesian learning by formation history as well. The topological differences characterized by Proposition 2 also hold except 2(a) – now it becomes possible that  $\Gamma_C$  and  $\Gamma_{IC}$  converge to entirely different networks even when  $n_h \geq 2$ . Finally, the welfare comparison shown in Proposition 3 stays the same under this updating rule.

## 6 Conclusion and Future Research

In this paper we analyzed the network formation process under agent heterogeneity and incomplete information. Our results are in stark contrast with the existing literature: instead of restricting the equilibrium network topologies to fall into one or two specific categories, our model generates a great variety of network types. Besides *what* networks can emerge as a result of convergence, we argue that it is also important to understand *how* a network gets formed, since we want to know, for instance, *why* some agents become central and others do not. While link formation and belief formation are usually treated as two independent processes to be studied separately, we combine them in our model and show that belief formation is in fact a key factor that could facilitate or deter link formation. Even if incomplete information vanishes in the long run, its impact on shaping the network topology is persistent.

Several future research topics can be built up on the basis of our model. One of these challenges is to pin down the structure of an efficient network and implement it in a game-theoretic setting. The usual definition on efficiency in networks adopted in the literature is *strong efficiency*, i.e. a network is strongly efficient if it maximizes the sum of agents' payoffs. In general, we know that a strongly efficient network must exist (though not necessarily be unique) because the set of possible network structures is finite. However, since payoffs are heterogeneous across agents according to the type vector, the exact topology of an efficient network becomes difficult to characterize; moreover, the efficient network may not be unique because in an agent-heterogeneous environment there could be multiple ways of generating the same level of social welfare.

Most importantly, we have assumed throughout, as does most of the literature, that agents are myopic rather than forward-looking. If it is assumed otherwise that agents are foresighted and are concerned about both their current and future welfare, then the aim of analysis essentially becomes solving an agent's dynamic optimization problem in the presence of other similarly foresighted agents. One can then easily anticipate a very different evolution pattern of network topologies as well as very different stable network topologies in the limit, for now link formation does not only serve as an action of maximizing the current expected payoff, but also as a way of acquiring information for potential future benefit.

## References

- Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. *Review of Economic Studies* 78(4):1201–1236
- Bala V, Goyal S (2000) A noncooperative model of network formation. *Econometrica* 68(5):1181–1229
- Ballester C, Calvo-Armengol A, Zenou Y (2006) Who's who in networks. wanted: The key player. *Econometrica* 74(5):1403–1417
- Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512
- Corbae D, Duffy J (2008) Experiments with network formation. *Games and Economic Behavior* 64(1):81–120
- Falk A, Kosfeld M (2003) It's all about connections: Evidence on network formation. *CEPR Discussion Papers* (3970)
- Fricke D, Finger K, Lux T (2013) On assortative and disassortative mixing in scale-free networks: The case of interbank credit networks. *Kiel Working Paper* (1830)
- Galeotti A (2006) One-way flow networks: The role of heterogeneity. *Economic Theory* 29(1):163–179
- Galeotti A, Goyal S (2010) The law of the few. *The American Economic Review* 100(4):1468–1492
- Galeotti A, Goyal S, Kamphorst J (2006) Network formation with heterogeneous players. *Games and Economic Behavior* 54:353–372
- Goeree JK, Riedl A, Ule A (2009) In search of stars: Network formation among heterogeneous agents. *Games and Economic Behavior* 67(2):455–466
- Haller H, Sarangi S (2003) Nash networks with heterogeneous agents. *Discussion Papers of DIW Berlin* (337)
- Jackson MO, Watts A (2002) The evolution of social and economic networks. *Journal of Economic Theory* 106:265–295
- Jackson MO, Wolinsky A (1996) A strategic model of social and economic networks. *Journal of Economic Theory* 71:44–74
- Leung M (2013) Two-step estimation of network-formation models with incomplete information. *Working Paper*
- McBride M (2006) Imperfect monitoring in communication networks. *Journal of Economic Theory* 126:97–119
- Mele A (2010) A structural model of segregation in social networks. *NET Institute Working Papers* (10-16)
- Rong R, Houser D (2012) Emergent star networks with ex ante homogeneous agents. *GMU Working Paper in Economics* (12-33)
- Watts A (2001) A dynamic model of network formation. *Games and Economic Behavior* 34:331–341
- Zhang Y, van der Schaar M (2012) Information production and link formation in social computing systems. *IEEE Journal on Selected Areas in Communications* 30(11):2136–2145
- Zhang Y, van der Schaar M (2013) Strategic networks: Information dissemination and link formation among self-interested agents. *IEEE Journal on Selected Areas in Communications* 31(6):1115–1123