Essays on the formation of Social Networks from a game theoretical approach

A dissertation presented
by

Antoni Rubí-Barceló

to
The Department of Economics
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of

Economics

Universitat Pompeu Fabra
Barcelona, Spain
October 2007

Dipòsit legal: B.17097-2008
ISBN: 978-84-691-4465-7

Thesis advisor

Antonio Cabrales Goitia

Author

Antoni Rubí-Barceló[1]

1

C/ Rector Bernat, 3, 2-a. 07320 Santa Maria del Camí (Balears)
antoni.rubi@upf.edu –Tel: 651 53 82 36

Essays on the formation of Social Networks from a game theoretical approach

Foreword

> *"I think that the causes explaining social life are to be sought mainly in the manner according to which the associated individuals are grouped."*
>
> *Emile Durkheim (1879)*

Frigyes Karinthy published in 1929 a short story titled *"Chains"* that informally analyzed a problem that would captivate future generations of mathematicians, sociologists, physicists and economists within the field of network theory. The Hungarian author believed that, due to technological advances in communications and travel, the modern world was shrinking because social networks could grow larger and span greater distances. Nowadays, it is widely accepted that we are all connected and influenced by others and that social networks play a key role in many different settings.

The pioneer in the formal study of networks was another distinguished Hungarian: the extraordinarily prolific mathematician Paul Erdös. His work provided the starting point for practically all social-science studies in this area. In the 1950's, Erdös and his colleague Alfred Rényi defined a graph as a series of points connected by lines. This abstract entity can represent many different things. As Philip Ball announces in *"Critical Mass"*, if points are cities and lines are roads, it is easy to comprehend the rules of the graph: it is like a map. The distances and directions of roads reflect those in geographical reality. But if points are people and lines represent social relationships among them, a two-dimensional graph cannot satisfy the same interpretation rules.

However, to understand the important features of social networks, what matters is not how we choose to draw the graph but the system of connections among individuals, which is referred by the mathematicians as the topology.

One of the first studies about actual social topologies was published in the 1970's by the social scientist Mark Granovetter. In the *"Strength of Weak Ties"*, the author highlighted the main features of friendship networks, and showed that individuals form a highly connected cluster with close friends, and this cluster is linked to other groups of people through few weaker links ("acquittance" links) which are essential in maintaining the connectivity of the network. Granovetter emphasized the potential economic significance of social networks in *"Economic Action and Social Structure: The problem of Embeddedness"* in 1985. He argued that seemingly irrational behavior can be rationalized by considering the existence of social relations among agents. By taking social networks into account, he suggested that many economic puzzles could be solved. His arguments raised the interest of economists in social networks in the 1990's and, since then, Network Economics has undergone an extraordinary growth.

The main objectives of this field can be summarized in the following three points:

- *To study the effects of social networks on the agents' decisions and on individual and aggregate outcomes.*

  Social links are usually interpreted as the paths through which positive or negative externalities flow from active agents to their neighbors. Understanding how agents' strategic behavior depends on these structures of connections and how social welfare is influenced by a specific topology constitutes the main purpose of many models. These models present games, to study local network effects, in which the structure of connections is fixed. The strategy of a player is exerting a level of effort or investment, and a player's payoff increases with effort, but it also depends on the level of effort exerted by neighbors. Calvó-Armengol and Zenou (2004) is an example of this kind of models. The authors analyze the role of social networks among criminals in promoting delinquent behavior. Ballester, Calvó-Armengol and Zenou (2006) help us identify the agent that exerts a higher influence in their neighbors and, in consequence, on the aggregate outcome in any type of network. Galleoti et al (2006) presents an extensive review of this type of games.

Another common interpretation of social networks considers the structure of connections as the means by which information flows among agents. Analyzing the best network topologies to maximize the aggregate welfare is also a purpose of many models of this type. For example, it is widely known that networks of friends and relatives are highly helpful in finding employment. Calvó-Armengol and Jackson (2004) analyzes how labor market outcomes are affected by social networks.

- *To provide based incentives explanations of actual social network topologies.*

The second main objective of Network Economics is to show how actual social networks can arise from the interaction of self-interested individuals that attempt to maximize their own payoffs. A necessary preliminary step to address this objective is to have a detailed description of the form of real social networks. With the advent of the information and communication revolution, electronic databases containing thousands of records are now available and from there, data on the relations among thousands of individuals can be obtained. The availability of network information has promoted a boom in the analysis of network structures, especially in physics. Newman (2003) presents an overview of these studies about social networks (for example, film actor networks or coauthorship networks among others) and other type of networks as the internet, e-mail networks or the electricity power grid. Granovetter (1973), Newman (2001a), (2001b) and (2004) or Goyal, van der Leij and Moraga (2006) are examples of this kind of studies.

Formal modeling of network formation can be divided into two categories. On one hand, there exists the physics-based modeling, in which agents are non-strategic. Individuals in these models do not make decisions. These models describe stochastic processes of network formation and have their origins in the random graph literature. Examples can be found in sociology and recently in computer science and statistical physics literatures[2]. The other way of modeling social networks is that followed by Network Economics. These models consider that agents behave strategically. Network Economics use game theoretical tools and analyze how real social networks can naturally arise from the interaction of self-interested individuals. Jackson and Wolinsky (1996) or Bala and Goyal (2000) are two classical references of these models.

- *To analyze the (in)compatibility of societal welfare with individual incentives to form and sever links.*

The formation of Social Networks depends on the decisions of many participants

[2]See Newman (2003) for a survey.

guided by their own interests. The study of the conflict between following self-interest and maximizing the aggregate welfare has been a recurrent issue in Network Economics. A better understanding of this problem could be useful to learn how to get a better alignment of individual and social incentives. See Jackson (2003) for a review about this topic.

This thesis includes three papers that reveal new insights on the conflict between individual incentives and social welfare. However, they mainly contribute to the second objective: offering a based incentive explanation of three different phenomena related to social networks.

Chapter 1

The first chapter introduces imperfect information into the network formation analysis. Most of the papers in Network Economics assume that agents involved in a network formation game are perfectly informed about all the relevant features of the game. In spite of that, empirical work by sociologists[3] suggest that individuals have limited "horizons of observability" in that they are more likely to correctly perceive others' links or personal characteristics if they are closer in the network. The objective of the first chapter is twofold: on the one hand, we present game theoretic tools for analyzing this type of games and develop a comparative analysis among them. On the other hand, we develop an application that highlights the importance of imperfect information in the explanation of any kind of segregation in social structures. That constitutes a new network approach to the so called statistical

[3] Launman (1969), Friedkin (1983)

discrimination theory which claims that discrimination can arise because of imperfect and asymmetric information, and without the need of asymmetric preferences.

Chapter 2

Chapter 2 addresses the economic puzzle of *structural holes* in social networks. This is a notion developed by sociology (see Burt (1992)). Structural holes refers to disconnections among agents; we say that two groups are separated by a structural hole when there is a lack of social connections between them. Empirical studies highlight that people whose links bridge holes in a social structure enjoy a disproportionately large payoff. How can such results be stable when agents are free to create or sever links? This chapter attempts to complete the answer to this question developed in Goyal and Vega (2006).

Chapter 3

In the last chapter of the thesis we develop a network formation model that focuses on scientific collaboration networks. Several empirical studies[4] have recently identified the main features of coauthorship networks in Economics, Biology, Physics and Mathematics. Although there is a large body of empirical research, there is a lack of foundational theoretical models that explain how self-interested researchers organize themselves as we observe in reality. The model shows that heterogeneity among researchers and the possibility of congestion can explain the main features of real scientific collaboration topologies.

[4]Newman (2001a), (2001b) and (2004) and Goyal, van der Leij and Moraga (2006).

Contents

Acknowledgments

Four years ago, I took on the challenge of writing a PhD thesis. During this process, I have learned much about theoretical modeling of networks through discussions, presentations and trial and errors in a very enriching environment in Barcelona and, specifically, at Universitat Pompeu Fabra. In these lines, I would like to express my gratitude to the people who made this thesis possible and this period unforgettable.

I started this process with Joan Esteban. During nine fruitful months, he provided me a huge amount of ideas that marked my beginnings. Our long meetings inspired my interest in science, and specifically in theoretical research. He spent a lot of time helping me and I am extremely grateful to him.

During the rest of the process, I have mainly learned from my supervisor, Antonio Cabrales. I am especially indebted to Antonio for his advice and patience during the hard times of my personal academic development. He stimulated me when no outcomes arose and he pushed me when it was necessary. It is clear that he was crucial in the development of the dissertation and in shaping me as a potential researcher.

Many others have also contributed to this work. This thesis benefited from comments and discussions with Coralio Ballester, Antoni Calvó-Armengol, Joan de Martí, David Grandadam, Rahmi Ilkiliç, Tomás Lejarraga and Aniol Llorente. I would also like to thank the students of the Graduate Program in Economics Finance and Management, in particular, Christian Castro, Pablo Fleiss, Ricardo Nunes, Diego W. Pererira, Antoni Sureda and Federico Todeschini for very enriching talks and for their valuable friendship. I am also grateful to my office-mates, Petya Platikanova, Marta Maras, Benedetto Molinari, Aniol Llorente and Davide Debortoli for contributing to an excellent work environment. Davide always was especially generous with his technical expertise. I am also grateful to Stefano Puddu and Harald Fadinger, with whom I shared a flat for two years, for their personal support and for stimulating my intellectual development. Without them I could never finish this thesis. I am especially grateful to Tomás Lejarraga for the most interesting discussions I ever had. I wish to thank Tomás for stimulating my thoughts during our talks at some terrace in Gracia or during our long running sessions in preparation for the 2007 Barcelona Marathon. He also contributed actively to this thesis.

I want to thank Antoni Riera for his personal support and guidance in the most difficult moments of this process. In spite of his very busy agenda, he was always there for a phone call. Toni specially motivated and helped me to return to Mallorca.

I am also grateful to the secretaries at the Graduate Program in Economics Finance and Management, for assisting me in many different ways. Marta Araque, Marta Aragay, Gemma Burballa, Mireia Monroig and Anna Ventura deserve special mention. Financial support from Fundación Ramón Areces is gratefully acknowledged.

I wish to thank my family for their unconditional support. Finally, I want to express my infinite gratitude to Margalida Forteza who has been by my side during all these years and contributed to this thesis to a great extent.

A na Margalida i a la meva família.

Chapter 1

Networks and Imperfect Information

Abstract

This paper introduces imperfect information into the game theoretical approach to network formation. The purpose is twofold: First, we discuss several equilibrium concepts that allow for limited information in a network formation game. We compare them and argue that Conjectural Pairwise Stability stands out from other concepts as the most appropriate for the analysis of social networks formation. Second, we apply the selected notion of equilibrium to a novell setting. We consider a model of social network formation in which agents are defined by two features: race and human quality. Race affects the players' informational structure whereas human quality involves payoff effects. This simple model allows us to illustrate how self-interested individuals can organize themselves forming segregated societies -although these social structures are inefficient under our payoff assumptions- in a setting where the unique asymmetry among race groups comes from the informational content of the messages received from the actual state of the world. This set up constitutes a game theoretical formalization of the *statistical discrimination* argument.

1.1 Introduction

Social and economic networks play an important role in many situations. Contributions to microeconomic theory have used network structures to formalize such diverse issues as the internal organization of firms, employment search and the structure of airline routes[1]. Most of the models of network formation in these scenarios share a common characteristic: perfect information. However, many real-world situations suggest that agents who interact in a network may ignore relevant characteristics that affect the final outcome of their interaction. For example, consider a buyer-seller network[2] in which agents ignore important information to estimate the value of their commercial relationships, for example, the quality of the products exchanged. One can also think of a network of contacts in a job search[3]. In this case, agents may not be able to observe to whom the links of their neighbors lead to, and therefore be unable to estimate the value of a personal contact as a means to finding a job. In general, social networks are a clear example of the relevance of imperfect information in these frameworks; all of us are involved in a set of interpersonal connections in which not only we are unaware of relevant features of our neighbors but also about the links of a large part of them. These examples suggest a need to extend the theoretical framework of network formation to allow for limited information.

Network Economics has formalized network formation through games in which the agents interact guided by their own interests. One of the key tools in this formaliza-

[1] Garicano (2000), Brueckner, Dyer and Spiller (1992) and Calvó-Armengol (2004) are examples of these issues.

[2] See Kranton and Minehart (2001) for a perfect information version of a buyer-seller network.

[3] See Calvó-Armengol (2004)

tion is the equilibrium concept that allows us to define the principles that determine the stable game resolutions, namely, those networks that do not offer profitable possibilities of deviation to their members. One of the aims of this paper, developed in Section 1.3, is to offer a detailed analysis of different stability notions. Specifically, we analyze and compare different equilibrium concepts applicable to social networks. This implies that, in addition to imperfect information, the concepts that will be discussed must capture the requirement of mutual consent in the creation of new ties. The concepts that we present can be classified into two categories according to the nature of agents' conjectures about the unknown features of the network. The concepts in the first group consider that agents follow an objective probability function to calculate the conditional probabilities of the different possible scenarios in order to evaluate the expected marginal payoff of potential deviations. The second group assumes that these conjectures are subjective, and in principle, agents can believe that the actual state of the world is any possible situation that does not contradict the information available about the network. The Conjectural Pairwise Stability (CPS) introduced by McBride (2005) is the leading example of this kind of equilibrium concepts. The comparison among the different stability concepts is twofold: on the one hand we build bridges between the different concepts within the first group. Then, we examine the pros and cons of CPS with respect to the concepts of the first group. As a result, we argue that CPS is the most appropriate equilibrium concept for the framework considered in this paper. For this reason, a theoretical application of CPS is developed in Section 1.4.

Section 1.4 presents the other main object of the paper. We attempt to show how imperfect information can be a key factor in explaining racial, sexual or any other type of discrimination in a social structure. In fact, this is the object of the statistical discrimination theory initiated by Phelps (1972)[4]. This line of research supports that (racial, sexual, etc.) discrimination in the labor market equilibrium is a consequence of imperfect information. In particular, the factor that drives discrimination is not the existence of racial or sexual biases in the employer's preferences but the informational asymmetries between the different groups of players involved in the labor market. Using the CPS concept introduced in Section 1.3, we develop a network formation model with imperfect information. This model can be considered a new game theoretical approach to statistical discrimination theory, given that both approaches consider the same source of discrimination. Results in Section 1.4 suggest that (racial, sexual, etc.) discrimination among agents can exist in the equilibrium networks as a consequence of imperfect information in a set up with no racial or sexual biases in the agents' preferences.

1.2 General setting

This section introduces basic notation and definitions to which we will refer throughout the paper.

Network relations among players are formally represented by graphs that involve nodes and links. Nodes represent agents and links capture the relations among them. The interpretation of a link is Jackson's and Wolinsky's (1996). The authors state that a link between two individuals is undirected, since both agents benefit from

[4] and also developed by Coate (1993) , Sattinger (1998) and Berck (2001).

its existence and participate on its cost. A link can be severed unilaterally but can only be created by mutual consent of two agents. Among other applications, this theoretical treatment allows for the interpretation of a link as a personal relationship between agents. This interpretation is developed in Section 1.4.

As suggested in the introduction, the incompleteness in the informational structure of our model can have two different sources: a player may be unaware of the existence of links that are sufficiently far away in the network; and she may also ignore the type of other agents. We assume that there is a set of individual types $\mathcal{T}$ that includes two or more elements. Let t be a typical element of this set.

1.2.1 Graphs

Let $N = \{1, ..., n\}$ be the finite set of players in the population. The set of all possible networks or graphs in N is $\mathcal{G}$ and g is a typical element of $\mathcal{G}$. The subset of N containing i and j is denoted by ij and is referred to as the link ij. If $ij \in g$, the nodes i and j are directly connected. If $ij \notin g$, the nodes i and j are non directly connected. A *path* (sometimes called *chain*) in g connecting i_1 and i_l is a set of distinct nodes $\{i_1, i_2, ..., i_l\} \subset N$ such that $\{i_1 i_2, i_2 i_3, ..., i_{l-1} i_l\} \subset g$. A nonempty network $g' \subset g$ is a *component* of g, if for all $i \in N(g')$ and $j \in N(g')$, $i \neq j$, there exists a path in g' connecting i and j, and for any $i \in N(g')$ and $j \in N(g)$, $ij \in g$ implies $ij \in g'$. Let $g + ij$ denote the network obtained by adding link ij to the existing network g and let $g - ij$ denote the network obtained by deleting link ij from the existing network. A link ij is *critical* to the graph g if $g - ij$ has more components than g.

1.2.2 Value function and allocation rule

Let w be a particular *state of the world* defined as a pair (g, t). Let W be the set of all the possible states of the world. Next, we focus on the total outcome of a graph and on how this is distributed among players. The value of a graph is represented by the function $v : W \to \mathbb{R}$. The set of all such functions is $\mathcal{V}$. Notice that this function assigns a real number to any state of the world. An allocation rule describes how the value associated to each network is distributed among players. It can be written as $Y : \mathcal{V} \times W \to \mathbb{R}^n$. For simplicity, we define $\gamma \equiv (v, Y)$.

Given a pair (v, t), a graph $g \in G$ is said to be more efficient than $g' \in G$ if and only if $v(g, t) > v(g', t)$.

1.3 Equilibrium concepts

As commented in the introduction the equilibrium concepts are classified into two main categories according to the agents' conjectures about the unknown features of the network. Concepts in the first group consider that there is an objective function, known to all agents, that assigns a probability to any possible state of the world given the information available. Within this group, we distinguish between concepts applicable to static games and concepts appropriate for dynamic games of network formation. We describe and compare these concepts, and draw results that relate different equilibrium notions. Concepts in the second group, assume that each player assigns a personal probability to any feasible state of the world given the information available. Here we introduce the concept of Conjectural Pairwise Stability (CPS).

Finally, we develop a comparative analysis between both groups of concepts and conclude that CPS is the most adequate concept for our purposes in Section 1.4.

Before moving to the next section, let us introduce some necessary notation. Let m_i denote the information that agent i receives from the actual state of the world. Formally, $m_i : W \rightarrow M_i$ is the i's message function that assigns to each state of the world a message m_i in message space M_i. In a full information setting, each state of the world w generates a different message to any individual and, as a consequence, any agent is able to detect the actual state of the world after receiving the message. In our setting, agents can receive the same message for different states of the world.

1.3.1 Objective conjectures

Let $P : W \rightarrow [0, 1]$ be a function that assigns a probability to any state of the world and that is common knowledge (*i.e.* known to all agents). Then, $P(w/m_i)$ denotes the conditional probability that agent i assigns to the state of the world w given the i's information about the actual state of the world.

Static case

The strategy of a player consists of making an announcement of intended links[5]. In a static game, players simultaneously decide their strategy and the network formation game is played once. In the traditional approach for analyzing the equilibrium of games with imperfect information, players ignore the exact payoff they obtain in the actual state of the world. In order to evaluate whether a network can be sustained in

[5]See more details below.

equilibrium, the traditional approach assumes that players decide in a moment in time when the actual payoff they should receive is not yet revealed (and the imperfection in the informational structure does not allow deriving it). In the following subsection we introduce several equilibrium concepts that follow this approach.

First we present an adaptation of the Pairwise Stability Concept developed by Jackson and Wolinsky (1996). The following definition embodies three ideas: (i) players can create or sever links; (ii) the formation of a link requires the consent of both parties involved (iii) but severance can be done unilaterally.

Definition 1 *A graph g is Bayes Pairwise Stable (BPS) with respect to γ, P and t if:*

(i) for all $ij \in g$,

$$\begin{aligned} \sum_{w\in W} P(w/m_i)\ Y_i(v,w) &\geq \sum_{w\in W} P(w/m_i)\ Y_i(v,g-ij,t) \\ \sum_{w\in W} P(w/m_j)\ Y_j(v,w) &\geq \sum_{w\in W} P(w/m_j)\ Y_j(v,g-ij,t) \end{aligned}$$

(ii) for all $ij \notin g$,

$$\begin{aligned} \textit{if} \sum_{w\in W} P(w/m_i)\ Y_i(v,w) &< \sum_{w\in W} P(w/m_i)\ Y_i(v,g+ij,t) \\ \textit{then} \sum_{w\in W} P(w/m_j)\ Y_j(v,w) &> \sum_{w\in W} P(w/m_j)\ Y_j(v,g+ij,t) \end{aligned}$$

The set of BPS networks with respect to γ, P and t is denoted by $BPS(\gamma, P, t)$. In words, a network is BPS if (i) the expected marginal payoff for severing a link is negative for any agent and (ii) no pair of agents mutually benefit (in expected terms) from the creation of a link between them.

Unlike the Pairwise Stability concept, the BPS notion assumes that any player i only knows part of the actual state of the world, in particular, the information contained in m_i. For this reason, when a player i analyzes whether to sever or create a link she has to take into account all possible states of the world and weigh them by their conditional probability in order to evaluate the expected payoff consequences of her actions.

Bayes Pairwise Stability is a relatively weak notion; it can be considered as a necessary condition for network stability. As such, it admits a relatively larger set of stable allocations than other, more restrictive, definitions. The strong restrictions on potential deviations generate this weakness of the concept. A Bayes Pairwise Stable network is only robust to single-link individual eliminations and to single-link bilateral creations. Nothing is said about the possibility of severing more than one link, or about simultaneous creation and elimination of links.

Next, we present an adaptation of a normal form game that has been widely used for cases with perfect information[6] and that was first proposed by Myerson (1991). The link formation game under imperfect information can be seen as a typical Bayesian Game in which nature plays first, assigning a particular type profile (t) to the set of players following the function $P(\cdot)$. Then, any agent i receives a message m_i that contains information about the actual state of the world (this information partially describes the network structure g and the type profile t[7]). Finally, agents decide which strategy (s_i) to play knowing the function $P(\cdot)$, the payoff function f^γ and the information set m_i.

---

[6]See Goyal and Joshi (2003) , Dutta, van den Nouweland, and Tijs (1998), Harrison and Munoz (2002), and an earlier use of the game in Qin (1996).

[7]The information contained in a message m_i depends on the specific assumptions of each particular model. In the application developed in the next section, we will show a particular informational structure.

A strategy is a mapping function $s_i : m_i \rightarrow A_i$, where m_i is the player i's information set and A_i is player i's action set with $|A_i| = 2^{N\setminus\{i\}}$. Let S_i be player i's strategy set. A typical strategy of player i, s_i, consists on the set of players with whom i wants to form a link. The payoff function is the mapping $f^{\gamma} : \prod_{i\in N} S_i \times \mathcal{T} \rightarrow \mathbb{R}^N$ given by:

$$f_i^{\gamma}(s,t) = Y_i(v, g(s), t)$$

for all $s \in \prod_{i\in N} S_i$, where $\gamma \equiv (v, Y)$ and

$$g(s) = \{(ij) : j \in s_i, i \in s_j\}.$$

The above definition states that a link between i and j is formed if and only if there is mutual consent. Thus, each strategy vector s generates a unique graph $g(s)$.

Once all these elements are defined, the Bayesian Linking Game (BLG) can be stated as the tuple $\{N, \{m_i\}_{i=1}^N, \{A_i\}_{i=1}^N, f^{\gamma}, P(\cdot)\}$. The first equilibrium concept that we present for this game is the Bayes-Nash Equilibrium (BNE). A BNE for BLG is a $s^* = (s_i^*)_{i=1}^N$ such that,

$$\sum_{w\in W} P(w/m_i)\ Y_i(v, g(s^*), t) \geq \sum_{w\in W} P(w/m_i)\ Y_i(v, g(s_i, s_{-i}^*), t)$$

for all $i = 1, 2, ..., N$ and for all $s_i \in S_i$. In words, a strategy vector is BNE if no player has incentives to individually deviate by severing one or more of her links.

A clear coordination problem takes place in this game, as a consequence of the multidimensional strategy space (players can announce any combination of links they wish) combined with the requirement of mutual consent in the formation of links. For this reason, BNE is too weak a concept to single out equilibrium networks. For instance, the empty network is always a BNE. To avoid this problem we can follow analogous steps to Goyal and Joshi (2003) and, building upon Bayes Pairwise Stability, we can further require that any mutually beneficial (in expected terms) link be formed in equilibrium. Thus, the following concept arises:

Definition 2 *A network $g \in \mathcal{G}$ is a pairwise Bayes-Nash equilibrium ($PBNE$) with respect to γ, P and t if and only if there exists a Bayes-Nash equilibrium strategy profile s^* that supports g, that is $g = g(s^*)$ and for all $ij \notin g$, if*

$$\sum_{w \in W} P(w/m_i) Y_i(v, w) < \sum_{w \in W} P(w/m_i) Y_i(v, g + ij, t)$$

then

$$\sum_{w \in W} P(w/I_j)\ Y_j(v, w) > \sum_{w \in W} P(w/I_j)\ Y_j(v, g + ij, t)$$

The set of $PBNE$ networks with respect to γ, P and t is denoted by $PBNE(\gamma, P, t)$.

$PBNE$ considers more general possibilities of deviation than the *Bayes Pairwise Stability* concept previously defined. PBNE networks are robust to bilateral commonly agreed one-link creation and to unilateral multi-link severance. So, PBNE is stricter than BPS. Next, we show how closely related these two concepts are.

Calvó-Armengol and Ilkiliç (2004) shows that, for the case of perfect information, the Pairwise Stability notion and the Pairwise-Nash Equilibrium are two sides of the same token whenever the payoff function holds the $\alpha-$convexity condition. Similarly we can define a new condition that does the same job with respect to BPS and PBNE.

Definition 3 *Let $\alpha \geq 0$. The network payoff function f_i^γ is α-Bayes convex in own* *current links on $\mathcal{A} \subseteq \mathcal{G}$ if and only if:*

$$\sum_{w \in W} P(w/m_i)[Y_i(v,w) - Y_i(v, g - ij_1 - ... - ij_l, t)] \geq$$
$$\alpha \sum_{p=1}^{l} \sum_{w \in W} P(w/m_i)[Y_i(v,w) - Y_i(v, g - ij_p, t)] \quad (1)$$

for all $i \in N$, $g \in \mathcal{A}$ and for any group of links $ij_1, ..., ij_l$ such that $ij_p \in g$ for $p \in \{1, 2, ..., l\}$.

The condition for α-Bayes convexity states that the joint expected returns from a group of links already in the network is higher than the sum of the expected marginal returns of each single link, scaled by α. This is the necessary and sufficient condition on the payoff function f^γ for the set of Bayes Pairwise Stable networks and the set of Pairwise Bayes-Nash equilibrium networks to coincide as stated below.

Proposition 1 *$BPS(\gamma, P, t) = PBNE(\gamma, P, t)$ if and only if f^γ is α-Bayes convex on $BPS(\gamma, P, t)$, for some $\alpha \geq 0$.*

The proof in the Appendix derives from the fact that when α-Bayes convexity condition holds, robustness to unilateral or to multilateral link severance are equivalent, eliminating the unique difference between these two concepts of stability.

Pseudo-dynamic case

Game Theory has also analyzed the network formation process as a dynamic game in which players interact repeatedly. In these games, where a new network substitutes the previous one after the agents' decisions, it seems unreasonable to assume that players will choose their links ignoring payoffs. This motivates the consideration of a new category of equilibrium concepts in which the actual payoff is part of a player's information set. Notice that now m_i includes more information than in the previous case.

Definition 4 *A graph g is dynamically Bayes Pairwise Stable (dBPS) with respect to v, t, P and Y if:*

(i) for all $ij \in g$,

$$\begin{aligned} Y_i(v,w) &\geq \sum_{w' \in W} P(w'/m_i)\ Y_i(v, g' - ij, t') \\ Y_j(v,w) &\geq \sum_{w' \in W} P(w'/m_j)\ Y_j(v, g' - ij, t') \end{aligned}$$

(ii) for all $ij \notin g$,

$$\begin{aligned} \textit{if } Y_i(v,w) &< \sum_{w' \in W} P(w'/m_i)\ Y_i(v, g' + ij, t') \\ \textit{then } Y_j(v,w) &> \sum_{w' \in W} P(w'/m_j)\ Y_j(v, g' + ij, t') \end{aligned}$$

The interpretation of this notion is analogous to the interpretation of the BPS concept discussed above.

If we consider a dynamic process and we assume that players make decisions based on the criteria (i) and (ii) of this definition, such a process can be qualified as naive best response dynamics, because the player's behavior is myopic in the following sense: a player might delete a link making herself better off, but this deletion may lead a second player to delete a second link which leaves the first player worse off relative to the starting position. If the first player foresees this, she might choose not to sever the link in the first place. This reasoning complicates the analysis considerably. However, overlooking this reasoning may well be justified and argued as reasonable in scenarios of large populations.

Assumptions about players' memory are relevant in a hypothetical dynamic model. Under a non-memory assumption, the type of a player can be forgotten from one period to the next. Even though in small populations this assumption may be too strong, it seems reasonable in large populations. Another dimension that is crucial in the justification of the non-memory assumption is the time between iterations in the network formation procedure. The justification of the absence of memory gains power as the time between iterations increases.

1.3.2 Subjective Conjectures

Instead of an objective probability function, we now assume that agents follow their own beliefs to assign a probability to each of the states of the world. As described above, these probabilities must not contradict the information available to each player. In this category we can include the *Conjectural Pairwise Stability* concept that was introduced by McBride (2005), after adapting the existing *Conjectural Equilibrium*[8] concept to the specific setting of network formation.

Let $\pi_i : W \rightarrow [0,1]$ be the subjective probability distribution over the possible states of the world and π_i denote i's beliefs. In order to respect the original formulation of the concept we assume, as in the pseudo-dynamic case, that the payoff information is included into the message. Moreover, the original configuration also specifies that each player must be aware of her own links. The Conjectural Pairwise Stability concept can be defined as follows:

[8]See Battigalli, Gilli, and Molinari (1992) for an extended discussion of the Conjectural Equilibrium concept.

Definition 5 *A graph $g \in G$ is Conjectural Pairwise Stable with respect to value function $v \in V$, utility functions $\{Y_i\}_{i \in N}$, message functions $\{m_i\}_{i \in N}$, and beliefs $\{\pi_i\}_{i \in I}$ if*

i *for all* $ij \in g$, $Y_i(v, w) \geq \sum_{w' \in W} \pi_i(w')\, Y_i(v, g' - ij, t')$ *and* $Y_j(v, w) \geq \sum_{w' \in W} \pi_j(w')$ $Y_j(v, g' - ij, t')$,

ii *for all* $ij \notin g$, *if* $Y_i(v, w) < \sum_{w' \in W} \pi_i(w')\ Y_i(v, g' + ij, t')$ *then* $Y_j(v, w) > \sum_{w' \in W} \pi_j(w')\ Y_j(v, g' + ij, t')$ *and*

iii *for each* i, $m_i(w') = m_i(w)$ *for any* $w' \in W$ *s.t.* $\pi_i(w') > 0$.

In words, a network is CPS if: (i) no player believes that she will be better off by deleting an existing link, (ii) in any pair of non-directly linked players, at least one player believes that she will be worse off by creating a new link between them and (iii) no player's beliefs are contradicted by her signal.

In what follows we present a comparative analysis between the concepts previously discussed and CPS. First, we must notice that none of the concepts introduced above require that probabilities attributed to each state of the world are justified. Instead, they only require that these probabilities are not contradicted by their messages. Rubinstein and Wolinsky (1994) and Gilli (1999) acknowledged this drawback for the Conjectural Equilibrium concept. They consider the imposition of common knowledge of rationality as a way to refine players' beliefs so that each player must reflect optimal play on the part of the other players. This imposition is referred to as the "rationalizability refinement". Imposing common knowledge of rationality involves making signal functions (not actual actions or types) common knowledge. Individual i must justify her beliefs about j's beliefs and actions given her beliefs about j's signal; and j must in turn rationalize her beliefs about k's actions and beliefs given her beliefs about k's signal; and so on. Imposing this kind of requirements in a set up with an objective probability function (as in the first group of concepts) is more difficult to justify.

Second, in the first group of concepts we know that the probability function $P(\cdot)$ is common knowledge. This feature is subject to criticism because it is difficult to imagine that, in an framework of imperfect information, all agents know the probability function that nature follows to assign the type profile. In this sense, the use of subjective conjectures may seem more appropriate.

The above discussion suggests that CPS is the most appropriate concept for the set up proposed in this paper. The following section applies CPS to a model of segregation in social structures.

1.4 Model of segregation in social structures

1.4.1 Introduction

There are many situations that can be formalized as a model of networks with imperfect information. One can think on a buyer-seller network[9] where the quality of products is uncertain to buyers. Another situation is the job contact network[10] in which players cannot observe direct contacts of their neighbors in order to evaluate the true value of a given link. In this job market environment, there is an extensive line of

[9]See Kranton and Minehart (2001)

[10]See Calvó-Armengol (2004)

research, initiated by Phelps (1972), that studies the effects of imperfect information on the job market equilibrium and, specifically, on the discrimination properties of this equilibrium. The models proposed by this line of research[11] are known as the statistical discrimination models, and identify two sources of discrimination in the labor market. The first source of discrimination depends on employers' beliefs about workers' productivity as a function of group identity (men versus women or whites versus blacks). The second source of discrimination depends on the ability that employers have to learn the productivity of a worker. Employers might be more familiar with one group and therefore be better able to learn the actual productivity level of a worker in that group. Previous research shows that such differences can lead not only to statistical discrimination but also to ex post quality differences between ex ante identical groups[12]. We apply the second source of discrimination (generated by asymmetric employers' learning ability, not by asymmetric employer's beliefs) to the argument presented in this paper.

The discrepancy in the ability of identifying the true worker productivities can be reasonably explained by Categorization Theory[13]. This theory asserts that the human mind processes information using a limited number of categories in which agents classify their experiences and, according to a representative agent from each category, make predictions about future events. To reduce prediction error, agents reserve a category for most common events, and classify the less frequent events into a heterogeneous category. The following example illustrates this idea. Let us think

[11] See Coate and Loury (1993), Sattinger (1998) and Berck (2001), among others.

[12] See Coate and Loury (1993).

[13] See Fryer and Jackson (2003) and Mullainathan (2002).

of a labor market in which workers are characterized by two parameters: race (black or white) and productivity (high or low). The combination of race and productivity yields four types of people: white-high, white-low, black-high and black-low. Let us assume that employers are only capable of processing three types, not four. Since the employer wants to reduce the error in her predictions about productivity of workers, she will reserve two of the three categories to the most common race in her society. That is, if an employer mainly interacts with whites her three categories will be white-high, white-low and black (pooling black-high and black-low into a wider black). Such categorization yields the second source of discrimination commented above. White workers' productivity will be perfectly perceived but not blacks'. Whenever possible, employers will hire white and highly productive workers.

The network model with imperfect information specified in this section can be seen as a statistical discrimination model where agents are not employers and workers but friends, and where links are not employment contracts but friendship among agents. The remainder of this section is organized as follows: Subsection 1.4.2 specifies the basic assumptions of this application. Subsection 1.4.3 presents the results of the model and Subsection 1.4.4 concludes this part with the most important remarks.

1.4.2 Basic Setting

Next we define a simple network model to study racial[14] segregation in equilibrium due to imperfect information. A social network is said to be racially segregated if there exist different *components* that split the population in groups that can be clearly

[14]This set up can also be applied to the study of other types of social segregation.

characterized by the race of its members. We develop a stability checking analysis using the above-defined Conjectural Pairwise Stability concept. In particular, we focus on checking whether racially segregated social structures can be sustained in a CPS for different frameworks. We also analyze the consequences of imposing the rationalizability refinement in each of these frameworks.

In this model individuals are characterized by two defining features: first, agents have a group identity (*i.e.* race) that has no payoff effects[15]; second, agents have a specific individual characteristic (*i.e.* human quality) that affects the payoff. For simplicity, we assume that race can only be black or white and human quality can only be high or low.

Imperfect information is generated in two dimensions. On the one hand, agents imperfectly monitor other individuals' actions. They ignore part of the structure of connections among the rest of agents. On the other hand, the human quality of other individuals can also be unknown. Following Categorization Theory, we assume that an agents' ability to capture others' true human quality depends on the race structure of their component (see definition in Section 1.2), *i.e.* the group of agents they are mainly interacting with. In particular, if a simple majority of individuals in a given component are black, its members will detect the true type of blacks easier than that of whites[16]. We assume that, for any individual, one race is more familiar than the other. Adapting the notation introduced in McBride (2004), $x/y_r/y_{nr}$ denotes that a

---

[15]no racial asymmetry is introduced at this point.

[16]At this point, it is natural to question the case of a component integrated by an equal number of whites and blacks. We assume that the familiar race for the members of such a component is any of the two groups. We exclude the case in which an agent finds both races equally familiar. These assumptions do not affect the results of this paper.

player can observe the ties that are within the geodesic distance x (we assume that $x \geq 1$), the human quality of the players of the more familiar race within the geodesic distance y_r ($y_r \geq 0$) and the human quality of the players of the other race within the geodesic distance $y_{nr} \geq 0$. Following the previous reasoning, $y_r \geq y_{nr}$.

We do not focus on the specific network topologies that can arise inside a given component. Instead, we study the possibility of observing different components and race segregation among them in equilibrium. For this reason, we do not need to specify a complete payoff function; instead we define several simple restrictions on the value received by agents. The agents' payoff holds the following three conditions. Given any network $g \in G$,

A1 let ij be a non critical link, $Y_k(v,g,t) > Y_k(v,g+ij,t)$ for any player k in the component of players i and j.

A2 let ij be a critical link,

(i) if i is a high-type agent, $Y_i(v,g,t) < Y_i(v,g+ij,t)$ only if in the component in which player j was before the creation of link ij, the number of high-type agents is higher or equal than β times the number of low-type agents. Analogously, $Y_i(v,g,t) < Y_i(v,g-ij,t)$ only if in the component in which player j is located after severing link ij, the number of high-type agents is lower than β times the number of low-type agents. Where $\beta \in (0,\infty)$.

(ii) if i is a low-type agent, the conditions are equivalent to the previous case but with respect to γ instead of β. We assume that $0 \leq \gamma < \beta$.

A3 there are no utility interactions among different components.

Avoiding the specification of a complete payoff function allows us to achieve high generality and analytic simplicity. Moreover, it allows us to avoid making implicit assumptions about the common knowledge of a function that may be difficult to justify in an imperfect information framework. Next, we explain the intuition behind these three assumptions:

A1 is the consequence of two implicit considerations: a link is costly and the payoff obtained by a player does not depend on the exact distribution of agents in a component (non-decay assumption). This simplifies the analysis of the stable intra-component structures and allows us to focus on the analysis of segregation.

A2 captures the externalities derived from cooperation. In essence, part (i) reflects the idea that the marginal value received by a high-type agent after creating a critical link will be strictly positive only if the added component contains a sufficiently high proportion of high-type agents relative to low-type agents (an analogous argument applies for deleting a critical link). β is the critical point that defines the sign of this marginal payoff. The higher it is β the higher the negative influence an additional low-type person exerts on the rest of agents in a component. In particular, when $\beta = 1$, a simple majority of high-type people in a certain component, will attract high-type players of other components. But if $\beta > 1$, a simple majority of high-type agents in a component is not enough for attracting high-type players of other components.

On the other hand, (ii) states that the key parameter for low-type people is γ. Since $\gamma < \beta$, low-type people are less demanding when deciding whether to create or sever a critical link. In the extreme case $\gamma = 0$, low-type people always increase their payoff after adding a critical link to another component. Notice the strong implication of this assumption: it implies that racial discrimination among low-type agents will never exist in a CPS network because agents will always have incentives to create links among them. Making such an extreme assumption can be useful to focus our attention on segregation among high-type people.

A3 implies that an agent's payoff is fully determined by the state of her own component.

1.4.3 Results

In this section we study the possibilities of having racial segregation in the Conjectural Pairwise Stable networks of different frameworks. In each framework, we simply vary the payoff parameters (the informational content of the messages remain constant) to offer a complete view of how these changes affects racial segregation.

Given the payoff assumptions, in order to have a CPS network any player i must be able to belief that:

a All her links are critical.

b All her links connect her to a set of players in which the number of high-type people is higher than the number of low-type people times β if i is high type or γ otherwise.

c Any other agent is in the same component or in a different one that contains a number of high-type people lower than the number of low-type people times β if i is high type or γ otherwise.

Therefore, in order to have a CPS network, the message that any player receives from the current network must not contradict any of the above belief conditions.

Moreover, for a network to be sustained as a rationalizable CPS we further require that each player's beliefs reflect optimal play on the part of the other players.

Before the analysis, we make the following remark:

Remark 1 *For a given t and $\{Y_i\}_{i \in N}$, if the messages $\{m'_i\}_{i \in N}$ contain more information than $\{m_i\}_{i \in N}$ then a CPS network under $\{m'_i\}_{i \in N}$ is also a CPS network under $\{m_i\}_{i \in N}$, but the converse is not necessarily true.*

The logic of this remark can be summarized as follows: the higher it is the informational content of the messages, the lower it is the number of unrealistic beliefs that an agent can conjecture about the actual state of the world. Thus, if the network meets the stricter requirements for CPS under $\{m'_i\}_{i \in N}$, it will certainly meet the requirements for CPS under $\{m_i\}_{i \in N}$. For this reason, showing that racial segregation can exist in CPS networks when messages contain a relatively high amount of information (as we do next) will allow us to conclude that the results also apply for other environments with less informative messages.

Case 1: $x = \infty$, $y_r = \infty$, $y_{nr} < \infty$, $\beta > 0$, $\gamma = 0$.

Here we assume that each agent has perfect information about the network structure ($x = \infty$) and the human quality of the members of the race they are more familiar with ($y_r = \infty$) but, with respect to the members of the race they are less familiar with, agents can only observe the human quality of the members of the same component ($y_{nr} < \infty$). Notice that in this setting, messages contain extensive information and this information is asymmetric, according to Categorization Theory. On the other hand, $\gamma = 0$ means that low-type agents always increase their payoff after creating a new critical link. As commented before, this implies that racial discrimination among low-type agents will never exist in a CPS network because they will always have incentives to create links among them. This allows us to focus our attention on the study of segregation among high-type people.

The list of necessary conditions for a network to be CPS is stated in the following result. These conditions have to hold in order to avoid that messages contradict beliefs (a), (b) and (c).

Proposition 2 *Suppose information and utility parameters to be those specified in case 1. A network $g \in G$ can be sustained in a CPS if:*

i All the links are critical,

ii no high-type agent is connected to a group of players in which the number of high-type agents is lower than β times the number of low-type agents

iii there are not low-type agents in more than one component and,

iv the racial structure of the components allows that no high-type agent (or no pair of high-type agents of different components) will be able to conclude that the number of high-type people in another component is higher or equal than β times the number of low-type agents.

Proof. These four points describe the necessary conditions to avoid a contradiction of beliefs (a), (b) and (c) described above.

Since agents can observe all the ties and types of members of their own component (*i.e.* perfect information) there is no place for unrealistic beliefs about the state of the world in their own component. In consequence, conditions (i) and (ii) directly follow from the utility assumptions.

Since $\gamma = 0$, low-type agents are willing to create links with other components. If (iii) holds, no new links between low-type agents will be created.

Condition (iv) avoids the remaining possibility of deviation consisting of creating a link between a high type and a low type (or between two high types) of different components. Given the utility assumptions presented above, whenever (iv) holds, the suitable beliefs to restrain from forming a new critical link can be sustained. ■

Once we have characterized the CPS networks, we show that, even in the case with almost perfect information, racial discrimination can exist in equilibrium and in consequence, given Remark 1, segregation can also exist in settings with less informative messages. Figures A.1-A.4 represent examples of networks that can be sustained in a CPS for $\beta = 1$. In these graphs and in all the following ones, circles represent high-type agents and squares are low-type agents; the color of these geometrical forms represents the agents' race.

For the purposes of this paper, the CPS network in figure A.2 is specially interesting and illustrative. To be sustained as a CPS, high-type agents of component h_1 (or h_2) should believe that in components h_2 and h_3 (or h_1 and h_3) the number of high-type agents is lower than the number of low-type agents times β (which is assumed to be 1). Since this belief is not contradicted by the message that members of h_1(or h_2) would receive, this network can be sustained as a CPS. In such a network, high-type agents are clearly segregated. However, low-type agents are not segregated because $\gamma = 0$, as argued before (this assumption will be changed in case 2).

Note that the network in figure A.2 is inefficient because the members of components h_1 and h_2 could increase their payoff without damaging anybody else by creating a critical link between both components. Therefore, imperfect information can be seen as an additional source of inefficiency in the models of network formation[17].

At this point, it is natural to formulate the following question: Can racial segregation exist in equilibrium if common knowledge of rationality is imposed? In order to answer this question, we will focus on figure A.2. As commented before, this network can be CPS under suitable beliefs. But, these beliefs are not rationalizable; under the assumption of common knowledge of rationality and given that players in h_1 (or h_2) observe all the ties, they cannot rationally think that in h_2 and h_3 (or h_1 and h_3) the number of high-type agents is lower than β times the number of low-type agents. If that was the case, some pair of low-type agents of different components (since $\beta > 0$ we know that such a pair of players exists) would have incentives to form a link between them (because $\gamma = 0$). Therefore, it should be the case that, one of these two components has not low-type agents at all. Since some low-type agent in h_3 can be observed by the individuals in h_1 (or h_2), they can rationally conclude that in h_2 (or h_1) all the members are high type individuals. In consequence, a link between h_1 and h_2 would be formed. Then, figure A.2 is not a rationalizable CPS network.

In general, by applying the rationalizability refinement to case 1, no networks can exist with more than one component without low-type agents. Adding condition (iii)

[17]See Jackson (2004) for a detailed survey on the studies about the stability-efficiency conflict in the models of network formation.

in Proposition 1 to the last statement reduces the maximum number of components in a rationalizable CPS network to two. Networks in figures A.3 and A.4 fulfill the necessary conditions for being *rationalizable CPS* under suitable beliefs. As we can see in figure A.4, it is still possible to sustain racial segregation after applying the rationalizability refinement[18]. In this structure of connections black-high-type agents in the component h_2 remain separated from the rest of high-type agents, although it would be more efficient to have a unique component including all the agents.

As commented before, since $\gamma = 0$, there is no possibility of racial segregation to exist among low-type agents in a CPS network. Next we consider the case where $\gamma > 0$. As expected, racial segregation among low-type agents in a CPS will exist, but $\gamma = 0$ generates other interesting effects on the set of rationalizable CPS networks.

Case 2: $x = \infty$**,** $y_r = \infty$**,** $y_{nr} < \infty$**,** $\beta > 0$**,** $\gamma > 0$**.**

The only difference between case 1 and 2 is that in case 2 γ is positive. This implies that low-type agents can be not willing to create critical links. In this case, this willingness depends on the type profile of other components. For this reason, imperfect information is now an argument that can also cause racial segregation among low-type agents in a CPS network.

The list of necessary conditions for a network to be CPS is stated in the following result. These conditions have to hold in order to avoid that messages contradict

[18] Network in figure A.4 can be sustained as a rationalizable CPS network because the members of h_1 can believe that in h_2 the number of high-type agents is lower than the number of low-type agents times β. Moreover the members of h_1 can belief that this reflects optimal play of the members of h_2. In particular, since agents in component h_1 can only observe the human quality of white agents in h_2, they can believe that all members of h_2 are low-type agents. This belief is not contradicted by the messages they receive from the network.

beliefs (a), (b) and (c).

Proposition 3 *Suppose information and utility parameters to be those specified in case 2. A network $g \in G$ can be sustained in a CPS if:*

i All the links are critical,

ii no high-type agent is connected to a group of players in which the number of high-type agents is lower than β times the number of low-type agents

iii no low-type agent is connected to a group of players in which the number of high-type agents is lower than γ times the number of low-type agents and,

iv the racial structure allows that no pair of agents of different components will be able to conclude that the number of high-type people in the other component is higher or equal than the number of low-type agents times β if they are of high type or γ otherwise.

Proof. Analogous to Proposition 2. ■

Figures A.6-A.9 show networks that can be sustained as CPS under suitable beliefs considering that $\beta = 1$ and $\gamma = \varepsilon$. The network in figure A.5 cannot be CPS because members of h_2 or h_4 will have incentives to cut all the links off (since $\gamma > 0$). As we can see in figure A.7, now it is also possible to observe racial segregation among low-type agents. This network represents a completely segregated society that can be sustained as a CPS if (1) the high-type members in h_1 (or h_2) belief that in component h_2 (or h_1) the number of high-type agents is lower than the number of low-type agents times β and (2) the low-type members in h_1 (or h_2) belief that in component h_2 (or

h_1) the number of high-type agents is lower than the number of low-type agents times γ; these beliefs are sustainable because they are not contradicted by any message.

Can racial segregation still exist after applying the rationalizability refinement? It seems that the introduction of $\gamma > 0$ increases the chances of having segregated equilibrium networks. But we observe another important effect. For $\gamma > 0$, part of the population is more demanding in the configuration of the network structures. Consequently, after applying the rationalizability refinement, the number of alternatives that a player can rationally believe about the others' behavior is drastically reduced with respect to case 1. Now, a player cannot rationally believe that in any other component with more than one member, the number of high-type agents is lower than β times the number of low type people. Figure A.7 illustrates this effect:

For this network to be CPS, high-type individuals in component h_1 have to believe that in h_2 the number of high-type people is lower than the number of low-type agents times β. We claim that players in h_1 cannot rationally believe this. If that was the situation in h_2, there would not exist any high-type agent in h_2 because such a player would have incentives to cut all her links off. But players in h_1 cannot rationally believe that all agents in h_2 are low type individuals, either. If this was the case, they would also have incentives to cut their links off because $\gamma > 0$. Therefore, there is no room for unrealistic beliefs and such a network cannot be rationalizable.

In general, we can conclude from case 2 that it is not possible to observe a network with more than one component containing more than one agent as a rationalizable CPS structure. That is the reason why networks in figures A.6, A.7 and A.9 are not rationalizable CPS. This restriction eliminates the possibility of racial segregation in any rationalizable CPS network[19].

McBride (2005) showed that the refinement power of the rationalizability requirement increases with the amount of information. In case 2 we observed an additional cause that increases the power of the rationalizability refinement. Such power increased, not because of a higher amount of information, but because part of the population became more demanding striving for a positive payoff.

In that case we observed that racial discrimination is not sustainable in a CPS. But there can be infinitely many features that affect the possibilities for racial segregation. Next section introduces a small and realistic modification of the initial assumptions about the utility function. Such modification increases the possibilities of ending up in a segregated society.

Case 3: $x = \infty$, $y_r = \infty$, $y_{nr} < \infty$, $\beta > 0$ and size effect on γ.

The only modification with respect to the assumptions of case 2 is that γ now depends on the size of the components. Specifically, we assume that $\gamma > 0$ when the number of low-type agents affected (directly and indirectly) by a deviation is low; but $\gamma = 0$ if this number exceeds a given lower bound k. This modification introduces a size effect on the utility of low-type people that forces them to join each other when the size of their group is sufficiently large. This can be induced by a feeling of group membership among low-type people that only arises when this group is quantitatively important. *A priori*, such configuration should increase the social cohesion, at least

[19]The only possibility of racial segregation is the marginal case in which there is an isolated high-type player of the minority race of the component with high-type players. This is shown in figure A.8. Evidently, this is inefficient because the total payoff will increase if this isolated high-type agent connects with the main component of the network.

among low-type people. However, results suggest that other effects come into play.

Case 3 involves one minor change with respect to Proposition 2: γ can now be zero. Figures A.2, A.6 and A.7 show examples of CPS networks for case 3, considering that $k \leq 6$ and $\beta = 1$. There are no significant deviations from case 2. It is straightforward that, also figures A.6 and A.7 can be sustained in CPS under suitable beliefs. The main difference with respect to case 2 is illustrated in figure A.2: this graph cannot be CPS in case 2 because members of the h_3 component have incentives to cut all their links off. But in case 3, due to a feeling of group membership among low-type agents, this network can be sustained in a CPS.

Applying the rationalizability refinement induces an important modification on the possibilities of racial segregation in equilibrium. To analyze such scenario, let us study the social structure in figure A.7. That graph represents a totally segregated society. Figure A.7 is now rationalizable because players of any component can have the rational (but unrealistic) belief that in the other component all the members are low-type agents. Since the number of members is higher than k (remember that $k \leq 6$) in both components and no agent's type can be observed by the members of the other component, such a belief represents a possible state of the world (not contradicted by the message received by any player) in which all the agents are behaving optimally. Then, figure A.7 is a rationalizable CPS network. Notably, this network is not efficient because the payoff of all the members can increase if the whole population is in a unique component. Contrarily, networks in figures A.2 and A.6 are not rationalizable CPS (members of h_1 (or h_2) cannot rationally believe that the number of high-type agents is lower than β times the number of low-type people in h_2 (or h_1)).

Surprisingly, we find that the attraction among low-type agents induced by the size effect on γ increases the possibilities of racial segregation in such networks.

1.4.4 Concluding remarks

In this section we presented a simple model to illustrate that imperfect and asymmetric information are a source of segregation and inefficiency in the configuration of social networks.

Remark 2 *Even in cases with highly informative messages, racial segregation can exist in a CPS network from the interaction among self-interested individuals with no racial asymmetries in their preferences.*

In the three cases studied, although messages contain extensive information, equilibrium networks can still be racially segregated. Given Remark 1, this implies that segregation would also exist in equilibria of settings with less informative messages and, in consequence, that this feature is a natural outcome of our model. On the other hand, this model provides a based-incentives explanation to how discrimination can exist in a framework with no racial asymmetries in agents' preferences. This is supported by Statistical Discrimination Theory.

Remark 3 *The imposition of common knowledge of rationality generally reduces, but not always eliminate, the possibilities of racial segregation in a CPS social network.*

The refinement power of this rationalizability requirement varies across cases. In case 1, racial segregation can exist in rationalizable CPS networks but in case 2 this refinement eliminates (almost) any possibility of segregation. The reason of this variation is not the amount of information in the different settings -which remains constant- but the change in the minimum requirements for obtaining a positive payoff. This change influences the amount of beliefs reflecting optimal play by the rest of individuals in the network; and thus influences the amount of networks that can be sustained in equilibrium.

Remark 4 *Imperfect and asymmetric information is a source of inefficiency.*

A central question in the literature of network formation concerns the conditions under which networks formed by players turn out to be efficient from a societal perspective. It is not surprising that imperfect information generates inefficiency. This study extends this phenomenon to a framework of network formation. Previous literature on network formation focuses on the inherent tension between stability and efficiency[20] in network formation games, but does not consider the role of imperfect information in this framework. As we have observed in the three cases analyzed, the informational structure of the network allows for the possibility of sustaining inefficient networks in equilibrium that would never be sustained under a set up with a higher amount of information.

1.5 General conclusion

This paper constitutes one of the first attempts to introduce imperfect information into the network formation analysis from a game theoretical approach. The work can

[20] See Jackson (2004) for a complete analysis of this issue.

be split into two clearly differentiated parts. In the first part, we present several equilibrium concepts that allow for the presence of imperfect information. Those concepts are classified into two main categories according on the nature of the agents' conjectures about the unknown features of the network. We develop a comparative analysis among the different equilibrium notions and conclude that the Conjectural Pairwise Stability concept introduced by McBride (2005) is the most adequate in this social network framework. As such, we apply it to a model of racial segregation developed in the second part of the paper. The aim of this simple model is to show how imperfect and asymmetric information can cause segregation (racial, sexual, religious, etc.) in the equilibrium social structures when there exist no other *a priori* differences among groups. In fact, this is the object of Statistical Discrimination Theory initiated by Phelps (1972). Our model can be seen as the first network approach to this theory. We use the basic assumptions postulated by Statistical Discrimination Theory and reproduce the segregated equilibrium networks that it predicts. Precisely, we show how self-interested individuals can organize themselves forming inefficient segregated societies in a setting where the unique asymmetry among groups affects the informational content of the messages received from the actual state of the world. Moreover, we take advantage of the simplicity of the model to analyze how different factors can affect the possibilities of observing segregation in equilibrium.

Chapter 2

Can you really get rich because of a structural hole?

Abstract

It has been empirically shown that structural holes in social networks enable potential large benefits to those individuals who bridge them. In a pioneering paper, Goyal and Vega (2006) attempt to find a based incentives explanation of this phenomenon. But their equilibrium network is not empirically robust: in reality, social networks tend to form a topology characterized by highly connected islands with a few links among them. Our paper shows the conditions under which structural holes, and players who benefit from them, exist in these realistic network topologies. We argue that the equilibrium notion used in Goyal and Vega (2006) is not appropriate in our setting. Then, we propose the widely used Pairwise-Nash Equilibrium notion to characterize the equilibrium networks and show that agents that bridge structural holes subsist in equilibrium only when neighborhoods are small.

2.1 Introduction

Networks provide answers to many economic questions. They are often means of communication and for the allocation of goods and services not traded in markets. For example, a network of personal contacts plays a critical role in obtaining information about job opportunities[1]; networks underlie the trade and exchange of goods in non-centralized markets[2] and also define the configuration of international alliances and trading agreements[3], among others.

In environments where networks provide a platform for the flow of information, two relevant aspects need to be considered: timing and control. With respect to timing, networks can accelerate the acquisition of information generating a first-mover advantage. People can seize such opportunities or pass information along to another member of the network who can benefit from it. Research environments are examples of this phenomenon. Control is another important feature. A person that is the unique contact between two different people or groups of people benefits from the control over the flow of information, adapting it to specific strategic interests. Timing and control suggest that the payoffs an agent obtains in a network are highly dependent on the position in the network and, in particular, on the agent's capacity to bridge gaps among agents. This argument is central in Granovetter (1974) and in the notion of structural holes introduced by Burt (1992). A structural hole is a disconnection among agents on a network structure. Several authors[4] provide empirical evidence

---

[1] See Granovetter (1974), Calvó-Armengol (2004)

[2] See Kranton and Minehart (2001), Charness, Corominas-Bosch and Frechette (2007)

[3] See Goyal and Joshi (1999)

[4] See Burt (1992), Mehra, Kilduff, and Bass (2003), Podolny and Baron (1997), Ahuja (2000)

that people who bridge structural holes in social networks have significantly higher payoffs. In particular, Burt (2004) shows, in a firm environment, that compensation, positive performance evaluations, promotions and good ideas are disproportionately in the hands of people whose networks span structural holes.

Economic theory has recently started to formalize the problem. How can structural holes and their associated large payoffs differentials be sustainable when *ex-ante* identical agents strategically decide their connections? This was studied by Goyal and Vega (2006). The authors find a specific setting in which the strategic formation of links leads to a star network (*i.e.* a network whose shape resembles a star). This shape benefits the central player with an extraordinary potential for obtaining uncommonly high payoffs. Goyal and Vega (2006) formalized the empirical findings about structural holes.

Real social networks present more complex topologies. Empirical research shows that these networks usually consist on densely connected groups (also called clusters) with a few links across them. Some examples of these clusters are communities in a geographic region, departments in a firm, groups within a profession and members of a sports team. Figure 2.1 corresponds to a division of labor familiar from Durkheim (1893) and it clearly illustrates these clusters.

The network in Figure 2.1 represents a view of the world which has been often put forward to explain the 'strength of weak ties' theory[5]. According to this theory, the world consists of families or communities with very strong ties between family members. These families are connected by trade relations or occupational colleague-

[5] See, for example, figure 2 in Granovetter (1973) or figure 1 in Friedkin (1980).

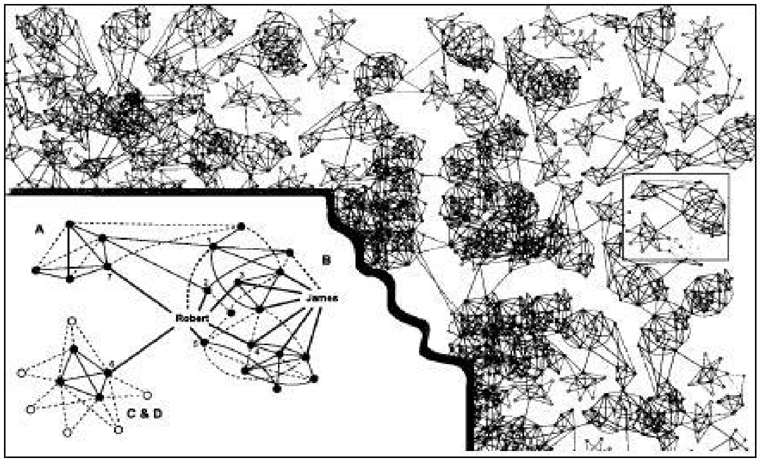


Figure 2.1: A real network with highly connected clusters

ships. However, these interfamily ties are typically weaker than intrafamily ties. In our analysis we omit the discussion about the strength of the ties, but we respect the network topology presented by these empirical studies. We do so to analyze the possibilities of having players who benefit from their position in the network by obtaining a larger payoff than the rest of players in the network.

Although these empirical observations suggest some kind of relationship between the distance (physical, ethnical, professional, etc.) among agents and the variables that affect link formation, this has not been deeply analyzed in the literature of network formation[6]. This paper introduces a basic underlying structure –in which agents are exogenously located– that respects the neighborhood-type empirical evidence by assuming that the cost of a link between two members within a neighborhood is lower than the cost of a link between members of different neighorhoods. We argue that

[6]Johnson and Gilles (2000) is an exception. This paper introduces a spatial cost topology in a network formation game.

the Bilateral Equilibrium concept used in Goyal and Vega (2006) is not appropriate for this setting because no network, except the one in which there is no link between different neighborhoods, can be sustained in equilibrium. This non-existence problem is an important drawback that motivates the consideration of a different concept: the widely used Pairwise-Nash Equilibrium (PNE)[7]. Our results show the necessary conditions for observing structural holes and agents who benefit from them in a PNE network. One of these conditions is that the size of the neighborhood should be sufficiently small. This condition generalizes the argument by Goyal and Vega (2006) that players bridging structural holes can exist in a setting with unipersonal neighborhoods. We show that players who bridge structural holes not only exist in equilibrium, but they can also obtain large payoffs differentials as illustrated by several examples at the end of the results' section.

The rest of the paper is organized as follows: in the next section we present the basic setting of the model and notation. In Section 2.3 we discuss the results in the following manner. First, we describe several preliminary results as stepping-stones to later arguments. Then, we analyze the negative consequences of applying the Bilateral Equilibrium concept to our setting and argue that such a concept is not sufficiently robust. In the last subsection of results, we develop the analysis applying the Pairwise-Nash Equilibrium concept and characterize the equilibrium network topologies under this notion. Finally, we show two examples that illustrate how agents that bridge structural holes and earn a large payoff can subsist in equilibrium. Section 2.4 is a conclusion.

[7]See Goyal and Joshi (2006), Calvó-Armengol (2004) and Bloch and Jackson (2005) for definitions and applications of pairwise-Nash networks.

2.2 The Model

2.2.1 General Set Up

The set up of the model is based on Goyal and Vega (2006).

Let $N = \{1, 2, ..., n\}$ be the finite set of ex-ante identical agents that make up the population in our model. These agents play a network-formation game with the following characteristics: the strategy of every player consists of making an announcement of intended links. $s_i = \{\{s_{ij}\}_{j \in N \setminus \{i\}}\}$ is the strategy of player i, where $s_{ij} \in \{0, 1\}$ and $s_{ij} = 1$ means that player i intends to form a link with player j, while $s_{ij} = 0$ means that player i does not intend to form such a link. Links represent pairwise relations among agents. A link between two individuals is undirected (both agents benefit from its existence and participate on its cost), can be severed by one of them unilaterally but can only be created by mutual consent of the two implied individuals. Formally, a link between two players i and j is formed if and only if $s_{ij}s_{ji} = 1$. Let $g_{ij} = 1$ denote the existence of the link between i and j while $g_{ij} = 0$ denotes the absence of such a link. Therefore, a strategy profile $s = \{s_1, s_2, ..., s_n\}$ induces a unique network $g(s)$. A *path* in g connecting i_1 and i_n is a set of distinct nodes $\{i_1, i_2, ..., i_n\} \subset N$ such that $g_{i_1 i_2} = g_{i_2 i_3} = ... = g_{i_{n-1} i_n} = 1$. All players with whom i has a path constitute the component of i in g, which is denoted by $C_i(g)$. If all the players are in a single $C_i(g)$, the network is said to be *connected*.

The utility function considers that any pair of players (i and j) connected by a path generate a unit of surplus. The distribution of this unit depends on the intermediaries between i and j. We assume that any two paths between any two players fully compete away the entire surplus (à la Bertrand competition). Therefore, an intermediary between i and j (say k) can retain part of the surplus generated by i and j if and only if this intermediary is part of all possible paths in g connecting i and j. If this condition holds, we will say that player k is an *essential* player for i and j. For example, in a star network[8] the central player is essential for any pair of players in the network because agents cannot avoid the central player in any path that links to another agent. Another extreme example is the cycle network[9], where there are no essential agents because every player has two different paths to arrive to any other individual in the cycle.

The two agents connected by a link must pay a cost $c(d)$ for it. This cost depends on the topological distance –introduced in the next subsection. Let $E(j,k;g)$ be the set of essential agents in g between j and k and let $e(j,k;g) = |E(j,k;g)|$. Then, for every strategy profile $s = (s_1, s_2, ..., s_n)$, net payoffs to player i are given by:

$$\Pi_i(s) = \sum_{j \in C_i(g)} \frac{1}{e(i,j;g)+2} + \sum_{j,k \in N} \frac{I_{\{i \in E(j,k)\}}}{e(j,k;g)+2} - \eta_i(g)c(d)$$

where $I_{\{i \in E(j,k)\}}$ is an indicator function specifying whether i is essential for j and k, and $\eta_i(g) \equiv |j \in N : j \neq i, g_{ij} = 1|$ denotes the number of players with whom i has a link in g. The first term represents i's *access payoffs* while the second term represents her *intermediation payoffs*.

In the network formation game introduced above players simultaneously announce all the links they wish to form. The resulting network is formed by the mutually announced links. This game is simple and intuitive. But, given that the creation of

[8]In a star network a unique agent is linked to all agents and no other agent has any additional link.

[9]In a cycle network all agents are linked to two other agents.

links requires the mutual consent of the two agents involved, a coordination problem takes place. As such, the game displays a multiplicity of Nash equilibria, and many different network geometries can arise endogenously. For this reason, the equilibrium concepts used in this paper are stricter than the Nash equilibrium. We use two refinements of the Nash equilibrium that allow for coordinated two-person deviations. One of these refinements, the Bilateral Equilibrium, is applied in Goyal and Vega (2006)[10]. The concept is defined as follows:

Definition 6 *A strategy profile s^b is a Bilateral Equilibrium (BE) if the following conditions hold:*

- *for any $i \in N$ and every $s_i \in S_i$, $\Pi_i(s^b) \geq \Pi_i(s_i, s^b_{-i})$*
- *for any pair of players $i, j \in N$ and every strategy pair (s_i, s_j),*

$$\Pi_i(s_i, s_j, s^b_{-i-j}) > \Pi_i(s^b_i, s^b_j, s^b_{-i-j}) \Rightarrow \Pi_j(s_i, s_j, s^b_{-i-j}) < \Pi_j(s^b_i, s^b_j, s^b_{-i-j}).$$

The networks sustained by a BE strategy profile $g(s^b)$ are robust to deviations consisting of bilateral commonly agreed one-link creation, to unilateral multilink severance and to deviations consisting of a simultaneous combination of the two previous deviations by any given pair of individuals. Therefore, the coordination possibilities of potential deviators are huge. For this reason, Goyal and Vega (2006) simplify the set of equilibrium networks to a great extent.

In the first part of the results section, we argue that this equilibrium concept is not adequate for the setting we propose in this model. One of the reasons to support this argument is that BE is too strict for this setting and (almost) no network can be

[10] Actually, they use a refinement of the Bilateral Equilibrium concept called *strict*.

sustained in a BE. Consequently, we focus on a weaker and widely used refinement of the Nash Equilibrium concept called Pairwise Nash Equilibrium. It is defined as follows:

Definition 7 *A strategy profile* s^{PN} *is a Pairwise-Nash Equilibrium (PNE) if the following conditions hold:*

∘ *for any* $i \in N$ *and every* $s_i \in S_i$, $\Pi_i(s^{PN}) \geq \Pi_i(s_i, s^{PN}_{-i})$

∘ *for any pair of players* $i, j \in N$ *and every strategy pair* (s_i, s_j) *in which* $s_i = \{\{s_{ik}\}_{k \in N \setminus \{i\}} : s_{ik} = s^{PN}_{ik} \;\; \forall\, k \neq j\}$ *and* $s_j = \{\{s_{jk}\}_{k \in N \setminus \{j\}} : s_{jk} = s^{PN}_{jk} \;\; \forall\, k \neq i\}$,

$$\Pi_i(s_i, s_j, s^{PN}_{-i-j}) > \Pi_i(s^{PN}_i, s^{PN}_j, s^{PN}_{-i-j}) \Rightarrow \Pi_j(s_i, s_j, s^{PN}_{-i-j}) < \Pi_j(s^{PN}_i, s^{PN}_j, s^{PN}_{-i-j}).$$

Networks generated by a PNE strategy profile $g(s^{PN})$ are robust to deviations of unilateral multilink severance (that is the usual Nash Equilibrium requirement) and to deviations of bilateral commonly agreed one-link creation. That is, a PNE network is a Nash Equilibrium network where no mutually beneficial link can be formed in equilibrium. Notice that PNE is weaker than BE since no simultaneous combinations are allowed.

2.2.2 Topological assumptions

As commented in the introduction, we are interested in reproducing the kind of network topologies that present densely connected clusters of agents. To this end, we assume that agents are exogenously located in an underlying structure, and that the cost of creating a link between two players depends on the distance between their locations in this underlying structure. From now on, such a distance will be referred to as topological distance. Notice that topological distances are exogenous because they are independent from the network formation game[11].

It is natural to assume that the cost of a link depends directly on topological distance. For example, if the topological distance has a geographic interpretation, the cost of a link between two nodes located in the same city will be lower than the cost of linking two people from different cities. Linking two cities that are far away involves transportation costs (time, gasoline, etc). But this concept of distance can have other interpretations. For example, it can represent the differences between professions, educational levels, religions, races, etc. One could argue that a link between two economists is cheaper than a link between an economist and a chemist. Linking two people with different professions involves communication costs (time, expression effort, etc).

In this paper we use a simple underlying structure on which agents are exogenously located. We assume that agents are distributed in neighborhoods or groups. The cost of a link depends on the neighborhood of the two implicated individuals: $c(d) = c^l$ if the two agents belong to the same neighborhood and $c(d) = c^h$ otherwise ($c^h > c^l$). As we already mentioned, we attempt to reproduce densely connected groups of players in equilibrium. This will happen when c^l is sufficiently small. For simplicity, we assume that $c^l = 0$. Let M be the total number of neighborhoods ($M \geq 3$). Let M_i be a typical neighborhood and m be the number of agents per neighborhood. We consider that $m > 1$.[12]

[11] Geodesic distance is the distance usually considered in the network formation games. The geodesic distance between two agents is defined as the number of nodes of the shortest path between them. This type of distance is endogenous because it depends on the network.

[12] The case $m = 1$ is already analyzed in Goyal and Vega (2006).

Before the analysis of the model, we review some graph-theoretic notions that will be used repeatedly throughout the paper. If a component $C_i(g)$ contains an essential player i then $C_i(g)$ can be split in, at least, three parts: two *i-groups* and i. Each player in $C_i(g)$ is included in one of these parts. Two players $j, k \in C_i(g)$ are members of different *i-groups* if i is essential for connecting them. Let $G_l(i)$ denote a typical *i-group*.

If $g_{ij} = 1$ for all pairs $i, j \in M_i$, the network among the members of M_i is said to be *complete*. A link between two players of different neighborhoods is said to be *external*. A neighborhood M_i is *essential* for connecting neighborhoods M_j and M_k if some member (not necessarily the same) of M_i lies on every path that links any member of M_j to any member of M_k. Analogously, if a component $C_i(g)$ contains some essential neighborhood M_i then $C_i(g)$ can be split in, at least, three parts: two *M_i-groups* and M_i. Each player in $C_i(g)$ is included in one of these parts. Two neighborhoods $M_j, M_k \in C_i(g)$ are members of different *M_i-groups* if M_i is essential for connecting them. Let $G_l(M_i)$ denote a typical *M_i-group*. If a neighborhood is not essential it can be extreme or non-extreme. M_i is *extreme* if all the external links of its players connect them to members of the same neighborhood. If the members of M_i have, at least, two external links to two different neighborhoods, M_i is *non-extreme*.

Finally, let us define two particular network topologies. A network with no external links is said to be *pseudo-empty*. A group of p neighborhoods constitute a *cycle* if they can be ordered in a list $M_1, M_2, ..., M_p$ such that M_p and M_1 are connected and M_i is linked to M_{i+1} for $i = \{1, 2, ..., p-1\}$ and there is no any other external link.

2.3 Results

As in Goyal and Vega (2006), agents in this model may exploit positional advantages if these provide them with the ability to block profitable bilateral interactions between pairs of players non-directly linked. Goyal and Vega (2006) shows that strategic decisions lead to the formation of a star as a prominent equilibrium architecture under the BE concept. In this structure, a single player is essential to connect every other pair of players, and this allows her to obtain, in equilibrium, a higher payoff than the rest of agents. The authors show how self-interested individuals can organize themselves forming topologies that enable potential large benefits to the individuals who bridge the "holes" among players. But, is the BE concept appropriate? In Section 2.3.2, we show that this concept is too strict to sustain a network in equilibrium (different from the pseudo-empty) in the realistic setting of this paper. This important drawback motivates the use of a different concept. PNE is a widely used alternative. In Section 2.3.3 we (i) develop the analysis of the network topologies that can be sustained in a PNE, (ii) show that individuals who bridge "structural holes" can exist in that topologies and (iii) list the conditions that must be hold for this to happen. Before these sections 2.3.2 and 2.3.3, we introduce some preliminary results.

2.3.1 Preliminary results

In this subsection we introduce two general results from the intra-group and inter-group equilibrium topologies under the two stability notions previously mentioned. We focus on network formation within and between neighborhoods.

Given that $c^l = 0$, we show the existence of BE and PNE intra-group structures and highlight their defining features.

Proposition 4 *For $c^l = 0$, an intra-neighborhood structure is a BE network if it holds the following two characteristics:*

- *there is no essential players between two agents of the same neighborhood and*
- *no pair of players can become essential after a simultaneous one-link creation and multi-link severance.*

An intra-neighborhood network is a PNE if the first condition holds.

Proof. First, notice that when there is no essential player: (i) no deviator can become essential after deleting one (or more) link(s) or after forming a new link and (ii) the access payoff is maximum. Since $c^l = 0$, then the first condition is sufficient for sustaining a PNE intra-neighborhood network. A BE network should be robust to deviations consisting of simultaneous one-link creation and multi-link severance. The second condition prevents that after such deviations some active player can become essential. Since $c^l = 0$, the access payoff is maximum without essential players and no new essential player can arise after any deviation, an intra-neighborhood network that holds these two conditions is a BE. ■

Notice that these two conditions can be easily held when there is a sufficiently high number of links among the players of the same group. In particular, a complete intra-neighborhood network always hold them, for any neighborhood size m. Therefore:

Remark 5 *For $c^l = 0$, there always exists some intra-neighborhood BE and PNE network*[13]*.*

---

[13] Contrarily to that, when $c^l = 0$ no intra-neighborhood network can be sustained as a Strict BE for $m > 3$.

Once we have proved the existence of BE and PNE intra-neighborhoods networks we focus on the analysis of the equilibrium inter-neighborhood structures. The following result is the first step on this direction.

Proposition 5 *For all m, a BE or PNE network is either connected or pseudo-empty.*

Proof. See Appendix. ■

In a BE or PNE network, a player can access either the whole population or only members of her own neighborhood. In the proof, we focus on showing that there cannot be a non-connected network with a multi-neighborhood component in a PNE or BE. Specifically, the proof shows that in this kind of networks, it is always profitable to create a new link that connects different components.

2.3.2 Bilateral Equilibrium networks

BE is a strict equilibrium concept because the equilibrium networks should be robust to bilateral deviations that allow for a very high degree of coordination among deviators. In this section we show that, BE's strictness is not appropriate for a setting with densely connected neighborhoods. This argument is developed in Theorem 1. Before presenting the result, we show a case that illustrates the irregularities derived from the use of this equilibrium concept.

Imagine two neighborhoods M_i and M_j that are directly linked and that there is no other path connecting them. The maximum number of links between these two neighborhoods in a BE network is two. Notice that for more than two links, there exists a player in one of the two neighborhoods who can delete one link (and, consequently, save c^h) without decreasing her access and intermediation payoffs. Moreover, the proof of Theorem 1 (below) shows that, for a large population, there cannot exist essential players in a BE. When this occurs, some pair of players have a strict incentive to create a link that circumvents the essential agent. Consequently, the unique remaining possibility is to have four players, $i, k \in M_i$ and $l, j \in M_j$, such that $g_{il} = g_{kj} = 1$. In such case, there are no essential players in the path connecting the two groups. But even this structure of connections cannot be sustained in a BE because players i and j have a strict incentive (see proof of Theorem 1) to delete the links g_{il} and g_{kj} and create the link g_{ij}. Notice that after this deviation takes place, players i and j become essential in the path connecting the two groups; then, we return to the previous case, and conclude that the new network is not a BE either. This phenomenon induces a cycle of deviations that raises several inconsistencies in the players' rational behavior: agents i and j would not delete their links (g_{il} and g_{kj}) and form g_{ij} if they knew that immediately after this deviation they would return to the previous status.

The argument against the use of the BE concept in this setting, is a non-existence problem:

Theorem 1 *Suppose n is large and $m > 1$. No network can be sustained in a BE, except for a pseudo-empty network when $c^h > \frac{1}{12}(3m^2 + 2m + 1)$.*

Proof. See Appendix. ■

The proof presented in the Appendix proceeds by showing that all networks other than the pseudo-empty network are not sustainable in a Bilateral Equilibrium. The arguments rely on three kinds of incentives: accessing others, gaining intermediation rents and avoiding intermediation payments. Next, we outline the intuition underlying the proof.

- *Essential players cannot exist in a BE network.*

  The starting point of the argument is that when there exists an essential player $i \in M_i$ the population can be split into, at least, two i-groups. The essential player should have, at least, two links to each of these i-groups; otherwise, another essential player $j \in M_j$ arises and then we can always find a pair of agents $i \in M_i$ and $k \in M_j$ (or $j \in M_j$ and $l \in M_i$) whose marginal payoffs for creating a link between them are proportional to population size. Therefore, marginal payoffs would be positive for a sufficiently large population and the network would not be stable. Given that i has, at least, two links to some i-group, we distinguish two different cases: if M_i is an essential neighborhood we can find two players in different M_i-groups and linked to a player in M_i who can delete these links and form a link between them. In such a way, the two players circumvent the essential player i with no additional cost because they have the same number of links as before. If M_i is non-essential, let $g_{ij} = g_{ik} = 1$ be the two links of i with some i-group. Consider the deviation in which j eliminates her link to i and forms a link to $l \in M_i$. Agent j circumvents an essential player with no additional cost and the marginal payoff of l is proportional to population size because l circumvent i to access any agent not in M_i. Therefore, marginal payoffs would be positive for a sufficiently large population and the network would not be stable.

  This result already implies that agents with disproportionately higher payoffs cannot exist in equilibrium. In this setting, a "hole" in the connection between agents will be immediately covered by several links that will vanish any possibility of enrichment.

- *In a BE, an essential neighborhood M_i cannot be connected to a M_i-group through a single neighborhood.*

  This follows from two observations. First, as commented in the previous point, there cannot be essential agents in a BE network. Second, if an essential neighborhood M_i is connected to some M_i-group through only one neighborhood (M_j), there cannot be more than two links between them; otherwise there is some agent who can delete one link (and save c^h) without generating a new essential agent. Then, the unique remaining possibility is to have two pairs of players $i, k \in M_i$ and $l, j \in M_j$ such that $g_{il} = g_{kj} = 1$. But as we mentioned above, i and j can become essential by deleting their links (g_{il} and g_{kj}) and forming a link between them. It is true that by doing so, the two players must make intermediation payments to others where none of these existed before. But we show that the intermediation rents dominate intermediation payments and as a result i and j obtain a positive marginal payoff.

  An important corollary derives from the last point: all neighborhoods should be included in some cycle and any essential neighborhood should be a member of more than one cycle of neighborhoods. The two remaining cases rule out the residual possibilities.

- *A network that contains more than one cycle of neighborhoods cannot be a BE.*

  The previous point also implies that, when there are two or more cycles of neighborhoods, any cycle must have at least one common neighborhood with another cycle. Let χ_1 and χ_2 be these two cycles. In the proof we show that it is always possible to find a deviation that transforms one of these two cycles (χ_1) into a line of neighborhoods. After such deviation the two active players become essential in the connection between the line and the cycle χ_2. Although the two active players now pay others for intermediation, new intermediation rents exceed intermediation payments; so the network with two cycles cannot be sustained in a BE.

- *A single cycle that contains all neighborhoods cannot be sustained in a BE.*

  In a cycle network, two players with external links that lie on opposite sides of the cycle have incentives to connect between them. By deleting their external links, they become central in the final line of neighborhoods. In this way, they earn large intermediation payoffs that exceed the loss from their access payoffs.

This non-existence problem is an important drawback of the BE applied to our setting. Therefore, we now focus on PNE.

2.3.3 Pairwise-Nash Equilibrium networks

As commented in Section 2.2, PNE is weaker than BE. Consequently, the set of PNE networks is larger. For instance, a cycle of neighborhoods with no essential agents can be sustained in a PNE. But in this paper we are interested on showing whether players who get a significantly larger payoff due to her strategic position on the network can exist in equilibrium for large populations. That is, we want to see whether essential players exists in equilibrium. In this section we show that this is possible and discuss the conditions that must hold in PNE.

The first result of this section limits the set of PNE networks by imposing a condition: the neighborhood size m has to be sufficiently small for an essential player to exist in a PNE network.

Proposition 6 *Suppose n and m are large. A PNE network cannot contain essential players.*

Proof. This proof is analogous to the proof of Proposition 14 contained in the proof of Theorem 1. By replacing BE with PNE, that proof is also valid for this proposition. ■

Then, essential players can only exist when the neighborhoods they connect are small. This result is in accordance with the findings by Goyal and Vega (2006) in the following sense. The authors' setting can be considered as a particular case of our model where the size of the neighborhood is extremely small, *i.e.* $m = 1$. In this case, the authors show that essential players naturally exist in equilibrium. Our Proposition 6 shows that a small neighborhood size is necessary for observing players with payoffs significantly larger than others.

The above proposition does not imply that rents of essential players cannot be large. Notice that essential players can obtain high intermediation payoffs when M is large. In fact, given that we are analyzing large populations, the case of a large M is the only case left to analyze. Next, we study the existence of essential players when M is large.

Proposition 7 *Suppose n and M are large. A PNE network cannot contain more* *than one essential player i. Moreover, a PNE topology with an essential player i cannot have two i-groups with a size proportional to M.*

Proof. By contradiction, let us assume that there are two essential players i and j in a PNE network. Notice that there must exist at least one i-group that does not contain j and one j-group that does not contain i. Take these two groups of agents and consider two players k and l that are contained in each of these two groups respectively. If the size of these groups is not proportional to M, k and l will have a strict incentive to form a link between them because, by doing so, they are able to circumvent an essential player (i and j, respectively) and reach the rest of the population which size is proportional to M. Since M is large, for any c^h, the deviation is profitable. If the size of these groups is proportional to M, the same deviation will also be profitable. Finally, if the size of only one of these two groups is proportional to M (say the i-group that does not contain j) j and k have a strict incentive to form a link between them. By doing so, j increases her intermediation payoffs obtained from the intermediation between the two groups and k circumvents an essential player to reach the rest of the population. Both marginal payoffs are proportional to M; then the deviation is profitable for a large number of neighborhoods and the network cannot be PNE, contradicting the initial statement.

On the other hand, let us assume that there is an essential player i and two i-groups with a size proportional to M. Two members contained in each of these two i-groups have incentives to form a link between them. By doing so, they circumvent the essential player i in order to reach the other group. Since the size of that group depends on M, the marginal payoff will be positive because M is assumed to be large.

■

Then, a network with two or more essential players cannot be sustained in a PNE. Next, we show that a network with an essential agent holding these conditions can actually be sustained in equilibrium. We rely on two examples to illustrate this phenomenon.

Example 1 *Imagine a network with a unique essential player i who has two links to each i-group. There is only one neighborhood in each i-group and there is no additional links. We claim that such a network is a PNE if the population is sufficiently large and the linking cost is not sufficiently low to justify an additional direct connection. Specifically, suppose that $m/6 < c^h < \frac{1}{12}[m(m(M-2)+2(m-1))+3\frac{mM-1}{M-1}]$. Then the payoffs of the central player are positive and equal to*

$$\frac{m^2(M-1)(M-2)}{6}+\frac{mM-1}{2}+\frac{(m-1)m(M-1)}{3}-2(M-1)c^h$$

If that player cuts some link off her marginal payoff would be equal to

$$c^h-\frac{m-1}{6}-\frac{(mM-m+2)(m-1)}{12}$$

which is negative for a sufficiently high M. Likewise, we conclude that the marginal payoff for cutting two links to a neighborhood is also negative. On the other hand, if a player in a peripheral neighborhood deletes one external link, then she obtains a marginal payoff equal to

$$c^h-\frac{1}{6}-\frac{mM-m-1}{12}$$

which is negative for a sufficiently large M. The creation of an additional link between two members of peripheral neighborhoods generates the following marginal payoff

$$\frac{m}{6}-c^h$$

which is negative given the conditions stated in this example. Since this is the most profitable outcome that can result from the creation of a new link, we conclude that this network is PNE under those conditions.

Apart from showing that there exist networks with essential agents that can be sustained in a PNE, this example shows that these agents can obtain a large payoff (significantly larger than payoffs obtained by the rest of agents). Yet, we can observe other equilibrium network topologies with essential agents.

Example 2 *Imagine a cycle of neighborhoods with only one group M_i that has a single player i with external links. Therefore, i would be essential connecting M_i and the rest of the population. This network is a PNE if the population is sufficiently large and $c^h > (m-1)/6$. Again, the marginal payoff for deleting a link depends negatively on M. Therefore, the marginal payoff is negative for a large M. On the other hand, the most profitable possibility for creating a new link (to add a link circumventing the essential player i) generates a marginal payoff to one of the deviators equal to:*

$$\frac{m-1}{6} - c^h$$

which is negative under the initial conditions stated above. Then, the network is PNE.

2.4 Conclusion

From sociological research we know that structural holes in social networks (that is, the lack of connections among agents) generate potential large benefits to those individuals who succeed in bridging them. This assertion has been shown empirically.

From an economic point of view, finding a based incentives explanation of the phenomenon is not a trivial issue: the existence of these agents would have to push other players to build their own bridges covering the hole, as new firms are attracted to a highly profitable market. Goyal and Vega (2006) formulated the first game theoretical approach to explain this puzzle. The authors presented a model of network formation where agents exploit positional advantages if these block profitable bilateral interactions between players who are not direct neighbors. In that model the star network arises as a prominent structure under Bilateral Equilibrium (BE). In such a network a single agent is essential to connect any pair of individuals and allows her to obtain a larger payoff than the rest of agents. This argument formalizes the above-mentioned empirical observation.

Social structures do not present the topologies obtained in the Goyal's and Vega's (2006) model. Social networks are usually formed by densely connected groups of agents interlinked among them. In this paper we show the conditions under which structural holes and players who benefit from them can exist. Our model shares the basic features with Goyal's and Vega's (2006) but assumes that agents are distributed in highly connected multipersonal neighborhoods, reproducing the empirical regularities of social networks. The first contribution of this paper is showing that the BE notion used by Goyal and Vega (2006) is not appropriate for our setting. This equilibrium concept allows for too many deviation possibilities and, as a consequence, no network (apart from the pseudo-empty) can be sustained in equilibrium. This non-existence problem is an important drawback of the BE that motivates the application of the Pairwise-Nash Equilibrium (PNE) instead. This is a weaker concept than BE and has been widely used in the literature of network formation.

The second set of results reveals the conditions for observing agents bridging structural holes (*i.e.* agents with higher rents) in PNE networks. First, we find that neighborhood size cannot be too large for these agents to exist in a PNE. This is in line with Goyal and Vega (2006). In their model, agents bridging structural holes exist in a setting that can be seen as a particular case of our model, in which neighborhoods are unipersonal. The restriction on neighborhood size does not prevent that agents bridging structural holes can enjoy a large payoff. We provide two examples that show that networks with structural holes can exist for large populations in a PNE.

Further research should shed light on additional explanations for the existence of agents with uncommonly higher rents in stable networks. For example, the existence of players with high communication abilities with different neighborhoods (that can be formalized through lower costs of link formation) could explain why there are agents that act as monopolists in the communication paths between several groups of agents, homologous to essential players presented before. This work takes a first step in the theoretical foundation of the argument that player's intrinsical features define their location and therefore, their payoffs from the network.

Chapter 3

Scientific Collaboration Networks: The role of Heterogeneity and Congestion

Abstract

We propose a dynamic model to analyze the formation of scientific collaboration networks. In this model, individuals continuously make decisions concerning the continuation of existing collaboration links and the formation of new links to other researchers through a link formation game. Ideas arrive exogenously to every node at a constant rate and agents are able to require the collaboration of one of the previously selected coauthors to increase the chances to publish articles based on those ideas. Agents are heterogeneous - they have different levels of productivity-, and they have a limited processing capability; so congestion can arise when a researcher receives a sufficiently high amount of collaboration requests. Consequently, the decision of whether to form a link must consider the trade off between the rewards (or costs) from collaborating with more (or less) productive agents and the costs (or rewards) from collaborating with more (or less) congested co-authors.

Focusing on the role of heterogeneity among agents' productivity and congestion problems derived from their limited processing capability we show how self-interested researchers can organize themselves forming the kind of scientific collaboration network topologies observed in reality.

3.1 Introduction

Social networks underlie many economic and social activities to the point that certain outcomes cannot be understood without taking into account the specific network structure. Examples and references are numerous[1]. One of the environments in which the key role of a social network is more evident is academics. In scientific production, the association with a group of competent colleagues to exchange information is a strong advantage in order to discover errors, raise research questions, and discern the appropriate ways to solve a problem. This unquestionable significance of networks in understanding scientific activity is one of the reasons that explain the extensive empirical work on this field. Today, in the advent of the information and communication revolution, data on scientific articles and researchers is stored in electronic databases containing thousands of records. With the use of these databases, empirical studies are able to reproduce coauthorship networks (in these networks a link between two researchers exists whenever there exists an article coauthored by them). From there, they are able to represent and analyze the main statistics of the collaboration among authors.

Empirical research about coauthorship networks is large[2]. Newman (2004), Newman (2001a) and Newman (2001b) analyze the defining statistics of coauthorship networks in Biology, Physics and Mathematics. Laband and Tollison (2000) focus on the importance of informal collaboration relationships in the comparison between networks in Economics and Biology. Hudson (1996) studies the reasons of the in-

[1] Calvó-Armengol and Jackson (2004) on learning about job openings through contacts or Kranton and Minehart (2001) on buyer-seller networks are only two examples

[2] Albert and Barabási (2002) offers a survey of empirical studies about any type of networks.

crease in the number of coauthors per paper in Economics. But the empirical work that most clearly shows these patterns of collaboration is Goyal, Van der Leij and Moraga (2006) (GVM hereafter). This work describes a detailed image of the features of actual coauthorship networks[3].

In spite of the great variety of empirical studies, there is a lack of foundational theoretical models that analyze how individual decisions contribute to the formation of scientific collaboration networks. To the best of our knowledge, chapter 4 in Van der Leij (2006) is the only attempt to compensate this deficiency. This paper, proposes a model that differs from Van der Leij (2006) but shares the same objective.

3.1.1 Characteristics of co-authorship networks

Before introducing the model, let us describe some of the key features of scientific collaboration networks. A surprising characteristic is the small average distance (measured by the shortest path length) between pairs of nodes. This stylized fact of social networks is captured in the famous "six degrees of separation" of John Gaure's play[4]. Scientific collaboration networks are not an exception to this phenomenon as GVM shows. The average distance in the Economics coauthorship network they analyzed was 9.47 with a total population of 33,027 nodes (*i.e.* researchers). This regularity extends to other fields. Newman (2004) shows that the average distances

[3]Although this empirical work refers to the field of Economics, we will argue that the main characteristics of co-authorship networks apply to other fields.

[4]Stanley Milgram (1967) pioneered the study of path length through a clever experiment where people had to send a letter to another person who was not directly known to them. The diameters of a variety of networks were measured. These include purely social networks, co-authorship networks, parts of the internet and parts of the world wide web. See Albert and Barabási (2002) for an illuminating account.

are 4.6 in Biology, 5.9 in Physics and 7.6 in Mathematics.

Another interesting feature of coauthorship networks refers to the degree-distribution of nodes which tends to exhibit "fat tails", *i.e.* there is a small part of the population accumulating a large proportion of links. In particular, GVM found that the 20% of most-linked authors in Economics account for about 60% of all the links. Newman (2004) shows that this phenomenon also extends to coauthorship networks in the fields of Biology, Physics and Mathematics. In each case, the distribution is fat tailed, with a small fraction of scientists having a very large number of collaborators.

GVM shows that the best-connected researchers collaborate extensively and most of their coauthors do not collaborate with each other. Moreover, the authors observe that these individuals are essential in maintaining the connectivity of the network. On the other hand, Newman (2004) found that most of the connections (64%) of an individual's shortest path to other researchers pass through the best-connected of their collaborators, and most of the remainders pass through the next-best connected.

These results lead GVM to conclude that: "the world of Economics is spanned by inter-linked stars" (an inter-linked star is a network in which some nodes connected among them accumulate a lot of links with other nodes who are not connected among themselves). Despite that there is no such conclusion referred to co-authorship networks in other fields, the similarity in the general results showed in Newman (2004) suggests a similar pattern in Biology, Physics and Mathematics. Moreover, GVM analyzes the evolution over the last thirty years and concludes that such a structure is stable over time.

3.1.2 Preview of the model and results

We show that the effects of two simple driving forces can explain the formation of scientific collaboration networks with an interlinked star topology. These two forces are caused by the heterogeneity among researchers and their limited processing capability. Specifically, we propose a dynamic model in which individuals periodically make decisions concerning the continuation of existing collaboration links and the formation of new links to other researchers through a link formation game. Once the network has been constituted, ideas arrive from outside of the network to every node at a constant rate and agents are able to require the collaboration of one of the previously selected coauthors to increase the chances to publish articles based on those ideas. As commented above, agents are heterogeneous –they have different levels of productivity–, and they have a limited processing capability; so congestion can arise when a researcher receives a sufficiently high amount of collaboration requests. Consequently, the decision of whether to form a link must consider the trade off between the rewards (or costs) from collaborating with more (or less) productive agents and the costs (or rewards) derived from more (or less) congested coauthors.

After defining the rules of the network formation game, the timing and procedure of the publication process, the payoff function and the equilibrium concept in section 3.2, we assume that the process reaches a Steady State and we characterize it. Then, we study which kind of network topologies can be sustained in Steady State. This analysis is divided into two parts: first, in section 3.3.1 we show several results that sharply narrow the set of potential equilibrium networks. In particular, these results show that the concentration of links towards the researchers with a higher productivity is a natural feature of equilibrium topologies in our model. Notice that this distribution of links generates a clearly unequal situation in which some agents do not receive ideas from others and where others receive numerous collaboration requests. Roughly speaking, we show that, in equilibrium, highly productive agents receive the minimum number of links necessary to exhaust their processing capability. Although the commented networks are highly unequal, we find the necessary conditions to reproduce this kind of equilibrium networks even for a highly homogeneous population of researchers.

In section 3.3.2 we go one step further in the simplification of the set of potential Steady State networks, and identify the conditions under which a single topology with the basic characteristics of actual scientific collaboration networks can be sustained as the unique equilibrium network of our model. Thus, our model naturally reproduces the scientific collaboration patterns observed in reality.

3.1.3 Literature Review

Theoretical models of social network formation can be classified into two groups. On one hand, there are the physics-based modeling of society. This approach treats agents as if they were just matter. That is, agents are non-strategic. This set has its origins in the random graph literature and has examples in sociology and recently in computer science and statistical physics. References of this kind of models are abundant[5] but we will focus on two of them. Jackson and Rogers (2006) proposes a nice, simple and general model of network formation. The authors combine random

[5]See Newman (2003) for a survey. Some examples are Watts (1999), Cooper and Frieze (2003) or Price (1976).

meeting and network-based meeting in a natural manner and analyze the relevance of these two forces in determining the formation of different kinds of networks (scientific collaboration structures are one of them). The second model we focus on is Arenas et al (2003). The authors present a stylized model of a problem-solving organization –whose internal communication structure is given by a network– that can suffer congestion. The authors develop a design problem to determine which kind of network architectures optimizes performance for any given problem arrival rate. Contrarily to our model, the network is fixed and players are non strategic.

The second classification of models involves strategic formation of networks and use game theoretic tools. That is, there is no exogenous prescription of how the network is formed but there is a definition of the rules of the game that agents have to play to form the network (see Jackson (2004) for a survey of this type of models). The model presented here belongs to this group of models. As introduced above, the work that more closely relates to our model is chapter 4 in Van der Leij (2006). This author also attempts to develop a theoretical model to explain the empirical regularities of research collaboration networks. In both models, heterogeneity across researchers plays a key role in explaining the results. Contrarily to our paper, Van der Leij constructs a static model in which the cost of link formation and the specific academic rewards scheme affect the equilibrium network topologies. Our model is dynamic and involves the possibility of congestion as the key factor (joint with agents' heterogeneity) for obtaining the results. Moreover, we do not require a minimum degree of heterogeneity among researchers (as Van der Leij (2006)) to reproduce the equilibrium networks observed in reality.

3.2 General setting

Let N be the set of nodes, interpreted as researchers, with $n = |N|$ and let i and j be typical members of this set. We assume that n is finite and arbitrarily large. Networks are modeled as directed graphs. A directed graph on N is an $N \times N$ matrix g where entry g_{ij} indicates whether a directed link exists from node i to node j; $g_{ij} = 1$ indicates the existence of such a directed link and $g_{ij} = 0$ indicates the absence of this directed link. Notice that we do not impose any specific value for g_{ii}; in particular, it is possible to have $g_{ii} = 1$ (see interpretation below). For any node $i \in N$, let $N_i(g) = \{j \in N : g_{ji} = 1\}$ be the set of players that have a link to i and $\eta_i(g) = |N_i(g)|$ denote the in-degree of i. On the other hand, let $M_i(g) = \{j \in N : g_{ij} = 1\}$ be the set of destinations of the links of i and $\mu_i(g) = |M_i(g)|$ denote the out-degree of i. Notice that $\eta_i(g)$ and $\mu_i(g)$ have to be natural numbers. We impose that $\mu_i(g) \geq 1$.[6]

Time is modeled continuously. However, for descriptive convenience, we split time in periods. The object of the agents in this model is to publish papers. This is their only source of payoffs. Specifically, a publication provides one unit of payoff, which is equally split among all its coauthors. A publication starts with an idea. Researchers receive ideas from outside the network at an independent positive rate ρ. These ideas are *open*, in the sense that they need to be processed to become a publication. Immediately after receiving these open ideas, agents send them to some previously selected destination. Agents can also choose to retain ideas. Here it is where the

[6]when $\nexists j \neq i$ such that $g_{ij} = 1$, then g_{ii} must be necessarily 1. When $\exists j \neq i$ such that $g_{ij} = 1$, g_{ii} can also be 1.

network plays its role, because a researcher i can only send her open ideas to some agent $j \in M_i(g)$ ($i \in M_i(g)$, *i.e.* $g_{ii} = 1$, means that agent i retains (part of) her own open ideas). We assume that all agents in $M_i(g)$ have the same probability of being selected as destination of a particular open idea obtained by i. The node chosen as destination is the researcher in charge of starting the publication process of this idea.

At any time, several open ideas may "wait" to be processed by certain node (as in a queue) because we assume that researchers have a limited processing capability. Specifically, there is an upper-bound in the number of ideas a node can process per period which we normalize to one[7]. Therefore, if a researcher receives a sufficiently high amount of collaboration requests (*i.e.* links), queues will be formed. Given this possibility, agents are provided with a decision rule to select the open idea they will process from their stock. We will take the simplest rule, that is, all open ideas in a queue have the same probability of being selected. Researchers also have a limited storage capability. In particular, each agent forgets an open idea with probability q. For this reason, not all open ideas received by a node will be finally processed.

Once an open idea is chosen to be processed, two outcomes are possible: it is published or rejected forever. In this setting, a publication can have at most two coauthors: the researcher who initially gets the open idea from out of the network and the destination of this open idea (notice that these two nodes can coincide). When an open idea is processed, the probability of being published by the coauthors (author) depends on their (her) talent. Let h be the vector of talent endowments and h_i be the i-th element of this vector interpreted as the agent i's amount of talent. We

[7]Notice that this assumption is not restrictive at all since the length of the period is not specified.

assume that h_i is exogenous, randomly generated following the probabilities described by a continuous distribution function[8] and that $h_i > 0$ for all $i \in N$. Vector h is fixed throughout the game. The relationship between talents and publication probabilities is determined by $f(\cdot)$. This is a strictly increasing probability function, holding $f(0) = 0$. This implies that the higher it is the amount of talent of a researcher/node the higher it is the probability of publishing the ideas being processed. So, $f(h_i + h_j)$ is the probability of publishing a particular idea processed by i (or j) and previously sent by j (or i). Notice that h_i can also be interpreted as the agent i's productivity.

Therefore, agents are characterized by two defining features: the endogenous size of their queue of open ideas waiting to be processed and the exogenous amount of talent.

3.2.1 Network formation game and timing

As commented before, time is modeled continuously but described in periods. At the beginning of a period, collaboration links are configured through the following network-formation game: all players $i \in N$ simultaneously announce the direct and directed links they wish to have either as origin or as destination. Formally, $S_i = \{0,1\}^{2n-1}$ is i's set of pure strategies. Let

$$s_i = (s^i_{i1}, s^i_{i2}, ..., s^i_{ii}, ..., s^i_{in}, s^i_{1i}, ..., s^i_{i-1,i}, s^i_{i+1,i}, ..., s^i_{ni}) \in S_i.$$

Then, $s^i_{ij} = 1$ if and only if player i wants to set up a directed link from i to j (and thus $s^i_{ij} = 0$, otherwise). As commented before $s^i_{ii} = 1$ is possible. A link,

[8] Notice that this implies that the probability of two agents having exactly the same amount of talent is zero.

which is assumed to be costless, from player i to player j is formed if and only if $s_{ij}^{i}s_{ij}^{j} = 1$. That is, we assume that mutual consent is needed to create a link. Let $S = S_1 \times ... \times S_n$. A pure strategy profile $s = (s_1, ..., s_n) \in S$ induces a directed network $g(s)$.

Once the new network is formed, any agent (say i) receives open ideas at a rate ρ and sends them to one selected destination. Simultaneously, node i selects and processes open ideas from her stock (if any) at a rate of one idea per period. At the end of each period, all stocked open ideas are forgotten with probability q. Just before the end of a period, the stock of open ideas of all nodes is updated.

3.2.2 Steady State analysis and payoff function

Suppose that the process reaches a Steady State. There are two defining properties of the Steady State: each researcher's stock of open ideas is constant and the network is stable.

Let o_i be the Steady State stock of open ideas waiting to be processed by node i. Under stationarity, the number of open ideas in stock behaves as a Markov process and the arrival and departure of ideas from and to each node i follow Poisson processes. Given that in Steady State all open ideas that arrive to a node eventually depart from it in finite time, we must observe that the arrival rate of open ideas must be equal to its departure rate. That is:

$$\rho \sum_{l \in N_i(g)} \frac{1}{\mu_l} = \begin{cases} 1 + qo_i & \text{, if } o_i \geq 1 \\ o_i(1+q) & \text{, otherwise} \end{cases} \quad \forall\, i \in N$$

The arrival rate of ideas to agent i is equal to the sum, over all nodes sending to i in

g, of the expected number of ideas they receive from out of the network per instant of time (ρ) times the probability of sending ideas to i. The departure rate is the sum of the processing rate and the rate of open ideas that node i forgets. Notice that when the stock of open ideas is lower than one the processing capability of a node will be restricted. In such a case, only o_i open ideas can be processed per period (on average). A node i is said to be *congested* if the Steady Staten stock of open ideas o_i is higher than one.

From the last expression, we write the Steady State stock of open ideas of a node as:

$$o_i = \begin{cases} \frac{\rho(\sum_{l \in N_i(g)} \frac{1}{\mu_l}) - 1}{q} & \text{, if } \sum_{l \in N_i(g)} \frac{1}{\mu_l} \geq \frac{q+1}{\rho} \\ \frac{\rho(\sum_{l \in N_i(g)} \frac{1}{\mu_l})}{1+q} & \text{, otherwise} \end{cases} \quad \forall\, i \in N \tag{1}$$

The stock of open ideas of a node is completely determined by the network structure (g) and also by q and ρ.

The other defining feature of the Steady State is network stability. Before defining the stability concept we introduce the payoff function. As commented above, researchers only obtain profits from the publication of ideas. For a given network structure g, the following expression defines the expected payoff agent i obtains per period when the stock of open ideas is constant for all agents[9]:

$$\Pi_i(g) = \Theta(i)[\sum_{l \in N_i(g) \setminus i} \frac{1}{\mu_l} cf(h_l + h_i) + g_{ii} \frac{1}{\mu_i} f(h_i)] + \frac{1}{\mu_i} \sum_{l \in M_i(g) \setminus i} \Theta(l) cf(h_l + h_i) \text{ with } c = \frac{1}{2} \tag{2}$$

[9]By using the per-period expected payoff to analyze the incentives to deviate from a particular network, we do not consider transition effects from one network to another. This simplification has minor implications, especially for cases in which transition does not last in time and/or the discounting rate is near to one.

$$\text{where } \Theta(i) = \begin{cases} \frac{1}{\sum_{k \in N_i(g)} \frac{1}{\mu_k}} & \text{, if } \sum_{k \in N_i(g)} \frac{1}{\mu_k} \geq \frac{q+1}{\rho} \\ \frac{\rho}{1+q} & \text{, if } 0 < \sum_{k \in N_i(g)} \frac{1}{\mu_k} < \frac{q+1}{\rho} \\ 0 & \text{, if } \sum_{k \in N_i(g)} \frac{1}{\mu_k} = 0 \end{cases}.$$

Agent i's publications can derive from her own stock of open ideas or from the open ideas that i previously sent to other researchers. These two different origins are represented by the two main parts of (2), respectively. For $g_{ji} = 1$, $\Theta(i)\frac{1}{\mu_j}$ can be interpreted as the Steady State probability that an open idea coming from node j is chosen to be processed by i. Given that all ideas in a given stock have the same probability of being selected, this probability is obtained by multiplying the share of ideas coming from j with respect to all ideas received by researcher i ($\frac{\frac{1}{\mu_j}}{\sum_{k \in N_i(g)} \frac{1}{\mu_k}}$) by the expected number of ideas that node i processes per period (1 if $\sum_{k \in N_i(g)} \frac{1}{\mu_k} \geq \frac{q+1}{\rho}$ and o_i otherwise[10]). On the other hand, $cf(h_j + h_i)$ (or $f(h_i)$) can be interpreted as the expected payoff obtained by each of the coauthors of a processed idea.

Notice that there are three main factors affecting agent i's expected utility: c influences the decision between retaining open ideas or sending them to other authors, the queue size (included in $\Theta(l)$ or $\Theta(i)$) affects the probability of processing an open idea, and the coauthor's amount of talent (h_l) affects the probability of publishing processed ideas.

The network stability concept used in this model is *Pairwise-Nash Equilibrium* (PNE hereafter). In a PNE network, no player must have incentives to deviate unilaterally (*i.e.* the usual Nash Equilibrium condition) but we further require that any mutually beneficial link be formed in equilibrium. PNE networks are robust to bi-

[10]Notice that for a given g, q and ρ the expected stock of open ideas is determined by (1).

lateral commonly agreed one-link creation and to unilateral deviations. Formally, a pure strategy $s^* = (s_1^*, ..., s_n^*)$ is a Nash Equilibrium of the game of network formation (previously described) if and only if $\Pi_i(g(s^*)) \geq \Pi_i(g(s_i, s_{-i}^*))$, for all $s_i \in S_i$ and $i \in N$. Let $g + ij$ be the network obtained by adding the link g_{ij} to g.

Definition 8 *A network g is a PNE network with respect to the network payoff function Π if and only if there exists a Nash equilibrium strategy profile s^* that supports g, that is, $g = g(s^*)$, and, for all pair of players i and j such that $g_{ij} = 0$ if $\Delta\Pi_i(g + ij) > 0$ then $\Delta\Pi_j(g + ij) < 0$.*

Given that s_{ii}^i is part of s_i $\forall i \in N$, a deviation consisting on a multi-link severance by i and a simultaneous creation of g_{ii} is an unilateral deviation because there is no need of mutual consent to form g_{ii}.

Notice that this is a relatively weak equilibrium concept[11]. Yet, we are able to isolate a single equilibrium network topology for a specific parameter space.

3.3 Results

The empirical study by Goyal, Van der Leij and Moraga (2006) describes a detailed image of the features of real coauthorship networks in Economics. The results of that paper "show that the world of Economics is spanned by inter-linked stars, that this feature is stable over time and that this is the main reason for small average distances". A similar conclusion applied to Biology, Physics and Mathematics can

[11]The concept of Bilateral Equilibrium (also called pairwise stable Equilibrium) introduced in Goyal and Joshi (2003) and used in Van der Leij (2006) is much stricter than PNE since it allows pairs of players to form and delete links simultaneously.

be extracted from Newman (2004). Our aim is to show how the model introduced in the previous section reproduces equilibrium structures with the features of actual scientific collaboration networks. In reality, these networks have a highly unequal distribution of links with a small fraction of researchers concentrating many links. In our model, this feature implies that some nodes receive many open ideas from others and, consequently, they are able to publish more papers than other researchers. Therefore, the payoffs of those nodes that receive a large proportion of open ideas are significantly larger than the payoffs of nodes that receive a small proportion of ideas. Reproducing such an unequal payoff distribution in the equilibrium of a setting in which players can discretionally create or sever links is relevant when the population is fairly homogeneous, *i.e.* when talent is fairly constant across agents. Our results show that for sufficiently low values of ρ (with respect to the processing rate which is one), these networks naturally arise from the interaction among similar agents. Evidently, the more heterogeneous population is, the easier it is to reproduce these unequal distributions of links in equilibrium and, as a consequence, the requirements on ρ are less stringent.

The results' section is divided into two parts. First, we present some preliminary results that determine the basic characteristics of equilibrium networks. We derive a Corollary that highlights how networks with a highly unequal distribution of links can be sustained in a PNE. This result shows that, in spite of having very low *a priori* differences among agents' talents, this kind of topologies can be PNE networks if ρ is sufficiently low. In these equilibrium networks high talent researchers receive an amount of links sufficient to exhaust their processing capability but sufficiently small to not suffer congestion. The second subsection shows that a particular interlinked star network, that holds the previously stated conditions, can naturally arise –from the interaction of self-interested researchers– as the unique PNE structure for any distribution of talent. Heterogeneity and congestion are the key factors that explain these results.

Before presenting the results, we comment the requirements about ρ. This paper shows that equilibrium networks have the form of an interlinked star when ρ is sufficiently small with respect to their processing capability, *i.e.* when agents receive a low number of ideas per period from outside the network in relation to their processing capability. Notice that the higher it is the heterogeneity among agents the less stringent it is the restriction on ρ for results to hold. We emphasize that this is not an unrealistic requirement. In real coauthorship networks, the most connected researchers have around 50 links in Economics (as shown by GVM) and even more in other fields as Biology (Newman, 2004). In our model, this means that each particular collaboration link contributes a small quantity of open ideas to their destination with respect to the processing rate. For this reason, researchers have to maintain the link with many collaborators in order to receive a sufficient amount of ideas to exhaust most of their processing capability.

3.3.1 Preliminary results

In the following three propositions we show that, when ρ is sufficiently low, PNE networks must hold several characteristics for any possible distribution of talents (h). In particular, these results establish an upper-bound in the number of out-degree and in-degree links of any node and a lower bound in the number of in-degree links of nodes with a relatively high level of talent. Therefore, they narrow the set of potential PNE networks. In each of the results we highlight that the higher it is the heterogeneity among agents the less stringent can be the requirements on ρ.

The first result refers to the number of out-degree links of a particular agent. In principle, the researchers of this model can send their open ideas to many collaborators, *i.e.* there is no upper-bound on μ_i $\forall i$. The following result approaches this issue.

Let $G^*_{h,\rho}$ be the set of PNE networks for a given pair (h, ρ).

Proposition 8 *For any functional form of* $f(\cdot)$*, for any pair* (h, ρ_0) *and any* $g \in G^*_{h,\rho_0}$ *in which* $\mu_i > 1$ *for some* $i \in N$*, there always exists a* $\bar{\rho}_1 < \rho_0$ *such that* $g \notin G^*_{h,\rho}$ $\forall \rho < \bar{\rho}_1$.

The proof (see in Appendix) proceeds as follows. We consider all possible cases that a researcher i with $\mu_i \geq 2$ can face. Then, we analyze her incentives to sever the link g_{ij} where $h_j < h_k$ for some $k \in M_i$. We show that, for any vector h whose values are extracted from a continuous distribution function, the marginal payoff derived from such a deviation tends to be positive as $\rho \to 0$. Using the concept of limit, we show how this is a reformulation of the statement of the proposition.

The intuition of the proof is simple. Heterogeneity implies that any agent i can rank the destinations of her open ideas with respect to their level of talent. In other words, i can rank all the agents in M_i with respect to the expected payoff obtained from the processed ideas transmitted to them. If i does not send all open ideas to the destination that maximizes this expected payoff, is because of another important factor: agents have limited processing capability. When an agent severs a link, she automatically increases the flow of open ideas sent to the rest of destinations and, consequently, increases the destinations' queue of ideas when congestion arises. This has a negative impact on the expected utility of the agent who initially deviates. In the proof we show that as $\rho \to 0$ the negative impact tends to vanish. Consequently, for a sufficiently low value of ρ, any agent i with $\mu_i \geq 2$ has incentives to deviate and send all open ideas to the most talented destination (who has a talent $\bar{h}$). The higher it is $\bar{h}$ with respect to the talent of the rest of destinations, the higher would be i's incentives to deviate and, therefore, the higher $\bar{\rho}_1$ can be.

This result illustrates that whatever h, we can always find a sufficiently low value of ρ (say $\bar{\rho}_1$) such that a network in which some agent has two or more out-degree links cannot be sustained in equilibrium for any $\rho < \bar{\rho}_1$. Moreover, increasing the heterogeneity among agents' talent can make the restriction on ρ less stringent. Therefore, an upper-bound on the equilibrium out-degree of nodes arises naturally. As introduced above, this result implies a dramatic simplification of the set of possible PNE networks.

Focusing on the in-degree of nodes we can go one step further in the simplification. The following result establishes an upper-bound for the amount of ideas (and indirectly the amount of links) a node can receive in a PNE network when ρ is sufficiently small.

Proposition 9 *For any functional form of $f(\cdot)$, for any pair (h, ρ_0) and any $g \in G^*_{h,\rho_0}$ in which $\sum_{l \in N_i} \frac{1}{\mu_l} \geq 1 + \frac{q+1}{\rho_0}$ for some $i \in N$, there always exists a $\bar{\rho}_2 < \rho_0$ such that $g \notin G^*_{h,\rho}$ $\forall \rho < \bar{\rho}_2$.*

The intuition of the proof (see in Appendix) is the following: when $\sum_{l \in N_i} \frac{1}{\mu_l} \geq 1 + \frac{q+1}{\rho_0}$ for some $i \in N$, the stock of open ideas of i will be higher than one, even after severing one in-degree link. Given that the maximum processing rate is normalized to one, researcher i can delete one in-degree link without damaging the average processing flow (which will continue to be one). Consequently, agent i can increase her average productivity if she severs an in-degree link coming from a researcher who holds two conditions: (i) her talent is below the average talent of the rest of agents in N_i and (ii) she does not receive a link from i. From Proposition 8 we know that for any $\rho < \bar{\rho}_1$, agent i will send all her open ideas to a unique destination. Then, in order to hold the two conditions above, we only need to have two agents below the average talent of the rest of agents in N_i. Evidently, this is easier when there is a high level of heterogeneity among the agents in N_i and i has a large number of in-degree links. In the proof of this proposition we show that for any arbitrarily homogeneous distribution of talents there always exists a $\bar{\rho}_2$ that assures the existence of such a pair of collaborators for any $\rho < \bar{\rho}_2$ and therefore assures that no researcher i can hold $\sum_{l \in N_i} \frac{1}{\mu_l} \geq 1 + \frac{q+1}{\rho_0}$.

The form of the production function has a direct effect on the incentives of collaboration. For a concave $f(\cdot)$, working with another researcher (rather than alone) increases the probability of publication less than proportionally with respect to the increase in the amount of talent. On the other hand, a convex $f(\cdot)$ implies that adding additional talent in the production process increases the publication probability more than proportionally. For this reason, we can say that a concave $f(\cdot)$ discourages agents to look for collaborators. Previous results are valid for any functional form of

$f(\cdot)$. Now, we focus on the case in which $f(\cdot)$ is linear or convex. In that way we specifically analyze the features of equilibrium networks in the case in which collaborating with other researchers does not imply a loss of productivity with respect to working alone. The results for this case narrow the set of potential PNE networks even more.

Proposition 10 *For a linear or convex $f(\cdot)$, a PNE network cannot have a player i with $\sum_{l \in N_i} \frac{1}{\mu_l} < \frac{q+1}{\rho} - 1$ when some agent j such that $h_j < h_i$ holds: $g_{jk} = 0$ $\forall k$ such that $h_k \geq h_i$.*

See proof in Appendix. When $f(\cdot)$ is linear or convex and without considering the effects of a potential congestion, researchers always prefer to have collaborators with higher talent. If $\sum_{l \in N_i} \frac{1}{\mu_l} < \frac{q+1}{\rho} - 1$ for some $i \in N$, agent i will not suffer congestion even after receiving a new link. For this reason, if all the researchers that receive ideas from another researcher j have a talent lower than h_i, then agent j has incentives to create a new link to agent i. Moreover, since i will not congest after receiving this new link, she will increase her processing rate and, therefore, her payoff will also increase. Thus, such pair of agents cannot exist in a PNE network. Again, heterogeneity and congestion are essential for shaping this result.

Considering this result, and Propositions 8 and 9 we can state the following:

Corollary 1 *For a linear or convex $f(\cdot)$ and for any vector h, highly talented researchers receive all links of the network in any $g \in G^*_{h,\rho}$ for any $\rho < \bar{\rho}_2$. Specifically, $\frac{q+1}{\rho} - 1 \leq \sum_{l \in N_i(g)} \frac{1}{\mu_l} < 1 + \frac{q+1}{\rho}$ for a highly talented researcher i and any $\rho < \bar{\rho}_2$.*

The first claim of the corollary directly derives from Propositions 8 and 10. The bounds of $\sum_{l \in N_i(g)} \frac{1}{\mu_l}$ are obtained in Propositions 9 and 10. This result illustrates the attraction force of highly talented researchers in this model. They will always have a minimum number of collaborators. On the other hand, this attraction force is restricted by the possibility of congestion derived from the agents' limited processing capability. For this reason, there is an upper-bound in the number of collaborators a node can have in equilibrium. We must also remember that the higher it is the heterogeneity among the levels of talent the higher $\bar{\rho}_2$ can be.

Notice that if $\sum_{l\in N_i}\frac{1}{\mu_l}=\frac{q+1}{\rho}$ for some $i\in N$, the stock of open ideas of agent i is one, which is exactly her processing rate. Then, this can be seen as the point in which agents receive the minimum number of ideas in order to process at their maximum rate. At that point agents maximize their processing rate without suffering congestion. The bounds established in Corollary 1 imply that in a PNE network highly talented researchers will have a number of in-degree links that allow them to be close to exhausting their processing capability without suffering congestion.

Notice also that (for a linear or convex $f(\cdot)$) there is no heterogeneity requirement on the distribution of talents to obtain Proposition 10. So, the attraction force of the best researchers exists regardless of the difference between their level of talent and that of the other researchers. That is, the result holds for any vector h extracted from a continuous distribution function.

Since $\frac{q+1}{\rho}-1\leq\sum_{l\in N_i(g)}\frac{1}{\mu_l}<\frac{q+1}{\rho}+1$, note that the lower it is ρ the higher it is $\sum_{l\in N_i(g)}\frac{1}{\mu_l}$. Consequently:

Corollary 2 *For a linear or convex $f(\cdot)$, arbitrarily small differences among agents' talent can generate highly unequal distributions of links in the PNE network candidates. The lower it is ρ, the higher it is the in-degree inequality.*

Moreover, because of the convexity of the relationship between ρ and $\sum_{l \in N_i(g)} \frac{1}{\mu_l}$, very small changes in ρ translate into large increases in the in-degree inequality among researchers. This contrasts with the results of Van der Leij (06) in which a minimum degree of heterogeneity among agents is required in order to reproduce the empirical results about in-degree inequality.

So, we have reproduced an empirical fact. In our Steady State network, some few agents can concentrate many links (especially for low values of ρ). But, does this mean that the network we obtain in equilibrium is an interlinked star as suggested by the empirical results of Goyal, Van der Leij and Moraga (2006)? By the previous results, it is not necessarily the case. For example, we can have a network formed by stars, in which the central agents (who receive a number of links respecting the bounds established in Corollary 1) are not connected between them. The following results focus on showing that for any arbitrarily homogeneous distribution of talents there exist a sufficiently low ρ under which this cannot happen. Yet, in the following subsection we go one step further and characterize a particular interlinked star network as the unique PNE for low values of ρ. Again, the more heterogeneous it is the distribution of talents the less stringent it is the condition on ρ.

3.3.2 Main results

First, we will characterize our PNE network candidate by developing some preliminary steps. Let G^s denote the set of interlinked star networks holding the following three properties:

- $\mu_i = 1 \; \forall i \in N$.

- $g_{ii} = 1$ if and only if $h_i > h_l$ $\forall l \in N$ and $l \neq i$.

- For any given pair (h_i, h_j) such that $h_i > h_j$ and $g_{ik} = g_{jl} = 1$ for any $k, l \in N$, it must be true that $h_k \geq h_l$.

The first condition states that all the nodes of the network have only one out-degree link. The second property implies that, except for the highest talented player, this links to some other agent. Finally, the third condition narrows the set of possible destinations of this out-degree link. In fact, the last two conditions imply that the highest talented agent receives her in-degree links from the nodes located just below her in the ranking of talents; the second player in this ranking receives her in-degree links from the nodes located just below the first group of players in this ranking; and so on. Depending on the distribution of links we can have different networks in G^s. The following result specifies that, when $\rho < \bar{\rho}_2$, there is a single distribution of links (*i.e.* a single network) in G^s holding two stability conditions.

Lemma 1 *For any given pair (h, ρ) such that $\rho < \bar{\rho}_2$, there is a unique network in G^s in which no agent i with $\eta_i > 0$ has incentives either to delete any in-degree link or to create a new in-degree link coming from some j such that $h_j < h_i$. We call this network g^s.*

See the proof in the Appendix. Notice that since $g^s \in G^s$, (i) this graph is an interlinked star network, that is, there is a group of authors linked among them that concentrate certain number of links and (ii) the coauthors of a given researcher do not collaborate with each other. This is, in essence, the basic topology of actual scientific collaboration networks according to the empirical studies commented above. The following two results show that g^s is a natural equilibrium outcome of our model.

Proposition 11 *For a linear or convex* $f(\cdot)$, *for any pair* (h, ρ_0) *and any* $g \in G^*_{h,\rho_0}$ *different from* g^s, *there always exists a* $\hat{\rho}_1 < \rho_0$ *such that* $g \notin G^*_{h,\rho}$ $\forall \rho < \hat{\rho}_1$.

See the proof in the Appendix. This result shows that whatever it is h, we can always find a sufficiently low value for ρ under which a network different from g^s cannot be sustained as a PNE. Therefore, the set of potential PNE networks for a sufficiently small ρ reduces to one single topology. In the proof we show that for any network $g \neq g^s$ we can always find a player (or a pair of players) whose marginal payoff for deviating tends to be positive as $\rho \to 0$. Notice that, using the definition of limit, this implies that such a network g cannot be PNE for any arbitrarily small ρ. Intuitively, the mechanisms underlying this result are the following. We first show that in a network different from g^s and with $\rho < \bar{\rho}_1$, there must exist an agent i who sends all her ideas to a node (say k) such that $h_k < h_j$, where j is the node that should receive ideas from i in g^s. If i does not deviate from this network by creating a new link g_{ij} is because of the possibility that j has a longer queue than k due to congestion. As ρ approaches zero the differences between agents' queue sizes (for example, j and k) become relatively smaller. We reach a point in which player i's incentives to deviate are basically driven by the differences between the talents of j and k. Since $h_k < h_j$, player i would have incentives to add g_{ij}. Evidently, the higher it is h_j relative to h_k the more incentives player i has to deviate. Therefore, the less stringent the condition on ρ would be. Then, a higher heterogeneity of talents can increase agent i's incentives to deviate.

However, for the deviation to take place, node j should have incentives to form the new link g_{ij}. In the worse case, agent j would only accept such a link if the average productivity from the ideas of her queue increases after the deviation. For a sufficiently low ρ, we can state that as $\rho \to 0$, this average productivity decreases. Then, there is a point in which this average is sufficiently low that the formation of g_{ij} pushes this average productivity up. Again, it is easy to show that a higher heterogeneity among the levels of talent would increase agent j's incentives to deviate.

After Proposition 11, the set of PNE candidates reduces to one single network for a sufficiently low ρ. The following result confirms that g^s is in fact a PNE for any arbitrarily homogeneous distribution of talents whenever ρ is sufficiently small.

Proposition 12 *For a linear or convex $f(\cdot)$ and for any pair (h, ρ_0), if $g^s \notin G^*_{h,\rho_0}$ there always exists a $\hat{\rho}_2 < \rho_0$ such that $g^s \in G^*_{h,\rho}$, $\forall \rho < \hat{\rho}_2$.*

See the proof in the Appendix. In the proof we review all the possible deviations from g^s. We find that, in the worse cases, the marginal payoff of potential deviators tends to be negative as $\rho \to 0$. Therefore for any vector h, we can always find a sufficiently low ρ such that no player has incentives to deviate from g^s. From an intuitive point of view, the proof can be explained as follows. There are two kinds of deviations from g^s. On the one hand, agent i can create an additional link to an agent with a lower talent than that of the previous destination or she can substitute her current out-degree link by g_{ii}. In both cases, this agent trades-off the potential benefits from avoiding or reducing the effects of congestion against the costs of reducing the productivity of the processed ideas due to the lower talent of the new destination. In the proof we show that the positive part of this trade-off tends to vanish as $\rho \to 0$, and in consequence, the marginal payoff for deviating tends to be negative. Again, increasing heterogeneity among agents would increase the differences in talent and thus, increase the negative part of the commented trade-off. As a consequence, $\hat{\rho}_2$ can be higher as we increase the differences among agents' talent. The other type of deviation consists of creating an additional link to an agent with a talent higher than that of the previous destination. In that case, we use Lemma 1 to conclude that the marginal payoff of the new destination to accept the link is negative.

The following corollary immediately emerges from the last two propositions. Let $\hat{\rho} \equiv min(\hat{\rho}_1, \hat{\rho}_2)$.

Corollary 3 *For a linear or convex $f(\cdot)$ and for any vector h, there exists a $\hat{\rho}$ such that g^s is the unique PNE network for any $\rho < \hat{\rho}$.*

Existence of g^s as a PNE comes from Proposition 12 and uniqueness comes from Proposition 11.

Notice again, that we do not need to impose any degree of heterogeneity among agents. Specifically any talent vector h extracted from a continuous distribution function can generate the previous result. Nevertheless, the higher it is the heterogeneity among agents the less stringent it can be the condition on the parameter ρ in order to get g^s as the unique PNE network. For these reasons we conclude that the kind of networks GVM observes in reality are a natural outcome from the interaction of self-interested researchers of this model.

Before ending this section of results, it is interesting to provide hints about the behavior of the model for a concave $f(\cdot)$. As commented above, concavity of $f(\cdot)$ discourages agents at the moment of looking for collaborations. Consequently, we cannot assure the stability of the interlinked star for all distributions of talent even for an arbitrarily small ρ. In particular, we would need a sufficiently high inequality across levels of talent to support such a network of collaborations. This inequality is also needed to support other structures with high in-degree inequality (non-interlinked stars) as stable networks.

Remark 6 *For $f(\cdot)$ concave, networks with high in-degree inequality (as an interlinked star) may not be PNE even for arbitrarily small ρ. To assure stability of this kind of networks we need a minimum degree of heterogeneity among researchers' talents.*

Other networks such as the empty network or the cycles can exist in equilibrium even with low values of ρ, especially when inequality across levels of talent is not so high. Summarizing, for a concave $f(\cdot)$ the key factor affecting the shape of the stable network is the inequality across levels of talent. Only highly unequal talent distributions will allow obtaining stable networks with a high in-degree inequality such as the interlinked star.

3.4 Discussion

3.4.1 Empirical patterns

Based on empirical patterns, Goyal, Van der Leij and Moraga (2006) reach the conclusion that the field of economic research is spanned by interlinked stars. In this paper we showed that a simple network formation model characterized by the limited processing capability of heterogeneous agents can reproduce the characteristics of the in-degree distribution of the so called interlinked star network in equilibrium. In this section, we discuss how our model may be extended to explain other empirical patterns.

One of the first empirical findings by GVM relates to the average number of collaborators. The giant component of the analyzed coauthorship network reveals an average of 2.48 in the 1970's and 3.06 in the 1990's[12]. None of the results of our model excludes the possibility of having these average numbers of collaborators. In fact, we can have equilibrium networks with 2, 3 or more collaborators per researcher. But, as Proposition 8 shows for low values of ρ, it would be especially difficult to have a researcher $i \in N$ with more than one out-degree link in a PNE network. Therefore, for low values of ρ our model can hardly reproduce this average number of collaborators.

A natural extension of the model will allow us to reproduce this empirical fact. Imagine a model in which there exist different types of talent and researchers are specialists, so they have a specific type. Moreover, ideas can require some specific type of talent to be published that does not necessarily coincide with the type of talent of the first receiver. For this reason, we can also classify the ideas on different types. To be specific let h_i^x denote the agent i's amount of talent of type x. Let $f_x(h_i^x + h_j^x)$ be the expected probability of publishing a processed idea of type x by researchers i and j. The rest of the model would not change with respect to the model described in section 3.2. In this new model, agents would have incentives to select specialist collaborators for each of the different types of ideas they can receive. Thus, equilibrium networks would be able to reproduce higher average numbers of collaborators. Moreover the

[12] Newman (2004) founds that the average number of collaborators in Biology, Physics and Mathematics was 18.1, 9.7 and 3.9 respectively.

average number of collaborators in equilibrium would positively depend on the degree of researchers' specialization. Therefore the model would be a formalization of the argument defending that one of the key factors explaining the increase in the flow of scientific collaboration is the increase in the specialization of researchers. This increase in the number of collaborators is a trend that GVM detected for the last 30 years in the field of economic research.

With respect to the degree distribution, GVM finds that such a distribution exhibits fat-tails, with a small fraction of scientists having a large number of collaborators. The same can be concluded for the fields of Biology, Physics and Mathematics, as Newman (2004) shows. Our model shows that the links concentrate in highly talented researchers. Decreasing the value of ρ increases the number of links directed towards each of these researchers and, in consequence, increases the inequality in the in-degree distribution. In equilibrium our model predicts that highly talented researchers have roughly the same number of links (see Corollary 1). Evidently, this is not the case in actual networks. This result arises because we assume that all players in our game have exactly the same processing capability, which we have normalized to 1. By allowing different processing rates, the model is able to reproduce equilibria with different in-degree levels for different players.

The last empirical pattern we discuss refers to clustering. GVM shows that *"the most connected individuals collaborated extensively and most of their coauthors did not collaborate with each other"*. Our model cannot provide an intuitive explanation to GVM's results. By focusing on the role of heterogeneity among players and their limited processing capability (as our model does), we can only provide an intuitive argument for explaining the number of links per node which is not a sufficient condition for explaining their level of clustering. In order to explain clustering, we would have to consider geographic or conceptual proximity among researchers. A simple extension of the model can capture this consideration: let us assume that researchers are distributed in broadly-defined groups. Two researchers can be members of the same group if they are in the same academic department, if they work on similar topics or if they share a common personal characteristic. If two researchers i and j are in the same group then $d_{ij} = 1$; otherwise $d_{ij} = 0$. We can reasonably argue that the collaboration between two members of the same group will be more productive than the collaboration between two more distant researchers. Formally, we can write the expected probability of publishing a processed idea by nodes i and j as $f(h_i + h_j + kd_{ij})$ for some $k > 0$. This simple extension allows the model to reproduce equilibrium networks with high clustering among members of the same group and low clustering among highly talented researchers of different groups. In Chapter 2, we presented a network formation model that takes into account these considerations[13].

3.4.2 Stability and efficiency

A network is efficient when it maximizes the aggregate payoff. The set of efficient networks usually does not coincide with the set of stable networks. In fact, one of the most usual analyses in the network formation models is the comparison between stable and efficient networks. Jackson (2004) collects a variety of examples. Jackson and Wolinsky (1996) develops a simple coauthorship network formation model as an

[13]In Chapter 2, agents are exogenously distributed in groups. However, we adopted a different payoff function than the one used here.

example to show how negative externalities can play an important role in the network formation process in the academic world, and in particular, in the conflict between stability and efficiency. In the efficient network that the model yields, researchers are distributed in pairs, *i.e.* two agents connected with each other and isolated from the rest. However, the stable network is over connected with respect to the efficient network. In that model, the advantageous strategy from an individualistic point of view does not coincide with the good strategy from an aggregate point of view. This phenomenon, is especially relevant in coauthorship networks because it implies that researchers' individual incentives damage aggregate scientific production. In this sense, Jackson and Wolinsky (1996) offers a pessimistic view of actual scientific collaboration networks. In the following lines we show that in our model, individual and aggregate incentives are more aligned than in Jackson's and Wolinsky's (1996) model. Yet, they do not fully coincide.

The results of Section 3.3 suggests that g^s is a stable network for a sufficiently small ρ when $f(\cdot)$ is linear or convex. In this section we derive the efficient network(s) for this specific case. This allows us to compare stable and efficient networks. *A priori*, we can say that g^s has favorable characteristics to maximize the aggregate payoff. In particular, for a linear or convex $f(\cdot)$, it seems appropriate that highly talented researchers collaborate with each other. But, is g^s the best structure of collaborations in order to maximize the aggregate payoff for any given pair (h, ρ)? The following result gives an important insight to answer this question.

Proposition 13 *For $g = g^s$ and for any vector h, there exists a ρ^* such that for any $\rho < \rho^*$, if we substitute a link g_{ij} by a new link g_{ik} then the marginal aggregate payoff* *decreases when* $h_k < h_j$ *and increases when* $h_k > h_j$.

See the proof in the Appendix. Independently of h, the accumulation of links to highly talented agents has positive implications for the aggregate payoff when ρ is sufficiently small. Evidently, the higher it is the difference in talent between highly talented researchers and the rest of researchers, the higher the aggregate incentives to accumulate links to them, and the less stringent is the requirement on ρ to hold the last proposition.

Once again, the trade off between the benefits of working with highly talented researchers and the costs of working with more congested coauthors come on stage. By changing g_{ij} by a new link g_{ik} such that $h_k < h_j$, the ideas of agent i can avoid congestion problems but they will have a lower probability of publication once processed. On the other hand, by changing g_{ij} by a new link g_{ik} such that $h_k > h_j$, k can suffer congestion problems but i increases the probability of publication of her processed ideas. But when $\rho \to 0$, congestion tends to disappear. Consequently, the aggregate marginal payoff increases if highly talented players receive more ideas and decreases when highly talented players receive less ideas.

The result shows that the PNE interlinked star network we obtained in the previous section is not efficient for any pair (h, ρ). It also shows that the way of increasing efficiency when the entrance rate of open ideas is sufficiently low and the differences among the levels of talent are sufficiently large, is accumulating more links to highly talented researchers than what would be individually desirable.

3.5 Conclusion

In spite of the large body of empirical research about scientific collaboration networks, there is a lack of foundational theoretical models that analyze how individual decisions contribute to scientific collaboration network formation. This paper proposes a dynamic model to analyze the formation of this kind of networks.

We focus on heterogeneity among agents' productivity and congestion derived from agent's limited processing capability. We show that self-interested researchers in this setting organize themselves in inter-linked stars, as it is suggested by empirical evidence.

Bibliography

[1] G. Ahuja. Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, 45(3):425–455, 2000.

[2] R. Albert and A.L. Barabási. Statistical mechanics of complex networks. *American Physical Society*, 74(1):47–97, 2002.

[3] A. Arenas, A. Cabrales, L. Danon, A. Díaz-Guilera, R. Guimerà, and F. Vega-Redondo. Optimal information transmission in organizations: Search and congestion. Working paper UPF 698, 2003.

[4] V. Bala and S. Goyal. A non-cooperative model of network formation. *Econometrica*, 68:1181–1230, 2000.

[5] P. Ball. *Critical Mass. How One Thing Leads to Another*. Arrow Books, 2004.

[6] C. Ballester, A. Calvó-Armengol, and Y. Zenou. Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.

[7] P. Battigalli, M. Gilli, and M.C. Molinari. Learning and convergence to equilibrium in repeated strategic interactions: An introductory survey. *Ricerche Economiche*, 46:335–378, 1992.

[8] J. B. Berck. Statistical discrimination in a labor market with job selection. working paper University of California, Berkeley, 2001.

[9] F. Bloch and M.O. Jackson. A noncooperative model of network formation. *Econometrica*, 68:1181–1230, 2000.

[10] J. K. Brueckner, N. J. Dyer, and P. T. Spiller. Fare determination in airline hub-and-spoke networks. *The RAND Journal of Economics*, 23(3):309–333, 1992.

[11] R. S. Burt. *Structural Holes: The Social Structure of Competition*. Harvard Universtity Press, 1992.

[12] R. S. Burt. Structural holes and good ideas. *American Journal of Sociology*, 110(2):349–99, 2004.

[13] A. Calvó-Armengol. Job contact networks. *Journal of Economic Theory*, 115:191–206, 2004.

[14] A. Calvó-Armengol and R. Ilkiliç. Pairwise-stability and nash equilibria in network formation. *mimeo*, 2004.

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

[16] A. Calvó-Armengol and Y. Zenou. Social networks and crime decisions: The role of social structure in facilitating delinquent behavior. *International Economic Journal*, 45(3):939–958, 2004.

[17] G. Charness, M. Corominas-Bosch, and G.R. Frechette. Bargaining and network structure: An experiment. *Journal of Economic Theory*, 136(1):28–65, 2007.

[18] S. Coate and G.C. Loury. Will affirmative-action policies eliminate negative stereotypes? *American Economic Review*, 83(5):1220–40, 1993.

[19] Colin Cooper and Alan Frieze. A general model of web graphs. *Random Structures and Algorithms*, 22(3):311 – 335, 2003.

[20] E. Durkheim. *Revue philosophique*, December 1879.

[21] A. Dutta, A. Van Den Nouweland, and S. Tijs. Link formation in cooperative situations. *International Journal of Game Theory*, 27:245–256, 1998.

[22] N. Friedkin. A test of the structural features of granovetter's 'strength of weak ties' theory. *Social Network*, 2:411–422, 1980.

[23] R. G. Fryer and M. O. Jackson. Categorical cognition: A psychological model of categories and identification in decision making. mimeo: Harvard University and Stanford University, 2003.

[24] A. Galleotti, S. Goyal, M.O. Jackson, F. Vega-Redondo, and L. Yariv. Network games. mimeo, University of Essex and California Institute of Technology, 2006.

[25] L. Garicano. Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904, 2000.

[26] M. Gilli. On non-nash equilibria. *Games and Economic Behavior*, 27:184–203, 1999.

[27] S. Goyal, M. Van der Leij, and J. L. Moraga-González. Economics: An emerging small world. *Journal of Political Economy*, 114:403–412, 2006.

[28] S. Goyal and S. Joshi. Bilateralism and free trade. mimeo: Tinbergen Econometric Institute report 176, 1999.

[29] S. Goyal and S. Joshi. Unequal connections. *International Journal of Game Theory*, 34:319–349, 2006.

[30] S. Goyal and F. Vega-Redondo. Structural holes in social networks. *forthcoming Journal of Economic Theory*, 2007.

[31] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78:1360–1380, 1973.

[32] M. Granovetter. Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3):481–510, 1985.

[33] R. Harrison and R. Munoz. Stability and equilibrium selection in a link formation game. mimeo: Georgetown University and University of Maryland, 2002.

[34] John Hudson. Trends in multi-authored papers in economics. *Journal of Economic Perspectives*, 10(3):153–58, 1996.

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

[36] M.O. Jackson. A survey of models of network formation: Stability and efficiency. In Gabrielle Demange and Myrna Wooders, editors, *Group Formation in Economics; Networks, Clubs and Coalitions*, chapter 1. Cambridge University Press, 2004.

[37] M.O. Jackson and B.W. Rogers. Meeting strangers and friends of friends: How random are social networks? *forthcoming American Economic Review*, 2006.

[38] C. Johnson and R. P. Gilles. Spatial social networks. *Review of Economic Design*, 5(3):273–299, 2000.

[39] R.E. Kranton and D.F. Minehart. A theory of buyer-seller networks. *American Economic Review*, 91(3):485–508, 2001.

[40] David N. Laband and Robert D. Tollison. Intelectual collaboration. *Journal of Political Economy*, 108(3):632–662, 2000.

[41] M. McBride. Position-specific information in social networks: Are you connected? mimeo: University of California, 2004.

[42] M. McBride. Conjectural pairwise stability: Why observation matters in mutual consent networks. mimeo: University of California, 2005.

[43] M. McBride. Imperfect monitoring in communication networks. *Journal of Economic Theory*, 126(1):97–119, 2006.

[44] A. Mehra, M. Kilduff, and D. Bass. The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly*, 26:121–146, 2003.

[45] S. Milgram. The small world problem. *Psychology Today*, 2:60–67, 1967.

[46] S. Mullainathan. Thinking through categories. *mimeo: Harvard University*, 2002.

[47] R. Myerson. *Game Theory: Analysis of Conflict*. Math. Operations Research, 1991.

[48] M. E. J. Newman. Scientific collaboration networks I: Network construction and fundamental results. *American Physical Society*, 64(1), 2001.

[49] M. E. J. Newman. Scientific collaboration networks II: Shortest paths, weighted networks, and centrality. *American Physical Society*, 64(1), 2001.

[50] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[51] M. E. J. Newman. Coauthorship networks and patterns of scientific collaboration. *PNAS*, 101(suppl. 1):5200–5205, 2004.

[52] E.S. Phelps. The statistical theory of racism and sexism. *American Economic Review*, 62(4):659–661, 1972.

[53] J. Podolny and J. Baron. Resources and relationships, social networks and mobility in the work place. *American Sociological Review*, 62:673–693, 1997.

[54] Derek J. de Solla Price. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27:292–306, 1976.

[55] C. Z. Qin. Endogenous formation of cooperation structures. *Journal of Economic Theory*, 69:218–226, 1996.

[56] A. Rubinstein and A. Wolinsky. Rationalizable conjectural equilibrium: Between nash and rationalizability. *Games and Economic Behavior*, 6:299–311, 1994.

[57] M. Sattinger. Statistical discrimination with employement criteria. *International Economic Review*, 39(1):205–237, 1998.

[58] M. Van der Leij. *The Economics of Networks: Theory and Empirics*. PhD thesis, Tinbergen Institute. Erasmus University, 2006.

[59] D. Watts. *Small Worlds*. Princeton University Press, 1999.

Appendix A

Chapter 1

A.1 Figures

Figure A.1:

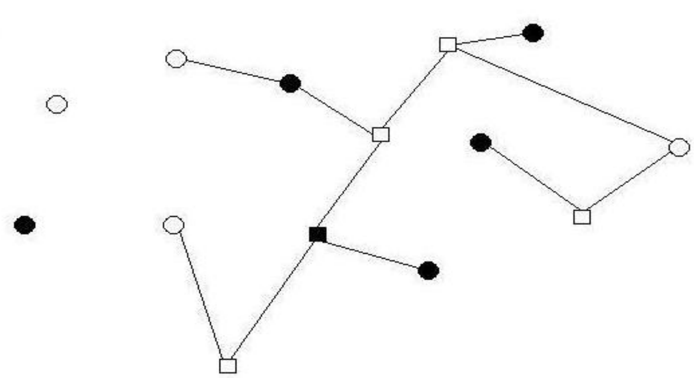


Figure A.2:

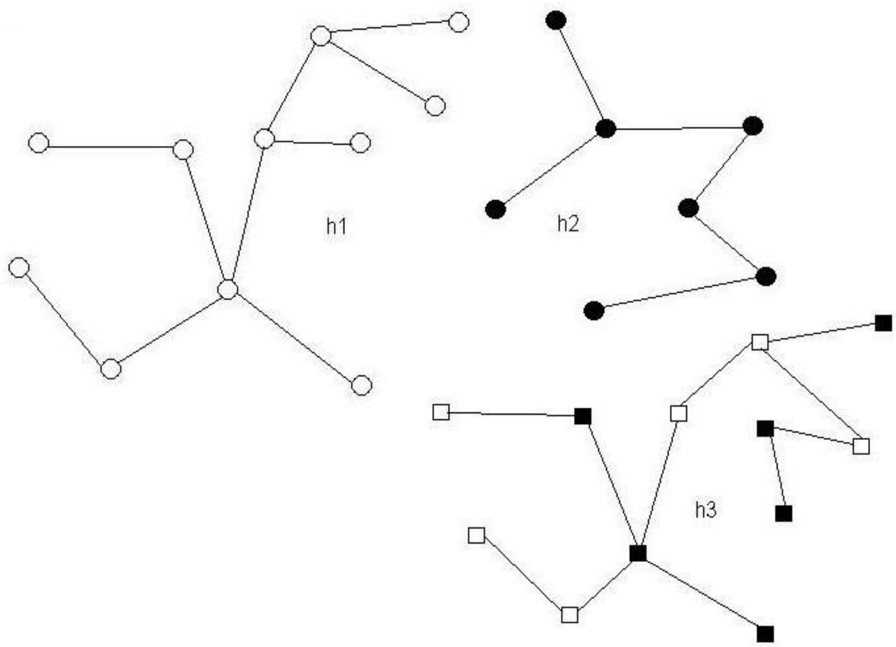


Figure A.3:

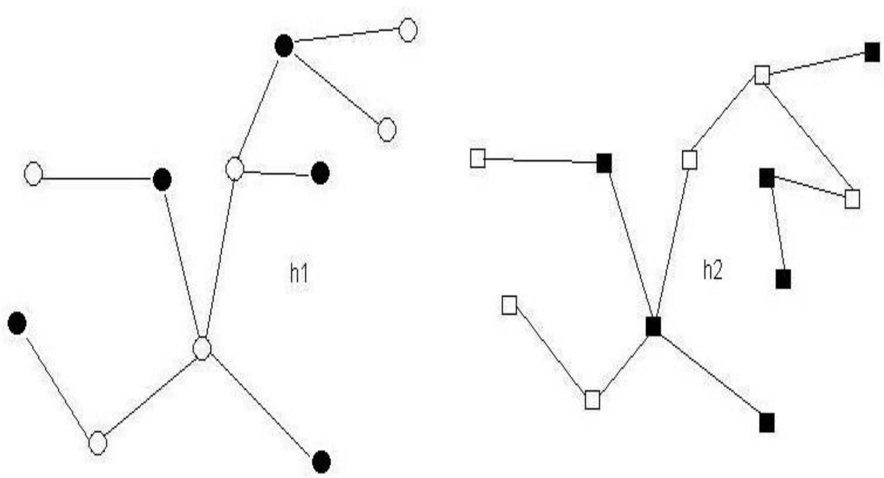


Figure A.4:

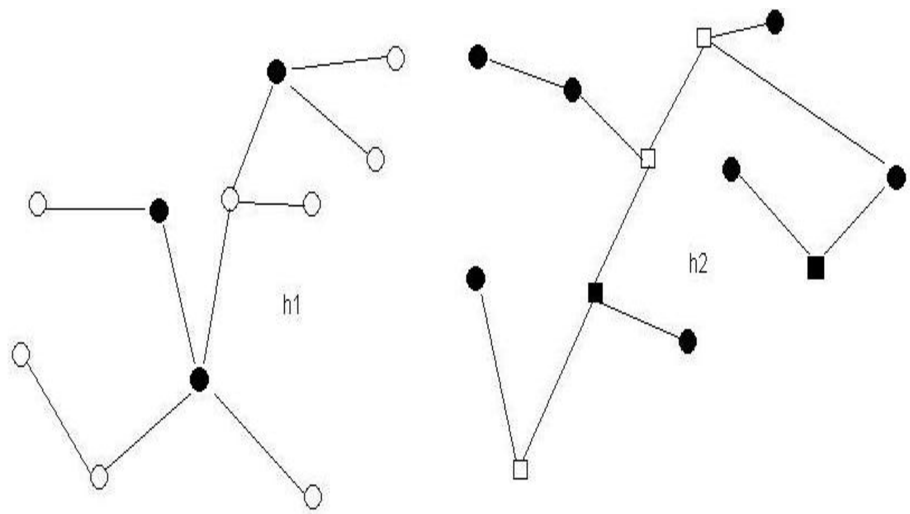


Figure A.5:

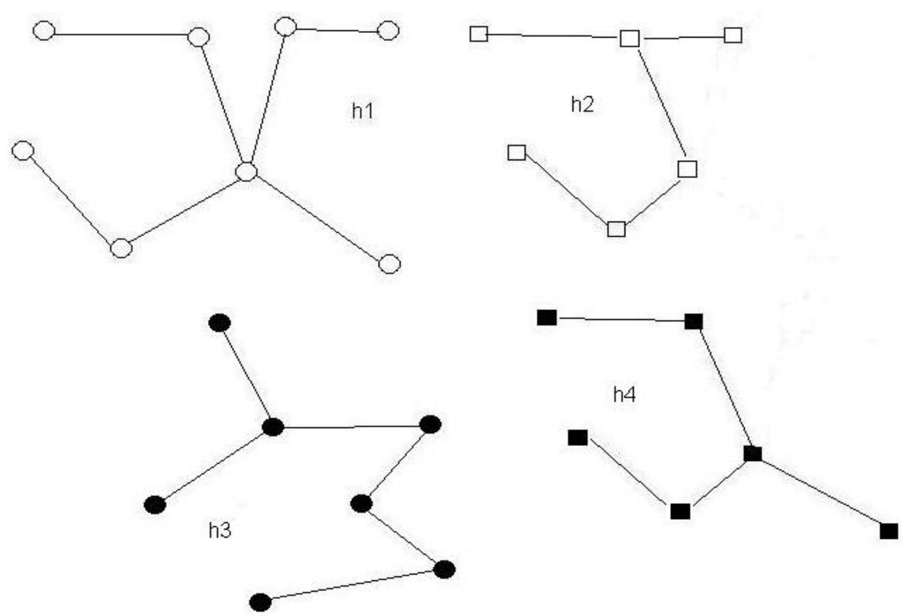


Figure A.6:

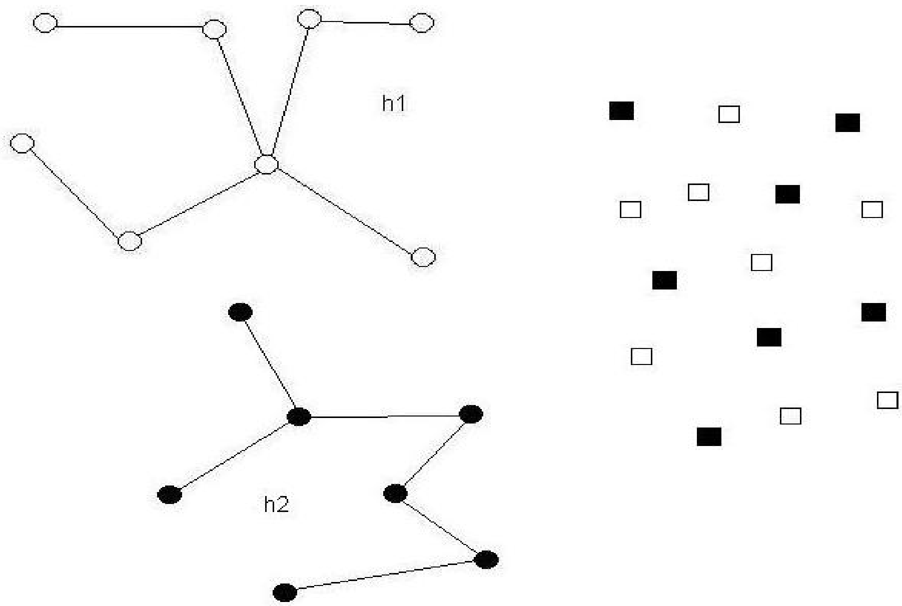


Figure A.7:

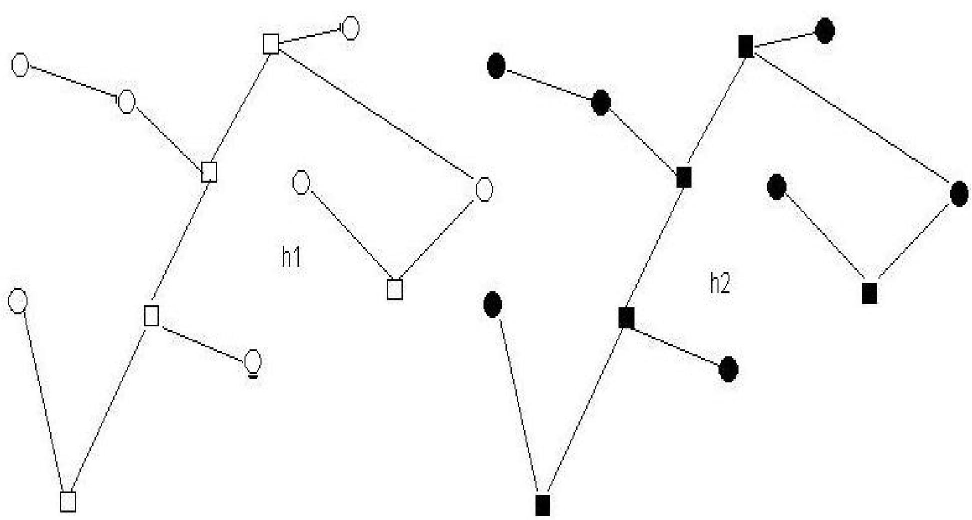


Figure A.8:

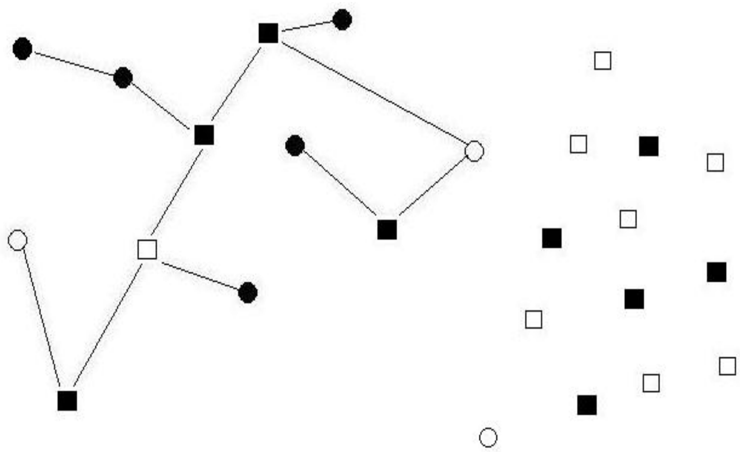
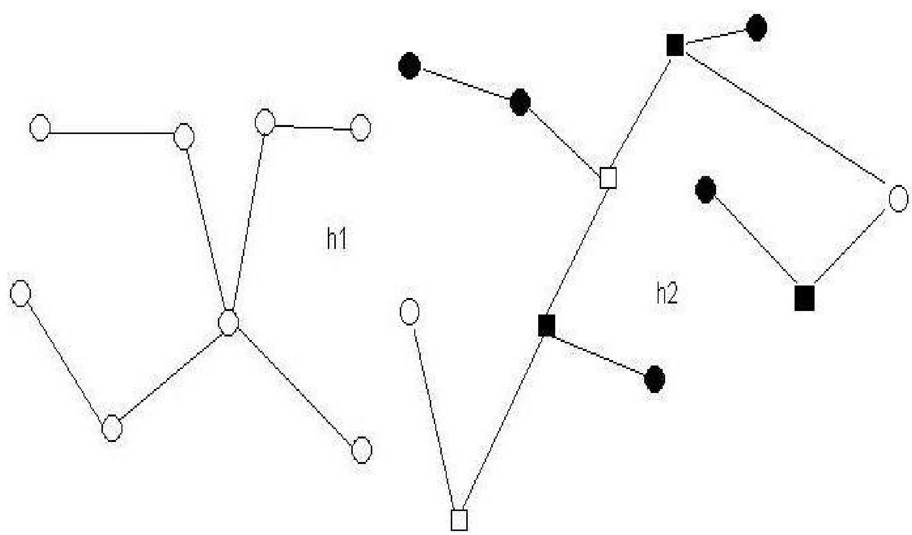


Figure A.9:



A.2 Proofs

Proof of Proposition 1. Notice that $PBNE(\gamma, P, t) \subseteq BPS(\gamma, P, t)$. Hence, if $BPS(\gamma, P, t) = \emptyset$, the result follows. Suppose that $BPS(\gamma, P, t) \neq \emptyset$ and let $g^* \in BPS(\gamma, P, t)$. Define

$$\phi(g^*, \gamma, P, t) \in \min\{\sum_{w \in W} P(w/I_i)[Y_i(v, g^*, t) - Y_i(v, g^* - ij, t)] : ij \in g^*, i \in \mathcal{N}\}$$

Bayes Pairwise Stability implies that $\phi(g^*, \gamma, P, t) \geq 0$. Suppose that f^γ is α-Bayes convex in own current links on $BPS(\gamma, P, t)$ for some $\alpha \geq 0$. Then, inequality (1) implies that $\sum_{t_{-I_i} \in T_{-I_i}} P(w/I_i)[Y_i(v, g, t) - Y_i(v, g - ij_1 - ... - ij_l, t)] \geq \alpha l \phi(g^*, \gamma, P, t) \geq 0$, for all $ij_1, ..., ij_l \in g^*$ and $i \in \mathcal{N}$. Moreover, by definition of Bayes Pairwise Stability if

$$\sum_{w \in W} P(w/I_i)\ Y_i(v, g, t) < \sum_{w \in W} P(w/I_i)\ Y_i(v, g + ij, t)$$

then

$$\sum_{w \in W} P(w/I_i)\ Y_j(v, g, t) > \sum_{w \in W} P(w/I_i)\ Y_j(v, g + ij, t)$$

for all $ij \notin g^*$ and $i \in \mathcal{N}$. Therefore, $g^* \in PBNE(\gamma, P, t)$.

On the other hand, suppose that there exists $g^* \in BPS(\gamma, P, t)$, such that for all $\alpha \geq 0$, inequality (1) does not hold for g^*. Then, for some $i \in \mathcal{N}$, there exists $ij_1, ..., ij_l \in g^*$ such that $\sum_{w \in W} P(w/I_i)[Y_i(v, g, t) - Y_i(v, g - ij_1 - ... - ij_l, t)] < 0$, implying that $g^* \notin PBNE(\gamma, P, t)$. ■

Appendix B

Chapter 2

B.1 Proofs

Proof of Proposition 5. This is proved by contradiction. Specifically, we will show that there cannot be a PNE network with two or more components in which one of them has two or more neighborhoods. Since PNE is weaker than BE, such a network cannot be sustained in a BE either.

First, we introduce two preliminary lemmas. Their proof is omitted here because they are immediate applications of two analogous lemmas set forth by Goyal and Vega (2006). Lemma 2 refers to the marginal payoff of critical links. These are links that define the unique path between the two players involved and whose deletion increase the number of components. By Proposition 4, critical links can only connect players from different neighborhoods in a PNE network; therefore critical links are not just the unique path between two players, they are also the unique path between neighborhoods.

Lemma 2 *Consider any network g. If $g_{ij} = 1$ and the link is critical, then the marginal payoff of the link g_{ij} for both players (i and j) is exactly the same.*

Lemma 3 *In a network g, any component has at least two non-essential neighborhoods.*

By contradiction, let assume that g is a PNE network with at least one inter-neighborhood link and suppose that there is more than one component. Let $\hat{C}$ be the largest component in g, which must therefore contain at least two neighborhoods. We claim that there is a neighborhood $M_j \notin \hat{C}$ and a player $j \in M_j$ that can establish a mutually profitable link to some player in $\hat{C}$, contradicting the initial assumption of stability. For simplicity, we will assume the less favorable case where M_j has no external connections, i.e. M_j is an isolated neighborhood.

By lemma 3, $\hat{C}$ has some non-essential neighborhood. Then there are two possibilities:

- First, one of these non-essential neighborhoods M_i is extreme (i.e. M_i has links to only one another neighborhood, say M_k). In essence, three basic cases can be distinguished at this point:

(i) there is a single critical link between M_i and M_k, say between agents $i \in M_i$ and $k \in M_k$.

(ii) there is a non critical link between M_i and M_k but there is only one player $k \in M_k$ who has links to members of M_i (say $g_{kl} = g_{ki} = 1$, where $l, i \in M_i$).

(iii) there is a non critical link between M_i and M_k with at least two agents in M_k linked to members of M_i.

Consider case (i). Since keeping the link between i and k is profitable for both, some player in M_j, say j, would find optimal to create a link to k if given the opportunity, and so would player k. This contradicts the initial statement of g as a PNE network.

Consider case (ii). Let $\mathcal{N}_i^r(g) = \{j \in C_i : e(i,j) = r\}$ be the players whom i accesses via r essential players and let $\eta_i^r(g) = |\mathcal{N}_i^r(g)|$.

Since g is a PNE network, it follows that k's marginal payoff of cutting one of her links to M_i off must be negative. That is:

$$\begin{aligned}\Delta\Pi_K &= c^h + (m-1)(\frac{1}{3} - \frac{1}{2}) + (m-1)(\eta_k^0(g) - m)(\frac{1}{4} - \frac{1}{3}) + \\ &\quad (m-1)\eta_k^1(g)(\frac{1}{5} - \frac{1}{4}) + ... + (m-1)\eta_k^R(g)(\frac{1}{R+4} - \frac{1}{R+3}) \\ &< 0\end{aligned}$$

where R is a finite positive number. This implies that:

$$c^h < \frac{1}{6}(m-1) - (m-1)(\eta_k^0(g) - m)(\frac{1}{4} - \frac{1}{3}) - (m-1)\sum_{i=1}^{R}\eta_k^i(g)(\frac{1}{i+4} - \frac{1}{i+3}) \quad \text{(i)}$$

On the other hand, player j has the following marginal payoff from forming a link to k:

$$\begin{aligned}\Delta\Pi_j &= \frac{1}{2} + \eta_k^0(g)\frac{1}{3} + \eta_k^1(g)\frac{1}{4} + ... + \eta_k^R(g)\frac{1}{R+3} + \\ &\quad (m-1)\frac{1}{3} + (m-1)\eta_k^0(g)\frac{1}{4} + (m-1)\eta_k^1(g)\frac{1}{5} + ... + (m-1)\eta_k^R(g)\frac{1}{R+4} - c^h\end{aligned}$$

Using inequality (i) we write:

$$\begin{aligned}
\Delta\Pi_j &> \frac{1}{2} + (\eta_k^0(g) + m - 1)\frac{1}{3} + \sum_{i=1}^{R} \eta_k^i(g)\frac{1}{i+3} + (m-1)\sum_{i=0}^{R}\eta_k^i(g)\frac{1}{i+4} - \\
&\quad \frac{1}{6}(m-1) + (m-1)(\eta_k^0(g) - m)(\frac{1}{4} - \frac{1}{3}) + (m-1)\sum_{i=1}^{R}\eta_k^i(g)(\frac{1}{i+4} - \frac{1}{i+3}) \\
&> (m-1)(\eta_k^0(g) - m)(\frac{1}{2} - \frac{1}{3}) + (m-1)\sum_{i=1}^{R}\eta_k^i(g)(\frac{2}{i+4} - \frac{1}{i+3})
\end{aligned}$$

Since $\eta_k^0(g) \geq m$ and $(\frac{2}{i+4} - \frac{1}{i+3}) > 0 \; \forall \; i \geq 0$, we conclude that $\Delta\Pi_j > 0$. By Lemma 2, agent k's marginal payoff from forming an additional link to j is also positive. Thus, the deviation is profitable for both players and this contradicts the statement of stability of g.

Consider case (iii). Imagine the less favorable case where an agent $k \in M_k$ has a link to $i \in M_i$ and there is another link between these two neighborhoods involving two other players. Agent k's marginal payoff from severing the link between i and k is:

$$\begin{aligned}
\Delta\Pi_k &= c^h + (m-1)(\frac{1}{4} - \frac{1}{2}) + (\frac{1}{3} - \frac{1}{2}) = \\
&\quad c^h - (m-1)\frac{1}{4} - \frac{1}{6}
\end{aligned}$$

Since g is a PNE network, $c^h < \frac{m}{4} - \frac{1}{12}$. Considering this condition, player j's marginal payoff for forming a link to a member of M_k holds:

$$\begin{aligned}
\Delta\Pi_j &> \frac{1}{2} + (2m-1)\frac{1}{3} + (m-1)(2m-1)\frac{1}{4} - c^h \\
&> \frac{1}{2} + 2(m-1)\frac{1}{3} + (m-1)^2\frac{1}{4} - c^h \\
&> \frac{1}{2} + 2(m-1)\frac{1}{3} + (m-1)^2\frac{1}{4} - \frac{m}{4} + \frac{1}{12} \\
&= \frac{1}{12}(3m^2 - m + 2)
\end{aligned}$$

From this point on, it is easy to check that $\Delta\Pi_j > 0$, $\forall m > 0$. By Lemma 2, player k also gets a positive payoff from this deviation. Thus, we reach again a contradiction with the statement of stability of g.[1]

- Second, there are no extreme neighborhoods. Let M_i be a non-essential neighborhood. Two subcases have to be considered:

(i) In M_i there is a single player (say i) with external links. By assumption, player i's number of external links is at least two. Player i's payoff can be written as:

$$\begin{aligned}\Pi_i &= \frac{\eta_i^0(g)}{2} + ... + \frac{\eta_i^r(g)}{r+2} + \\ &(m-1)\frac{(\eta_i^0(g) - (m-1))}{3} + (m-1)\frac{\eta_i^1(g)}{4} + ... + (m-1)\frac{\eta_i^r(g)}{r+3} - \eta_i(g)c^h\end{aligned}$$

Since g is a PNE network, it follows that:

$$\frac{1}{\eta_i(g)}[\frac{\eta_i^0(g)}{2} + ... + \frac{\eta_i^r(g)}{r+2} + (m-1)\frac{(\eta_i^0(g) - (m-1))}{3} + (m-1)\frac{\eta_i^1(g)}{4} + ... + (m-1)\frac{\eta_i^r(g)}{r+3}] \geq c^h \quad \text{(ii)}$$

Agent j's marginal payoff for forming a link between i and j is:

$$\begin{aligned}\Delta\Pi_j &= \frac{1}{2} + \frac{\eta_i^0(g)}{3} + ... + \frac{\eta_i^r(g)}{r+3} + \frac{m-1}{3} + \\ &(m-1)\frac{\eta_i^0(g)}{4} + ... + (m-1)\frac{\eta_i^r(g)}{r+4} - c^h \\ &> \frac{\eta_i^0(g)}{4} + ... + \frac{\eta_i^r(g)}{2r+4} + \\ &(m-1)\frac{\eta_i^0(g)}{6} + ... + (m-1)\frac{\eta_i^r(g)}{2r+6} - c^h \\ &\geq \frac{1}{\eta_i(g)}[\frac{\eta_i^0(g)}{2} + ... + \frac{\eta_i^r(g)}{r+2} + (m-1)\frac{\eta_i^0(g)}{3} + ... + (m-1)\frac{\eta_i^r(g)}{r+3}] - c^h \geq 0\end{aligned}$$

[1]Notice that if $m = 1$ only case (i) is possible.

The first inequality is immediate, while we use $\eta_i(g) \geq 2$ for deriving the second inequality. The third inequality follows from condition (ii). Again, by applying Lemma 2, both i and j have incentives to form a link between them contradicting the assumed stability of the initial network.

(ii) In M_i there is more than one player with external links. Let $i \in M_i$ be a player with at least one external link. Such a player must exist since M_i is part of a component with at least two neighborhoods. Player i's payoff holds:

$$\Pi_i \geq \frac{1}{2}\eta_i^0(g) + \frac{1}{3}\eta_i^1(g) + ... + \frac{1}{r+2}\eta_i^r(g) - c^h \geq 0 \qquad \text{(iii)}$$

Consider the deviation consisting on forming a link between i and j. Agent j's marginal payoff holds:

$$\begin{aligned}
\Delta\Pi_j &= \frac{1}{2} + \frac{1}{3}\eta_i^0(g) + \frac{1}{4}\eta_i^1(g) + ... + \frac{1}{r+3}\eta_i^r(g) + \\
&\quad (m-1)[\frac{1}{3} + \frac{1}{4}\eta_i^0(g) + \frac{1}{5}\eta_i^1(g) + ... + \frac{1}{r+4}\eta_i^r(g)] - c^h \\
&= \frac{1}{2} + \eta_i^0(g)(\frac{1}{3} + \frac{m-1}{4}) + ... + \eta_i^r(g)(\frac{1}{r+3} + \frac{m-1}{r+4}) - c^h \\
&\geq \frac{1}{2} + \eta_i^0(g)(\frac{1}{3} + \frac{1}{4}) + ... + \eta_i^r(g)(\frac{1}{r+3} + \frac{1}{r+4}) - c^h \\
&\geq \frac{1}{2} + \eta_i^0(g)(\frac{1}{3} + \frac{1}{4} - \frac{1}{2}) + ... + \eta_i^r(g)(\frac{1}{r+3} + \frac{1}{r+4} - \frac{1}{r+2})
\end{aligned}$$

where the last inequality follows from condition (iii). Since $\frac{1}{p} + \frac{1}{p+1} - \frac{1}{p-1} > 0$, $\forall p \geq 3$, we can conclude that $\Delta\Pi_j \geq 0$. By Lemma 2, player i will also have incentives to deviate. Thus g is not a PNE network. This contradiction completes the proof.

■

Proof of Theorem 1. This proof can be divided into four parts. Each of them is embodied by a corresponding proposition that rules out a set of potential BE networks. But first, let us show that a pseudo-empty network can be sustained in a BE for $c^h > \frac{1}{12}(3m^2 + 2m + 1)$. The unique possible deviation consists of creating a link between two agents (say i and j) of two different neighborhoods. The marginal payoff for these two agents is:

$$\Delta\Pi_j = \Delta\Pi_i = \frac{1}{2} + \frac{2(m-1)}{3} + \frac{(m-1)^2}{4} - c^h$$

By simple algebra, it follows that $\Delta\Pi_j = \Delta\Pi_i < 0$ if and only if the inequality stated above holds.

Below we focus on showing that, for a large n, no other topology can be sustained as a BE network for any c^h. Next proposition excludes the possibility of having essential players in a BE.

Proposition 14 *Suppose n is large and $m > 1$. A BE network cannot have essential players.*

Proof. From now on, it is assumed that $m > 1$. Notice that having a large population (n) will imply that m or M (or both) are very large. In what follows, we show that when one of these conditions holds, no player can be essential in a BE network.

Lemma 4 *For a large m, there cannot exist essential players in a BE network.*

Proof. By contradiction let us assume that g is a BE network with an essential player $i \in M_i$ and with an arbitrarily large m. Notice that there are, at least, two i-groups. The smallest group, at least, $m - 1$ players (the rest of members of M_i).

Since m is large, there is always the possibility to create a new link between two players of two different i-groups circumventing the essential player i. After that link is created, the deviators increase their access payoff with respect to the members of the other i-group. Since the minimum size of these groups depends positively on m, we conclude that, for any c^h, there is always a sufficiently large neighborhood size (m) that makes that deviation profitable. This contradicts the initial statement of stability and concludes the proof. ■

Now it is only left to show that the above result is reproduced when M is large. This is done in the next steps. First we present the following preliminary result:

Lemma 5 *For a large M, every essential player i has at least two links to each i-group in a BE network.*

Proof. By contradiction, imagine that g is a BE network with an essential player i linked to some i-group (say $G_j(i)$) through a single link $g_{ij} = 1$. The total population can be split into two groups: $G_j(i)$ and the rest of the society. For a fixed m, notice that a large total population implies that at least one of these two groups is also large. Let us assume that $G_j(i)$ is such a group. Consider that agents j and k (where $k \in M_i$ and $k \neq i$) deviate by creating a link between them. We claim that such a deviation is profitable for the involved players when M is large. After this deviation j will increase her intermediation payoff because she will avoid the essential player i in the intermediation between $G_j(i)$ and the rest of the society. Player k will also circumvent an essential player in order to access to the group $G_j(i)$. Then, for a given c^h, there is a sufficiently large $G_j(i)$ under which both deviators will obtain a positive marginal payoff. If $G_j(i)$ has a limited size, we can proceed symmetrically with respect to the other group (the rest of the society not included in $G_j(i)$) and find an analogous profitable deviation contradicting the initial statement of stability and concluding this proof. ■

The next two lemmas rule out the remaining possibility of observing essential players in BE networks for a large population and conclude the proof of the proposition.

Lemma 6 *For a large M, there cannot exist essential players in essential neighborhoods in a BE network.*

Proof. Let g be a BE network. Assume that g contains an essential neighborhood M_k with an essential player. Then, there is a pair of neighborhoods (M_i and M_j) located in different M_k-groups and linked to M_k. Moreover, there is player $k \in M_k$ who is essential for the connection between the members of M_i and M_j. Following the previous result, for a large M there are at least two links between each of these two M_k-groups (including M_i and M_j respectively) and M_k. Take two players from M_i and M_j linked to the essential neighborhood M_k. Then consider the deviation consisting of deleting their links to M_k and forming a link between them. Without increasing their costs they will circumvent the essential player k to access the other M_k-group. Therefore, they will strictly increase their access payoff. So, they have incentives to deviate, contradicting the initial statement of stability of g. ■

To conclude this proof, it is only left to show that there cannot exist essential players in non-essential neighborhoods. This proof is developed below:

Lemma 7 *For a large M, there cannot be essential players in non-essential neighborhoods in a BE network.*

Proof. By contradiction, let g be a BE network with an essential player (say i) in a non-essential neighborhood (M_i). There are two types of non-essential neighborhoods: (i) extreme and (ii) non-extreme. Below we develop them in turn:

(i) M_i is an extreme neighborhood. We have seen that i must have, at least, two external links. Let $g_{ij} = g_{ik} = 1$, where $j, k \in M_j$, be these two external links. Consider the deviation in which j severs the link g_{ij} and forms a new link with some player l such that $l \in M_i$ and $l \neq i$. The marginal payoff for player j will be positive, given that without increasing cost, one essential player (i) can be avoided in order to reach the members of M_i. On the other hand, after forming the link, player l would eliminate an essential player (i) in order to reach the rest of neighborhoods assuming the cost of an additional link (c^h). Again, for a sufficiently large population of neighborhoods, player l would also have incentives to deviate. Consequently, an extreme neighborhood cannot include an essential player in a BE network.

(ii) M_i is a non-extreme and non-essential neighborhood. Notice that $i \in M_i$ can be essential if and only if she is the unique agent in M_i with external links. Let $g_{ij} = 1$ where $j \notin M_i$. Consider a deviation consisting of severing the link between i and j, and forming a link between j and k where $k \in M_i$ and $k \neq i$. After that deviation takes place, j will have two different paths to communicate with any member of M_i. Then, the marginal payoff of player j will be positive since she avoids an essential player without increasing the cost. Player k will eliminate an essential player to access the rest of neighborhoods with an additional cost of c^h. Thus, for a large M, k will also deviate. Consequently, this kind of neighborhoods cannot contain an essential player either.

In each of the two points we reach a contradiction with the initial statement of stability of g, concluding the proof of the lemma and the proof of the proposition. ■

■

Let us move to the second part of the proof of the theorem. The following proposition excludes another large set of potential BE networks.

Proposition 15 *Suppose n is large. In a BE network, an essential neighborhood M_i cannot be connected to a M_i-group through a single neighborhood.*

Proof. By contradiction, let g be a BE network with an essential neighborhood M_i connected to some M_i-group (say $G_j(M_i)$) through a single neighborhood (say M_j). By the previous proposition we know that there are no essential players in g. Therefore, there must exist (at least) two agents in M_i linked to two different members of M_j. First, note that if there are more than two links, for example three links between M_i and M_j, there always exists some player in M_i (or M_j) who can sever a link without generating a new essential player. Thus, this deviator would obtain a positive net marginal payoff of c^h. Consequently, this network with three links between M_i and M_j is not a BE.

On the other hand, let g be a network where only two players $i, k \in M_i$ are connected with two different agents $l, j \in M_j$. Let us assume that $g_{il} = g_{kj} = 1$. Consider the following deviation: i and j sever their external links (g_{il} and g_{kj}) and simultaneously form a link between them. After deviating, agents i and j become essential in the connection between $G_j(M_i)$ and the rest of the society. Agent i's access payoff will decrease 1/6 with respect to any member of $G_j(M_i)$ (except j). But, after deviating, player i will get an intermediation payoff of 1/4 from the relationship between that member of $G_j(M_i)$ and any other member not included in $G_j(M_i)$. Since i, has the same number of links as before, it follows that her marginal payoff is positive. The argument for player j is symmetric. Thus, g is not a BE. This contradicts the initial statement of stability and concludes the proof. ■

Given this proposition, notice that all neighborhoods should be included in some cycle and any essential neighborhood should be a member of more than one cycle of neighborhoods. Thus, we still have to distinguish two remaining possible cases: a network with two or more cycles of neighborhoods and a single cycle including all neighborhoods. Next we present the last two propositions that take these up in turn.

Proposition 16 *A network that contains more than one cycle of neighborhoods cannot be a BE.*

Proof. By Proposition 14, there cannot be essential agents in a BE. Therefore, we focus on the cases without essential agents. Suppose that g is a BE network without essential players and with two or more cycles of neighborhoods. Let $\chi_1 = (M_1, M_2, ..., M_p)$ be an ordered set of neighborhoods in one cycle. Since g is a BE any cycle of p neighborhoods should have exactly p external links. Since g is connected, by the previous lemma, it follows that another cycle χ_2 must have some common neighborhood with χ_1. Let $\chi_2 = (\hat{M}_1, \hat{M}_2, ..., \hat{M}_q)$. We split this proof into two parts:

i First, let us consider that there is a single common neighborhood M_i between χ_1 and χ_2. Since there are no essential agents, M_i must have, at least, a pair of agents linked to each of the two cycles. More than two agents cannot exist in a BE as showed in the proof of the previous proposition. Consequently, there can be two, three or four agents in M_i with external links. Given the similarity between cases, we only consider the first possibility (i.e. only two members of M_i have external links).

Let i and j be the members of M_i with external links. In particular, $g_{ik} = g_{jl} = 1$ where k and l are members of neighborhoods in χ_1 (not included in χ_2) and $g_{ih} = g_{jt} = 1$ where h and t are members of neighborhoods in χ_2 (not included in χ_1). By the previous proposition, in a BE network, k and l should be members of different neighborhoods and so should h and t. Consider the following deviation: i and l sever their links to k and j respectively and simultaneously form a link between them. Notice that after this deviation takes place, we observe a line of $p-1$ neighborhoods (say L_1) connected to cycle χ_2 through two essential agents, i and l. Notice also that we can define a symmetric deviation that transforms the network into a line of $q-1$ neighborhoods (say L_2) connected to cycle χ_1 through two essential agents, i and t.

A sufficient condition for having a positive agent i's marginal payoff is that the new intermediation payoff that i obtains from the relationship between the members of χ_2 and the most distant neighborhood in L_1, exceeds the loss in the i's access payoff with respect to this neighborhood. This condition can be written as follows:

$$(A_2-1)\frac{1}{2p-1}+(A_2-1)(m-1)\frac{1}{2p}>(\frac{1}{2}-\frac{1}{2p-2})+(m-1)(\frac{1}{2}-\frac{1}{2p-1})$$

where $A_2 = mq$, i.e. the number of agents in χ_2. By simple algebra, we find that a sufficient condition for this inequality to hold is:

$$mq \geq p$$

On the other hand, the symmetric deviation introduced above generates a positive marginal payoff to agent i if:

$$mp \geq q$$

It is straight-forward that one of the two conditions stated above must hold. By contradiction let us assume that none of them holds, i.e. $mq < p$ and $mp < q$. Then it follows that:

$$m(q+p) < p+q$$

which is a contradiction given that $m > 1$. Then, one of the two symmetric deviations must be profitable for player i. Next we consider $mq \geq p$.

Below we analyze the marginal payoff of the other active agent l. In particular, we present two conditions which are sufficient for having a positive l's marginal payoff and show that they always hold:

- First, the new l's intermediation payoff between the members of her neighborhood and χ_2 should exceed the loss in her access payoff with respect to χ_2. This is written as follows:

  $$\frac{m-1}{3} + \frac{(m-1)(A_2-1)}{4} > \frac{A_2-1}{6}$$

  It is easy to see that this inequality holds for $m > 1$.

– On the other hand, the new l's intermediation payoff between the members of the most distant neighborhood in L_1 (say M_1) and the members of χ_2 should exceed the loss in her access payoff with respect to M_1. This can be written as follows:

$$\frac{A_2 - 1}{2p-1} + (A_2 - 1)(m-1)\frac{1}{2p} + \frac{1}{2p-2} + (m-1)\frac{1}{2p-1} > \left(\frac{1}{2} - \frac{1}{2p-3}\right) + (m-1)\left(\frac{1}{2} - \frac{1}{2p-2}\right)$$

A few algebraic operations lead to the conclusion that this inequality holds for $mq \geq p$.

Therefore, both i and l have incentives to deviate. This contradicts the initial statement of stability. Then, we cannot observe two cycles with only one common neighborhood in a BE.

ii If g has two or more common neighborhoods, except for three cases, it is always possible to describe two symmetric deviations analogously to the previous case. The corresponding sufficient conditions for having a positive net marginal payoff for the involved players are: $mq \geq p-(t-1)$ and $mp \geq q-(t-1)$ where m, p and q are defined as above and t is the number of common neighborhoods. As before, one of these two conditions must hold. Therefore, those networks cannot exist in a BE. The three cases in which these symmetric deviations cannot be defined are:

– Two cycles have two common neighborhoods M_i and M_j. Let $i \in M_i$ and $j \in M_j$ be such that $g_{ij} = 1$. Both agents have another external link to one of the cycles (the same one in both cases).

In such a case, both agents have a strict incentive to delete the link g_{ij}. This allows them to save c^h with no loss, because no new essential player arises after the deviation.

– Two cycles of neighborhoods (χ_1 and χ_2, defined as above) have t common neighborhoods $(M_i, \bar{M}_1, \bar{M}_2, ..., \bar{M}_{t-2}, M_j)$, where $t \geq 2$. Let $i \in M_i$ and $j \in M_j$ be the agents who have a link to a member of $\bar{M}_1$ and $\bar{M}_{t-2}$, respectively. Only one of these two agents (say i) has an external link to a member of one cycle (say χ_1). Moreover, in M_j only another agent (say k) has external links to members of χ_1 (not in χ_2) and χ_2 (not in χ_1). Notice that in such a case we cannot define two symmetric deviations as we did in part (i) because agent i becomes essential after any deviation that do not alter the cycle χ_1. Such an essential player does not arise in the deviations that transform part of χ_1 into a line of neighborhoods. However, we claim that at least one of the deviations considered in part (i) is profitable for the players involved. In order to show this claim, let us consider the case of only two common neighborhoods (M_i and M_j) in which $i \in M_i$ and $j \in M_j$ are such that $g_{ij} = 1$.

The analysis of the deviation that transforms the network into a line of $p-2$ neighborhoods connected to the cycle χ_2 is analogous to the analysis developed for case (i). Then, we can conclude that this deviation generates a positive marginal payoff for the players involved if:

$$mq \geq p - 1 \quad \text{(a)}$$

So, let us analyze the deviation that transforms the network into a line $q-2$

neighborhoods L_2 connected to the cycle χ_1 through two essential players (see case (i) for a description of the deviation). In particular, let us consider that $k \in M_j$ and l are the deviators who become essential in the connection between χ_1 and L_2. In addition, after the deviation player i will become essential between the members of her neighborhood (say M_i) and the rest of the society.

Proceeding analogously to case (i), we argue that a sufficient condition for having a positive agent k's marginal payoff is that the new k's intermediation payoff that comes from the relationship between the members of χ_1 and the most distant neighborhood in L_2 (say $\hat{M}_1$) exceeds the loss in the k's access payoff with respect to $\hat{M}_1$ and with respect to M_i. This condition is written as follows:

$$(A_1 - m)\frac{1}{2q-3} + (A_1 - m + 1)(m-1)\frac{1}{2q-2} + (m-1)^2\frac{1}{2q-1} > \left(\frac{1}{2} - \frac{1}{2q-4}\right) + (m-1)\left(\frac{1}{2} - \frac{1}{2q-3}\right) + (m-1)\left(\frac{1}{2} - \frac{1}{3}\right)$$

where $A_1 = mp$, i.e. the number of agents in χ_1. By simple algebra, we find that a sufficient condition for this inequality to hold is:

$$mp \geq \frac{4}{3}q - \frac{2}{3} \tag{b}$$

Assume condition (b) holds. Then, for the deviation to occur, the other deviator (l) should also obtain a positive marginal payoff. As above, we present two conditions which are sufficient for obtaining a positive l's marginal payoff and show that they always hold:

First, the new l's intermediation payoff between the members of her neighborhood and χ_1 should exceed the loss in her access payoff with respect to χ_1.

This is written as follows:

$$\frac{(m-1)^2}{5} + \frac{(m-1)(A_1-m)}{4} > \frac{A_1-m}{6} + \frac{m-1}{4}$$

This inequality holds for $m > 1$.

On the other hand, the new l's intermediation payoff from the relationship between the members of the most distant neighborhood in L_2 (say $\hat{M}_1$) and the members of χ_1 should exceed the loss in her access payoff with respect to $\hat{M}_1$. This is written as follows:

$$\frac{A_1-1}{2q-3} + (A_1-m+1)(m-1)\frac{1}{2q-2} + \frac{1}{2q-4} + \frac{(m-1)^2}{2q-1} >$$
$$(\frac{1}{2} - \frac{1}{2q-5}) + (m-1)(\frac{1}{2} - \frac{1}{2q-4})$$

After some algebra we conclude that this inequality holds for $mp \geq \frac{4}{3}q - \frac{2}{3}$. Consequently, agent l's marginal payoff is also positive when condition (b) holds.

Proceeding as before, we show that (since $m \geq 2$) one of these two conditions, (a) or (b), must hold. Therefore, one of the two deviations must be profitable for the involved players and g cannot be sustained as a BE, reaching the desired contradiction.

– Two cycles of neighborhoods (χ_1 and χ_2, defined as above) have t common neighborhoods $(M_i, \bar{M}_1, \bar{M}_2, ..., \bar{M}_{t-2}, M_j)$, where $t \geq 3$. Let $i \in M_i$ and $j \in M_j$ be the agents who have a link to a member of $\bar{M}_1$ and $\bar{M}_{t-2}$, respectively. Let $g_{ik} = g_{jl} = 1$ where $k \in \bar{M}_1$ and $l \in \bar{M}_{t-2}$. Both i and j have another external link to members of two different neighborhoods of the same cycle (say χ_1). Let $g_{ir} = g_{js} = 1$ be these two links.

We claim that such a network cannot be a BE. Let us consider the following deviation: k and j sever their links to i and l respectively and simultaneously form a link between them. Notice that after the deviation takes place, the subset of common neighborhoods $(\bar{M}_1, \bar{M}_2, ..., \bar{M}_{t-2})$ becomes a line of $t-2$ neighborhoods connected to the rest of neighborhoods, who form a cycle with no essential players (χ_3). Moreover, k and j become essential in the connection between these two parts of the network. As we have already seen, the deviators obtain a positive marginal payoff if:

$$A_3 \geq t - 1$$

where $A_3 = m(p + q - 2t + 2)$, i.e. the number of agents in χ_3.

Notice that we can define another deviation in which r and j sever their external links to i and s respectively and form a new link between them. After the deviation takes place, we observe a cycle (χ_2) of neighborhoods and a line of $p - t$ neighborhoods connected through two essential players, t and j. They obtain a positive marginal payoff if:

$$A_2 \geq p - t + 1$$

where $A_2 = mq$, i.e. the number of agents in χ_2.

One of these two conditions stated above must hold. By contradiction, let us assume that none of them holds. Then the following inequality must be true:

$$m(p + 2(q - t + 1)) \leq p$$

Since $m > 1$ and, by definition, $q > t$ this is a contradiction. Therefore, one of the deviations should be profitable for the players involved and the network cannot be a BE.

■

Proposition 17 *A single cycle containing all neighborhoods cannot be sustained in a BE.*

Proof. By Proposition 14, there cannot exist essential players in a BE. Therefore, we focus on the case without essential players. In this case, each neighborhood has two agents linked to two different neighborhoods. Let $\chi = (\hat{M}_1, \hat{M}_2, ..., \hat{M}_M)$ be the ordered set of neighborhoods in the cycle such that a member of $\hat{M}_i$ is linked to a member of $\hat{M}_{i+1}$ for all $i \in \{1, ..., M-1\}$ and a member of $\hat{M}_M$ is linked to a member of $\hat{M}_1$. Consider two players $i \in \hat{M}_i$ and $j \in \hat{M}_j$ such that: (i) $g_{i,k} = g_{j,l} = 1$, where $k \in \hat{M}_{i+1}$ and $l \in \hat{M}_{j+1}$, and (ii) $\hat{M}_i$ and $\hat{M}_j$ are farthest apart in terms of the minimum number of neighborhoods between them in the cycle. Now, consider the deviation in which i and j delete their external links to k and l respectively and form a link to each other. Notice that, by doing so, they create a line. Assume for simplicity that M is even, so that there are $(M-2)/2$ to one side of $\hat{M}_i$ and $(M-2)/2$ to the other side of $\hat{M}_j$ in the line created. Let $L = (M_1, M_2, ..., M_M)$ be the ordered set of neighborhoods in this line. Notice that the neighborhoods $M_{M/2}$ and $M_{M/2+1}$ of the set L correspond to the neighborhoods $\hat{M}_i$ and $\hat{M}_j$ of the set χ. We now show that players i and j will strictly increase their payoff with this coordinated deviation.

We proceed in two steps: the first step is to show that the individual payoffs are strictly increasing as we move toward the center of the line. The payoffs of an individual player consist of two components, the returns from accessing others and the returns for being essential on paths between pairs of other players. Number the players with external links in the line introduced above as $1, 2, ..., 2M-2$. Notice that the two players with external links in the center of this line are in positions $M-1$ and M.

The access returns to player l (assume for simplicity that l is even) are given by:

$$\frac{m-1}{l+1}+\frac{1}{l}+\frac{m-1}{l-1}+\frac{1}{l-2}+...+\frac{1}{2}+\frac{m-1}{2}+\frac{1}{3}+\frac{m-1}{4}+...+\frac{1}{2M-l-1}+\frac{m-1}{2M-l}$$

while access returns to player $l+2$ are given by:

$$\frac{m-1}{l+3}+\frac{1}{l+2}+\frac{m-1}{l+1}+\frac{1}{l}+...+\frac{1}{2}+\frac{m-1}{2}+\frac{1}{3}+\frac{m-1}{4}+...+\frac{1}{2M-l-3}+\frac{m-1}{2M-l-2}$$

It follows that access returns for player $l+2$ are larger than access returns for player l if $l < M - 3/2$. For l odd we can proceed analogously.

We now turn to the returns for being essential. The essentialness payoff to player l (assume for simplicity that l is even) are given by:

$$\sum_{i \in M_1}^{i \in M_{l/2}} \sum_{j \in M_{l/2+2}}^{j \in M_M} \frac{1}{e(i,j)+2} + \sum_{i \in M_1}^{i \in M_{l/2}} \sum_{\substack{k \in M_{l/2+1} \\ k \neq l}} \frac{1}{e(i,k)+2}$$

Similarly, the essentialness payoff to player $l+1$ is:

$$\sum_{i \in M_1}^{i \in M_{l/2}} \sum_{j \in M_{l/2+2}}^{j \in M_M} \frac{1}{e(i,j)+2} + \sum_{j \in M_{l/2+2}}^{j \in M_M} \sum_{\substack{k \in M_{l/2+1} \\ k \neq l}} \frac{1}{e(j,k)+2}$$

The first part of the essentialness payoffs to the two players are equal, while the second part of the payoffs are greater for player $l+1$ if $l < M/2$. For l odd we can proceed analogously.

To show that i and j indeed obtain a positive net marginal payoff, note that the aggregate gross payoffs obtained in both cases are the same. The above argument implies that i and j enjoy a higher share of total gross value in the line as compared to the other players. This implies that agents i and j earn a higher gross payoff in the line. Since their linking cost is the same in both cases, it follows that they obtain higher net payoffs as well, which completes the proof. ■ ■

Appendix C

Chapter 3

C.1 Proofs

Proof of Proposition 8. Let g be a PNE for a given pair (h, ρ_0). Let $i \in N$ be a member of this network who has $\mu_i \geq 2$. We claim that there exists a $\bar{\rho}_1 < \rho_0$ such that g will not be PNE for any pair (h, ρ) such that $\rho < \bar{\rho}_1$.

Let us focus on the case in which $\mu_i = 2$ (the cases in which $\mu_i > 2$ can be proved analogously). Let players k and j be the destinations of these two links, that is, $g_{ij} = g_{ik} = 1$. Here we consider the case in which $i \neq j$ and $i \neq k$. The case in which $i = j$ or $i = k$ is analogous, thus omitted. Next we analyze i's incentives to deviate. At this point we have to distinguish between the following two cases:

i $\sum_{l \in N_r(g)} \frac{1}{\mu_l} \geq \frac{q+1}{\rho}$ for $r = k, j$.

The i's marginal payoff for cutting the link g_{ij} off is positive if and only if:

$$\frac{f(h_i + h_k)}{\sum_{l \in N_k} \frac{1}{\mu_l} + \frac{1}{2}} > \frac{1}{2}\left[\frac{f(h_i + h_k)}{\sum_{l \in N_k} \frac{1}{\mu_l}} + \frac{f(h_i + h_j)}{\sum_{l \in N_j} \frac{1}{\mu_l}}\right] \qquad \text{(i)}$$

On the other hand, i's marginal payoff for cutting the link g_{ik} off is positive if and only if:

$$\frac{f(h_i+h_j)}{\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{2}} > \frac{1}{2}[\frac{f(h_i+h_k)}{\sum_{l\in N_k}\frac{1}{\mu_l}} + \frac{f(h_i+h_j)}{\sum_{l\in N_j}\frac{1}{\mu_l}}] \tag{ii}$$

Given that $0 < \frac{1}{\sum_{l\in N_r(g)}\frac{1}{\mu_l}} \leq \frac{\rho}{q+1}$ for $r=k,j$ we can say that:

$$lim_{\rho\to 0}(\frac{1}{\sum_{l\in N_r}\frac{1}{\mu_l}+\frac{1}{2}} - \frac{1}{\sum_{l\in N_r(g)}\frac{1}{\mu_l}}) = 0$$

Following the definition of limit we can say that for any $\varepsilon > 0$, there exists a ρ' such that $|\frac{1}{\sum_{l\in N_r}\frac{1}{\mu_l}+\frac{1}{2}} - \frac{1}{\sum_{l\in N_r(g)}\frac{1}{\mu_l}}| < \varepsilon$ for $r=k,j$, for any $\rho < \rho'$. Given that $h_j \neq h_k$, notice that if the LHS of conditions (i) and (ii) were equal to $\frac{f(h_i+h_r)}{\sum_{l\in N_r}\frac{1}{\mu_l}}$ for $r=k,j$ respectively, (i) and (ii) would be complementary. So, we conclude that for any specific vector h, and in particular for any specific triple (h_i,h_j,h_k), there exists a ρ' such that some of the two conditions (i) and (ii) has to hold for any $\rho < \rho'$. That is, for a sufficiently small ρ agent i will have incentives to deviate and cut one of her out-degree links off.

ii $\sum_{l\in N_j(g)}\frac{1}{\mu_l} \geq \frac{q+1}{\rho}$ and $\frac{q+1}{\rho} - \frac{1}{2} \leq \sum_{l\in N_k(g)}\frac{1}{\mu_l} < \frac{q+1}{\rho}$ (or vice versa).

The i's marginal payoff for cutting the link g_{ij} off is positive if and only if:

$$\frac{f(h_i+h_k)}{\sum_{l\in N_k}\frac{1}{\mu_l}+\frac{1}{2}} > \frac{1}{2}[\frac{\rho}{1+q}f(h_i+h_k) + \frac{f(h_i+h_j)}{\sum_{l\in N_j}\frac{1}{\mu_l}}] \tag{iii}$$

The i's marginal payoff for cutting the link g_{ik} off is positive if and only if:

$$\frac{f(h_i+h_j)}{\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{2}} > \frac{1}{2}[\frac{\rho}{1+q}f(h_i+h_k) + \frac{f(h_i+h_j)}{\sum_{l\in N_j}\frac{1}{\mu_l}}] \tag{iv}$$

Given that $0 < \frac{1}{\sum_{l\in N_r(g)}\frac{1}{\mu_l}+\frac{1}{2}} \leq \frac{\rho}{q+1}$ for $r=k,j$ we can say that $lim_{\rho\to 0}(\frac{1}{\sum_{l\in N_k}\frac{1}{\mu_l}+\frac{1}{2}} - \frac{\rho}{1+q}) = 0$ and $lim_{\rho\to 0}(\frac{1}{\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{2}} - \frac{1}{\sum_{l\in N_j}\frac{1}{\mu_l}}) = 0$. Following the definition of limit we can say that for any $\varepsilon > 0$, there exists a ρ'' such that $|\frac{1}{\sum_{l\in N_k}\frac{1}{\mu_l}+\frac{1}{2}} - \frac{\rho}{1+q}| < \varepsilon$ and $|\frac{1}{\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{2}} - \frac{1}{\sum_{l\in N_j}\frac{1}{\mu_l}}| < \varepsilon$ for any $\rho < \rho''$. Given that $h_j \neq h_k$, notice that if the LHS of conditions (iii) and (iv) were equal to $\frac{\rho f(h_i+h_k)}{1+q}$ and $\frac{f(h_i+h_j)}{\sum_{l\in N_j}\frac{1}{\mu_l}}$ respectively, (iii) and (iv) would be complementary. So, we conclude that for any specific vector h, and in particular for any specific triple (h_i, h_j, h_k), there exists a ρ'' such that some of the two conditions (iii) and (iv) has to hold for any $\rho < \rho''$. That is, for a sufficiently small ρ agent i will have incentives to deviate and cut one of her out-degree links off.

iii $\frac{q+1}{\rho} - \frac{1}{2} \leq \sum_{l\in N_r(g)}\frac{1}{\mu_l} < \frac{q+1}{\rho}$ for $r = k, j$.

iv $\sum_{l\in N_j(g)}\frac{1}{\mu_l} \geq \frac{q+1}{\rho}$ and $\sum_{l\in N_k(g)}\frac{1}{\mu_l} < \frac{q+1}{\rho} - \frac{1}{2}$.

v $\frac{q+1}{\rho} - \frac{1}{2} \leq \sum_{l\in N_j(g)}\frac{1}{\mu_l} < \frac{q+1}{\rho}$ and $\sum_{l\in N_k(g)}\frac{1}{\mu_l} < \frac{q+1}{\rho} - \frac{1}{2}$.

The proof for these three cases proceeds analogously to the previous one.

vi $\sum_{l\in N_r(g)}\frac{1}{\mu_l} < \frac{q+1}{\rho} - \frac{1}{2}$ for $r = k, j$.

In this case it is easy to observe that agent i will have incentives to sever the link with the agent with the lowest level of talent for any given ρ.

Therefore, in any possible case in which $\exists i \in N$ such that $\mu_i \geq 2$, we can find a sufficiently low value for ρ under which there is a profitable deviation. Defining $\bar{\rho}_1$ as the minimum of all these values of ρ, the proof of the proposition is done. ■

Proof of Proposition 9. Suppose we have an agent i such that $\sum_{l\in N_i}\frac{1}{\mu_l} \geq 1 + \frac{q+1}{\rho_0}$ in a network $g \in G^*_{h,\rho_0}$ for a given pair (h, ρ_0). First, we claim that there is a ρ (say

$\bar{\rho}_2$) such that for any $\rho < \bar{\rho}_2$, $\exists k \in N_i$ such that $k \neq i$, $k \notin M_i$ and:

$$cf(h_i + h_k) < \frac{1}{\sum_{l \in N_i} \frac{1}{\mu_l} - \frac{1}{\mu_k}} \Big(\sum_{l \in N_i \setminus \{i,k\}} \frac{1}{\mu_l} cf(h_l + h_i) + g_{ii} \frac{1}{\mu_i} f(h_i) \Big) \tag{v}$$

In words, for a sufficiently low ρ, some player k with a link towards i ($k \neq i$) has a joint productivity with i below the average productivity of i without considering researcher k. Notice that from Proposition 8, there exists a ρ (that we call $\bar{\rho}_1$) such that $\mu_i = 1$ in any $g \in G^*_{h,\rho}$ for any $\rho < \bar{\rho}_1$. Therefore, if at least two agents $j, k \in N_i$ have a joint productivity with i below the average of the rest of players in N_i, we know that one of them will not be in M_i for any $\rho < \bar{\rho}_1$, thus one of the players holds the conditions stated above. However, we can formulate a counter example. Imagine there is a single agent (say j) whose joint productivity with i is below the average of the rest of players in N_i. Moreover, let $j \in M_i$. In such a case, there exists no player k holding the conditions stated above. Next, we show that for a sufficiently low ρ this case cannot exist when $\sum_{l \in N_i} \frac{1}{\mu_l} \geq 1 + \frac{q+1}{\rho}$.

Consider an $\varepsilon > 0$ arbitrarily small. Imagine that $cf(h_j + h_i) = \varepsilon$ for $j \neq i$ and $f(h_j) = \varepsilon$ for $j = i$. Let $k \in N_i$, $k \neq j$ and:

$$\begin{cases} \varepsilon < cf(h_k + h_i) < cf(h_l + h_i) \ \forall l \in N_i \setminus \{i,k,j\} \text{ and } cf(h_k + h_i) < f(h_i) & \text{for } k \neq i \\ \varepsilon < f(h_k) < cf(h_l + h_i) \ \forall l \in N_i \setminus \{k,j\} & \text{for } k = i \end{cases}$$

Let us assume that $cf(h_k + h_i) = b$ for $k \neq i$ (or $f(k) = b$ for $k = i$). We know that there is an $\varepsilon > 0$ arbitrarily small under which $cf(h_l + h_i) > b + \varepsilon$ $\forall l \in N_i \setminus \{i, j, k\}$.Then, the average productivity of the players of N_i without including agent k is higher than:

$$\frac{1}{\sum_{l \in N_i} \frac{1}{\mu_l} - \frac{1}{\mu_k}} \Big(\varepsilon + \Big(\sum_{l \in N_i} \frac{1}{\mu_l} - \frac{1}{\mu_j} - \frac{1}{\mu_k} \Big)(b + \varepsilon) \Big)$$

After simple algebra we conclude that this expression is higher than b when the following condition holds:

$$b+\varepsilon > (b+\frac{b}{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}})\frac{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}}{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}-\frac{1}{\mu_j}+1}$$

Given that $\mu_j \geq 1$, $\frac{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}}{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}-\frac{1}{\mu_j}+1} \leq 1$. Since $\sum_{l\in N_i}\frac{1}{\mu_l} \geq 1+\frac{q+1}{\rho}$, we conclude that $lim_{\rho\to o}\frac{b}{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}} = 0$. This implies that for any $\varepsilon > 0$, we can always find a value of ρ (say ρ') such that $|\frac{b}{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}}| < \varepsilon$ for any $\rho < \rho'$. Therefore, for any $\rho < \rho'$ we have at least two agents (j and k) who hold condition (v). As commented above, one of them must fulfill the conditions stated in the beginning of the proof when $\rho < \bar{\rho}_1$. Let $\bar{\rho}_2 \equiv min(\bar{\rho}_1, \rho')$. That concludes the proof of the initial claim.

Our next step is to show that a network that holds these conditions cannot be sustained as a PNE for $\rho < \bar{\rho}_2$.

Agent i's marginal payoff for cutting the link g_{ki} off is:

$$\Delta\Pi_i = \frac{1}{\sum_{l\in N_i}\frac{1}{\mu_l}-\frac{1}{\mu_k}}[\sum_{l\in N_i\backslash\{i,k\}}\frac{1}{\mu_l}cf(h_l+h_i)+g_{ii}\frac{1}{\mu_i}f(h_i)] - \frac{1}{\sum_{l\in N_i}\frac{1}{\mu_l}}[\sum_{l\in N_i\backslash\{i\}}\frac{1}{\mu_l}cf(h_l+h_i)+g_{ii}\frac{1}{\mu_i}f(h_i)]$$

After simple algebra we can say that $\Delta\Pi_i > 0$ if and only if condition (v) holds. This happens by definition of k. Thus, agent i has incentives to deviate. ■

Proof of Proposition 10. Before proving this proposition we formulate an additional lemma.

Lemma 8 *If* $c \geq \frac{1}{2}$, $f(0) = 0$ *and* $f(\cdot)$ *is linear or convex, then* $cf(h_k+h_l) > f(h_l)$ *for* $h_k > h_l$.

Proof. *For a linear or convex $f(\cdot)$ with $f(0)=0$ and $c=\frac{1}{2}$, the inequality $cf(h_k+h_l)>f(h_l)$ reduces to $h_k>h_l$ after simple algebra. For $f(\cdot)$ convex and $c>\frac{1}{2}$, the difference between $cf(h_k+h_l)$ and $f(h_l)$ will be higher; therefore, the inequality of the statement also holds.* ■

By contradiction let us assume that we have a PNE network in which agent i holds $\sum_{l\in N_i}\frac{1}{\mu_l}<\frac{q+1}{\rho}-1$ and there exists an agent j with $h_j<h_i$ who holds: $g_{jk}=0$ $\forall k$ such that $h_k\geq h_i$. Imagine the deviation in which player j proposes to player i the formation of a new link. In such case, the marginal utility for player j is:

$$\Delta\Pi_j=\frac{1}{\mu_j+1}[\sum_{l\in M_j\setminus\{j\}}\hat{\Theta}(l)cf(h_l+h_j)+g_{jj}\hat{\Theta}(j)f(h_j)+\frac{\rho}{1+q}cf(h_j+h_i)]$$
$$-\frac{1}{\mu_j}[\sum_{l\in M_j\setminus\{j\}}\Theta(l)cf(h_l+h_j)+g_{jj}\Theta(j)f(h_j)]$$

where $\hat{\Theta}(l)$ corresponds to the variable $\Theta(l)$ after the deviation. After simple algebra we can write that $\Delta\Pi_j>0$ if and only if:

$$\frac{\rho}{1+q}cf(h_j+h_i)+\sum_{l\in M_j\setminus\{j\}}(\hat{\Theta}(l)-\Theta(l))cf(h_l+h_j)+g_{jj}(\hat{\Theta}(j)-\Theta(j))f(h_j)>$$
$$\frac{1}{\mu_j}[\sum_{l\in M_j\setminus\{j\}}\Theta(l)cf(h_l+h_j)+g_{jj}\Theta(j)f(h_j)]$$

On the one hand, by definition we know that $\Theta(l)$ is lower or equal than $\frac{\rho}{1+q}$ for any $l\in N$, and by assumption we also know that $h_l<h_i$ $\forall l\in M_j$. On the other hand, by Lemma 8 we can say that for a linear or convex $f(\cdot)$ and $c\geq\frac{1}{2}$, $cf(h_j+h_i)>f(h_j)$. These conditions imply that $\frac{\rho}{1+q}cf(h_j+h_i)>\frac{1}{\mu_j}[\sum_{l\in M_j\setminus\{j\}}\Theta(l)cf(h_l+h_j)+g_{jj}\Theta(j)f(h_j)]$. Given that for any $l\in M_j$, $\hat{\Theta}(l)>\Theta(l)$, we conclude that $\Delta\Pi_j>0$.

To complete the proof, we need to show that agent i will have incentives to form the link. Her marginal utility from accepting it is:

$$\Delta\Pi_i = \frac{\rho}{1+q}[\sum_{l\in N_i\setminus\{i\}} \frac{1}{\mu_l} cf(h_i+h_l) + g_{ii}\frac{1}{\mu_i}f(h_i) + \frac{1}{\mu_j+1}cf(h_i+h_j)]$$

$$+ \frac{c}{\mu_i} \sum_{\substack{l\in M_i\setminus\{i\}\\ l\in M_j}} \hat{\Theta}(l)cf(h_l+h_i)$$

$$- [\frac{\rho}{1+q}[\sum_{l\in N_i\setminus\{i\}} \frac{1}{\mu_l}cf(h_i+h_l) + g_{ii}\frac{1}{\mu_i}f(h_i)] + \frac{c}{\mu_i}\sum_{\substack{l\in M_i\setminus\{i\}\\ l\in M_j}} \Theta(l)cf(h_l+h_i)]$$

where $\hat{\Theta}(l)$ is defined as above. After simple algebra:

$$\Delta\Pi_i = \frac{\rho}{1+q}cf(h_i+h_j) + \frac{c}{\mu_i}\sum_{\substack{l\in M_i\setminus\{i\}\\ l\in M_j}} (\hat{\Theta}(l) - \Theta(l))cf(h_l+h_i)$$

Since $\hat{\Theta}(l) > \Theta(l)$ for any $l \in M_j$, we conclude that $\Delta\Pi_i > 0$, contradicting the initial statement of stability. ■

Proof of Lemma 1. In this proof we show that for any given pair (h, ρ) such that $\rho < \bar{\rho}_2$, if $g \in G^s$ there is only one possible distribution of links under which no player has incentives either to delete an in-degree or to create a new in-degree link coming from a player with a lower talent. This implies that only one network in G^s can hold this condition. To demonstrate this, we prove that any node i must have a given amount of in-degree links to not have incentives to change η_i,

Imagine a player i and a given in-degree $\eta_i > 0$. Given that $\rho < \bar{\rho}_2 \leq \bar{\rho}_1$, players can have at most a single out-degree link in any PNE network (Proposition 8). Then, $\sum_{l\in N_i} \frac{1}{\mu_l} = \eta_i \; \forall i \in N$. Moreover, since $\rho < \bar{\rho}_2$, we can say that $\frac{q+1}{\rho} - 1 \leq \eta_i < \frac{q+1}{\rho} + 1$ (Corollary 1). Given that η_i can only take natural numbers, there are at most two possible values for η_i in any PNE (say $\bar{\eta}$ and $\bar{\eta} - 1$). We claim that for only one of these two values agent i will not have incentives to change her in-degree for a given pair (h, ρ). First, the following must hold:

$$\frac{q+1}{\rho} - 1 \leq \bar{\eta} - 1 < \frac{q+1}{\rho} \leq \bar{\eta} < \frac{q+1}{\rho} + 1$$

Given these inequalities, we conclude that in any $g \in G^s$ no player i with $\eta_i = \bar{\eta}$ will have incentives to accept an additional in-degree link. Moreover, no player i with $\eta_i = \bar{\eta} - 1$ will have incentives to delete one existing in-degree link. Therefore, there are only two possibilities of deviation. An agent i can have incentives to cut one in-degree link off when $\eta_i = \bar{\eta}$ or she can have incentives to accept some additional in-degree link when $\eta_i = \bar{\eta} - 1$. Analyzing the marginal payoff of both deviations, we observe that if one is positive the other must be negative.

Let $\eta_i = \bar{\eta}$ and $g_{ji} = 1$ in a given network g_1.The agent i's marginal payoff for deleting the in-degree link g_{ji} is:

$$\Delta\Pi_i = \frac{\rho}{1+q}[\sum_{l \in N_i(g_1)\backslash\{i,j\}} cf(h_i + h_l) + g_{ii}f(h_i)] - \frac{1}{\bar{\eta}}[\sum_{l \in N_i(g_1)\backslash\{i\}} cf(h_i + h_l) + g_{ii}f(h_i)]$$

After simple algebra, $\Delta\Pi_i > 0$ if and only if:

$$(\frac{\bar{\eta}\rho}{1+q} - 1)[\sum_{l \in N_i(g_1)\backslash\{i,j\}} cf(h_i + h_l) + g_{ii}f(h_i)] > cf(h_i + h_j)$$

On the other hand, let $\eta_i = \bar{\eta} - 1$ and $g_{ji} = 0$ in a given network g_2 for some agent j such that $h_j < h_i$. The agent i's marginal payoff for creating the link g_{ji} is:

$$\Delta\Pi_i = \frac{1}{\bar{\eta}}[\sum_{l \in N_i(g_2)\backslash\{i\}} cf(h_i+h_l) + cf(h_i+h_j) + g_{ii}f(h_i)] - \frac{\rho}{1+q}[\sum_{l \in N_i(g_2)\backslash\{i\}} cf(h_i+h_l) + g_{ii}f(h_i)]$$

After simple algebra, $\Delta\Pi_i > 0$ if and only if:

$$(\frac{\bar{\eta}\rho}{1+q} - 1)[\sum_{l \in N_i(g_2)\backslash\{i\}} cf(h_i + h_l) + g_{ii}f(h_i)] < cf(h_i + h_j)$$

Notice that $\sum_{l \in N_i(g_2)\setminus\{i\}} cf(h_i + h_l) = \sum_{l \in N_i(g_1)\setminus\{i,j\}} cf(h_i + h_l)$. Therefore, one (and only one) of the two previous inequalities will hold for any vector h and any possible pair of players i and j such that $h_j < h_i$. Then, in G^s, any agent i can receive a single number of in-degree links in order to have no incentives neither to cut some in-degree link off nor to add some new in-degree link from an agent with a lower talent. This implies that only a single network in G^s holds these conditions. ■

Proof of Proposition 11. Let g be a PNE network for some pair (h, ρ_0), i.e. $g \in G^*_{h,\rho_0}$. Let g^s_{ij} denote the link g_{ij} in the network g^s ($g^s_{ij} = 1$ if and only if node i have a link towards j in g^s). Imagine that g is different from g^s. This implies that there exists an agent (say i) such that $g_{il} \neq g^s_{il}$ for some $l \in N$. There are various scenarios in which this holds. In the following lines, we show that there always exists a sufficiently low value of ρ (say $\hat{\rho}_1$) under which none of these scenarios can be sustained in a PNE for any $\rho < \hat{\rho}_1$. Let $j \in N$ be such that $g^s_{ij} = 1$.

(a) The first scenario is that agent i has more than one out-degree link. Following Proposition 8, we know that there exists a ρ (we called $\bar{\rho}_1$) such that such a network cannot be sustained as a PNE for any $\rho < \bar{\rho}_1$.

(b) Second, agent i can have a link to k, i.e. $g_{ik} = 1$, and $h_k < h_j$. Now let us show that this second case cannot hold in a PNE network. Two subcases need to be considered:

- $k = i$. Consider that player i creates a new link g_{ij}. Let us consider the extreme (and less favorable) case in which $o_i \leq 1$ and $\sum_{l \in N_j} \frac{1}{\mu_l} \geq \frac{q+1}{\rho}$. The agent i's marginal payoff from this deviation is:

$$\Delta\Pi_i = \frac{c}{\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{\mu_i+1}}\frac{1}{\mu_i+1}f(h_j+h_i) + \frac{\rho}{1+q}[\sum_{l\in N_i\setminus\{i\}}\frac{1}{\mu_l}cf(h_l+h_i)+\frac{1}{\mu_i+1}f(h_i)] - \frac{\rho}{1+q}[\sum_{l\in N_i\setminus\{i\}}\frac{1}{\mu_l}cf(h_l+h_i)+\frac{1}{\mu_i}f(h_i)]$$

We conclude that $\Delta\Pi_i > 0$ if and only if:

$$cf(h_j+h_i) > \frac{\rho(\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{\mu_i})}{1+q}\frac{1}{\mu_i}f(h_i) \quad \text{(vi)}$$

Given Proposition 9, $\exists\bar{\rho}_2$ such that $\sum_{l\in N_j}\frac{1}{\mu_l} < 1+\frac{q+1}{\rho}$ for any $\rho < \bar{\rho}_2$ in a PNE. Given that $\sum_{l\in N_j}\frac{1}{\mu_l} \geq \frac{q+1}{\rho}$, we can write that for any $\rho < \bar{\rho}_2$, $\frac{\rho(\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{\mu_i})}{1+q} \in [1+\frac{\rho}{1+q}\frac{1}{\mu_i}, 1+\frac{\rho}{1+q}(1+\frac{1}{\mu_i}))$. So, $lim_{\rho\to 0}\frac{\rho(\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{\mu_i})}{1+q} = 1$. On the other hand, from Lemma 8 we know that $cf(h_j+h_i) > f(h_i)$ for $f(\cdot)$ linear or convex and $c \geq \frac{1}{2}$. Following the definition of limit, we conclude that for a linear or convex $f(\cdot)$ and for any $\varepsilon > 0$, there always exists a $\rho' \leq \bar{\rho}_2$ such that $|\frac{\rho(\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{\mu_i})}{1+q} - 1| < \varepsilon$, $\forall\rho < \rho'$. Therefore, given that $\mu_i \geq 1$, we can state that, for any vector h, and in particular for any pair (h_j, h_i) we can find a sufficiently low value of ρ (say ρ') such that $\frac{\rho(\sum_{l\in N_i}\frac{1}{\mu_l}+\frac{1}{\mu_{i+1}})}{1+q}$ is sufficiently close to 1 to hold condition (vi) for any $\rho < \rho'$ (when $f(\cdot)$ is linear or convex).

– $k \neq i$. Let us assume the extreme (and less favorable) case in which $o_k < 1$ and $\sum_{l\in N_j}\frac{1}{\mu_l} \geq \frac{q+1}{\rho}$. The marginal payoff derived from adding the link g_{ij} to the network is positive when:

$$f(h_j+h_i) > \frac{(\sum_{l\in N_j}\frac{1}{\mu_l}+\frac{1}{\mu_i})\rho}{1+q}\frac{1}{\mu_i}f(h_i+h_k) \quad \text{(vii)}$$

From the previous point we know that for $\rho < \bar{\rho}_2$, $lim_{\rho \to 0} \frac{\rho(\sum_{l \in N_j} \frac{1}{\mu_l} + \frac{1}{\mu_i})}{1+q} = 1$. Given that in case (b) $h_j > h_k$ and that $\mu_i \geq 1$, following the definition of limit we conclude that there always exists a $\rho'' < \bar{\rho}_2$ such that $\frac{\rho(\sum_{l \in N_i} \frac{1}{\mu_l} + \frac{1}{\mu_{i+1}})}{1+q}$ is sufficiently close to 1 to hold condition (vii) for any $\rho < \rho''$.

(c) There is a third case in which i has a unique link to k and $h_k > h_j$. For any $\rho < \bar{\rho}_1$, no other node has more than one link. Since g is a PNE, this implies that if $g_{ik} = 1$, either $g_{ki} = 1$ and then k is in case (b) or $\exists l \in N$ such that $g^s_{lk} = 1$ and $g_{lk} = 0$; otherwise, agent k would have incentives to cut some in-degree link off (from the definition of g^s). If $\exists l \in N$ such that $g^s_{lk} = 1$ and $g_{lk} = 0$, notice that $g_{lr} = 1$ for some $r \in N$. If $h_r < h_k$ agent l is in case (b). If $h_r > h_k$ we are able to repeat the same argument as before. Since n is finite, we eventually reach an iteration in which some player would be in case (b). Consequently, if a player is in case (c), for a sufficiently low ρ there must exist a different player in case (b).

But for the link g_{ij} to be formed, node j must agree. If $o_j \leq 1$ after the deviation, player j's marginal payoff will be positive. On the other hand, let us consider the case in which $\sum_{l \in N_j} \frac{1}{\mu_l} \geq \frac{q+1}{\rho}$ (the case in which $\frac{q+1}{\rho} - 1 \leq \sum_{l \in N_j} \frac{1}{\mu_l} < \frac{q+1}{\rho}$ is analogous). Player j's marginal payoff for forming $g_{i,j}$ is:

$$\Delta\Pi_j = \frac{1}{\sum_{l \in N_j} \frac{1}{\mu_l} + \frac{1}{\mu_i+1}} [\sum_{l \in N_j \setminus \{j\}} c\frac{1}{\mu_l} f(h_l + h_j) + g_{jj}\frac{1}{\mu_j} f(h_j) + \frac{1}{\mu_i + 1} cf(h_i + h_j)] - \frac{1}{\sum_{l \in N_j} \frac{1}{\mu_l}} [\sum_{l \in N_j \setminus \{j\}} c\frac{1}{\mu_l} f(h_l + h_j) + g_{jj}\frac{1}{\mu_j} f(h_j)]$$

After simple algebra we conclude that $\Delta\Pi_j > 0$ if and only if:

$$cf(h_i + h_j) > \frac{\frac{1}{\mu_i+1}}{\sum_{l \in N_j} \frac{1}{\mu_l}} [\sum_{l \in N_j \setminus \{j\}} c\frac{1}{\mu_l} f(h_l + h_j) + g_{jj}\frac{1}{\mu_j} f(h_j)] \qquad \text{(viii)}$$

where the RHS can be interpreted as $\frac{1}{\mu_i+1}$ times the average productivity of the ideas stored in the queue of j. Given that agent j receives some link in g^s, we conclude that j will have a relatively high level of talent. Consequently, we can say that there should be a ρ^o such that for any $\rho < \rho^o$ the average productivity of the ideas of the queue of i depends positively on ρ (and, in consequence, it depends negatively on $\sum_{l \in N_j} \frac{1}{\mu_l}$). Therefore, since $\mu_i + 1 \geq 1$, it should exist a ρ''' such that for any $\rho < \rho'''$ condition (viii) holds.

Given that $\rho' \leq \bar{\rho}_2$ and $\rho'' \leq \bar{\rho}_2 \leq \bar{\rho}_1$, if we define $\hat{\rho}_1$ as $min(\rho', \rho'', \rho''')$ the initial claim is proved. ■

Proof of Proposition 12. For g^s to be a PNE network, it should be robust to each of the three deviations of the following list. We show that, for a sufficiently low ρ, g^s is robust to each of them. Let $g^s_{ij} = 1$

a Agent i changes the destination of her open ideas from j to herself (by definition of g^s, $h_i < h_j$). We claim that agent i's marginal payoff obtained from this deviation is negative for a sufficiently low value of ρ, and therefore, network g^s is robust to such a deviation for a sufficiently low ρ. To show it, let us consider the extreme (and less favorable) case in which researcher j holds $\sum_{l \in N_j} \frac{1}{\mu_l} \geq \frac{1+q}{\rho}$ and agent i holds $\sum_{l \in N_i} \frac{1}{\mu_l} + 1 < \frac{1+q}{\rho}$. In that case, i's marginal payoff will be negative when the following condition holds:

$$f(h_i) < \frac{1+q}{\rho \sum_{l \in N_j} \frac{1}{\mu_l}} c f(h_i + h_j) \qquad \text{(ix)}$$

Given that j receives some link in g^s (then, she has a relatively high level of talent), we can use Corollary 1 to state that $\frac{q+1}{\rho} - 1 \leq \sum_{l \in N_j(g^s)} \frac{1}{\mu_l} < 1 + \frac{q+1}{\rho}$

for a $\rho < \bar{\rho}_2$. After simple algebra, we observe that it is equivalent to say that $\frac{1+q}{1+q+\rho} < \frac{1+q}{\rho \sum_{l \in N_j} \frac{1}{\mu_l}} \leq \frac{1+q}{1+q-\rho}$. So, we conclude that for $\rho < \bar{\rho}_2$, $lim_{\rho \to 0} \frac{1+q}{\rho \sum_{l \in N_j} \frac{1}{\mu_l}} = 1$. On the other hand, from Lemma 8 we know that $cf(h_j + h_i) > f(h_i)$ for $f(\cdot)$ linear or convex, $h_i < h_j$ and $c \geq \frac{1}{2}$. Following the definition of limit, we can say that for any $\varepsilon > 0$, there always exists a $\rho' < \bar{\rho}_2$ such that $|\frac{1+q}{\rho \sum_{l \in N_j} \frac{1}{\mu_l}} - 1| < \varepsilon$, $\forall \rho < \rho'$. Then, for any h vector, and in particular for any pair (h_i, h_j) we can always find a sufficiently low value of ρ (say ρ') such that condition (ix) is hold for any $\rho < \rho'$, when $f(\cdot)$ is linear or convex. In this case, agent i will not have incentives to deviate.

b Agent i deletes one (or more) in-degree link. Given Lemma 1, the agent i's marginal payoff for deviating will be negative. Thus, g^s is robust to such deviation.

c Agent i proposes an additional link to some agent k with $h_k > h_j$. Given Lemma 1, agent k will have a negative marginal payoff for accepting the link g_{ik}. Thus g^s is also robust to such deviation by definition.

d Node i proposes the formation of the additional link g_{ik} to some player k with $h_k < h_j$[1]. We claim that the agent i's marginal payoff obtained from this deviation is negative for a sufficiently low value of ρ, and therefore, network g^s is also robust to such a deviation for a sufficiently low ρ. Let us consider the extreme (and less favorable) case in which $\sum_{l \in N_k} \frac{1}{\mu_l} + \frac{1}{2} < \frac{1+q}{\rho}$ and $\sum_{l \in N_j} \frac{1}{\mu_l} - \frac{1}{2} \geq \frac{1+q}{\rho}$. Researcher

[1]Notice that k is not necessarily different from i. The case in which $k = i$ is considered in part (e).

i's marginal utility obtained from that deviation would be:

$$\Delta\Pi_i = \frac{c}{\mu_i + 1}[\frac{1}{\sum_{l\in N_j}\frac{1}{\mu_l} - \frac{1}{2}} f(h_i + h_j) + \frac{\rho}{1+q} f(h_i + h_k)] - \frac{1}{\mu_i \sum_{l\in N_j}\frac{1}{\mu_l}} cf(h_i + h_j)$$

where $\mu_i = 1$. After simple algebra, we observe that $\Delta\Pi_i > 0$ if and only if:

$$f(h_i + h_k) > f(h_i + h_j)\frac{1+q}{\rho}\frac{\sum_{l\in N_j}\frac{1}{\mu_l} - 1}{(\sum_{l\in N_j}\frac{1}{\mu_l} - \frac{1}{2})(\sum_{l\in N_j}\frac{1}{\mu_l})} \qquad \text{(x)}$$

We know that $h_k < h_j$. On the other hand, given that j has a relatively high level of talent, we can use Corollary 1 to state that $\frac{q+1}{\rho} - 1 \leq \sum_{l\in N_j(g^s)}\frac{1}{\mu_l} < 1 + \frac{q+1}{\rho}$ for $\rho < \bar{\rho}_2$. Then we conclude that, for $\rho < \bar{\rho}_2$, $lim_{\rho\to o}\frac{1+q}{\rho}\frac{\sum_{l\in N_j}\frac{1}{\mu_l}-1}{(\sum_{l\in N_j}\frac{1}{\mu_l}-\frac{1}{2})(\sum_{l\in N_j}\frac{1}{\mu_l})} = 1$. Since $h_k < h_j$ and following the definition of limit, we conclude that for any vector h there always exists a sufficiently low value of ρ (say $\rho'' < \bar{\rho}_2$) such that condition (x) holds for any $\rho < \rho''$ and, as a consequence, g^s is robust to this deviation.

e Node i deviates by forming an additional link towards herself. We claim that the marginal payoff obtained from this deviation is negative for a sufficiently low value of ρ, and therefore, network g^s is also robust to such a deviation when ρ is sufficiently low. Let us consider the extreme (and less favorable) case in which $\sum_{l\in N_i}\frac{1}{\mu_l} + \frac{1}{2} < \frac{1+q}{\rho}$ and $\sum_{l\in N_j}\frac{1}{\mu_l} - \frac{1}{2} \geq \frac{1+q}{\rho}$. The marginal payoff obtained by i would be:

$$\Delta\Pi_i = \frac{1}{\mu_i + 1}[\frac{c}{\sum_{l\in N_j}\frac{1}{\mu_l} - \frac{1}{2}} f(h_i + h_j) + \frac{\rho}{1+q} f(h_i)] - \frac{c}{\mu_i}\frac{1}{\sum_{l\in N_j}\frac{1}{\mu_l}} f(h_i + h_j)$$

After simple algebra, we see that $\Delta\Pi_i > 0$ if and only if:

$$f(h_i) > cf(h_i + h_j)\frac{1+q}{\rho}\frac{\sum_{l\in N_j}\frac{1}{\mu_l} - 1}{(\sum_{l\in N_j}\frac{1}{\mu_l} - \frac{1}{2})(\sum_{l\in N_j}\frac{1}{\mu_l})} \quad \text{(xi)}$$

From the definition of g^s we know that $h_i < h_j$. Repeating the same arguments as before we conclude that $lim_{\rho\to o}\frac{1+q}{\rho}\frac{\sum_{l\in N_j}\frac{1}{\mu_l}-1}{(\sum_{l\in N_j}\frac{1}{\mu_l}-\frac{1}{2})(\sum_{l\in N_j}\frac{1}{\mu_l})} = 1$. On the other hand, from Lemma 8 we know that $cf(h_j+h_i) > f(h_i)$ for $f(\cdot)$ linear or convex, $h_i < h_j$ and $c \geq \frac{1}{2}$. Following the definition of limit, we can say that for any $\varepsilon > 0$, there always exists a ρ''' ($\rho''' < \bar{\rho}_2$) such that $|\frac{1+q}{\rho}\frac{\sum_{l\in N_j}\frac{1}{\mu_l}-1}{(\sum_{l\in N_j}\frac{1}{\mu_l}-\frac{1}{2})(\sum_{l\in N_j}\frac{1}{\mu_l})} - 1| < \varepsilon$, $\forall\rho < \rho'''$. Then, for any h vector, and in particular for any pair (h_i, h_j) we can always find a sufficiently low value of ρ (say ρ''') such that condition (xi) is hold for any $\rho < \rho'''$, when $f(\cdot)$ is linear or convex. In this case agent i will not have incentives to deviate; so g^s is robust to this deviation.

Defining $\hat{\rho}_2$ as $min(\rho', \rho'', \rho''')$ the claim of the proposition is proved. ■

Proof of Proposition 13. Let us divide the proof in two steps. First, we want to show that in g^s if we substitute the link g_{ij} by another one, say g_{ik}, such that $h_j > h_k$, the aggregate payoff will decrease when ρ is sufficiently small. Given g^s, $\mu_r = 1$ and $\sum_{l\in N_r}\frac{1}{\mu_i} = \eta_r$, $\forall r \in N$. Following the definition of g^s, if $\eta_r > 0$ then η_r can take one of these two possible values, $\eta_r = \bar{\eta}$ or $\eta_r = \bar{\eta} - 1$. In the proof of lemma 1 we show that:

$$\frac{q+1}{\rho} - 1 \leq \bar{\eta} - 1 < \frac{q+1}{\rho} \leq \bar{\eta} < \frac{q+1}{\rho} + 1$$

Let us assume the less favorable case (the one in which this marginal aggregate payoff is maximum) in which agent j has $\frac{q+1}{\rho} \leq \sum_{l\in N_j}\frac{1}{\mu_l} < \frac{q+1}{\rho} + 1$ and $o_k < 1$ even after receiving the additional link. In such a case:

$$\triangle \sum_{i \in N} \Pi_i = \frac{\rho}{1+q}[\sum_{l \in N_j \setminus \{i,j\}} cf(h_l + h_j) + g_{jj} f(h_j)] + \frac{2\rho}{1+q} cf(h_i + h_k) - [\frac{1}{\sum_{l \in N_j} \frac{1}{\mu_l}}[\sum_{l \in N_j \setminus \{j\}} cf(h_l + h_j) + g_{jj} f(h_j)] + \frac{1}{\sum_{l \in N_j} \frac{1}{\mu_l}} cf(h_i + h_j)]$$

After simple algebra $\triangle \sum_{i \in N} \Pi_i > 0$ if and only if:

$$(\frac{\rho \sum_{l \in N_j} \frac{1}{\mu_l}}{1+q} - 1)[\sum_{l \in N_j \setminus \{i,j\}} cf(h_l + h_j) + g_{jj} f(h_j)] > 2[cf(h_i + h_j) - \frac{\rho \sum_{l \in N_j} \frac{1}{\mu_l}}{1+q} cf(h_i + h_k)]$$

Since $\frac{q+1}{\rho} \leq \sum_{l \in N_j} \frac{1}{\mu_l} < \frac{q+1}{\rho} + 1$, $\frac{\rho \sum_{l \in N_j} \frac{1}{\mu_l}}{1+q} \in [1, 1 + \frac{\rho}{1+q})$. Then we can say that $lim_{\rho \to 0} \frac{\rho \sum_{l \in N_j} \frac{1}{\mu_l}}{1+q} = 1$. Given that $h_j > h_k$, we can say that $\exists \rho'$ such that the RHS will be higher than certain ε (> 0) for any $\rho < \rho'$. On the other hand, $lim_{\rho \to 0}(\frac{\rho \sum_{l \in N_j} \frac{1}{\mu_l}}{1+q} - 1) = 0$. Consequently, for the previous $\varepsilon > 0$, $\exists \rho''$ such that the LHS will be lower than ε for any $\rho < \rho''$. Then, we can always find a value of ρ (say ρ''') such that the last inequality will not hold for any h and for any $\rho < \rho'''$.

Second, we want to show that in g^s if we substitute the link g_{ij} by another one, say g_{ik}, such that $h_j < h_k$, the aggregate payoff will increase when ρ is sufficiently small. Let us assume the less favorable case (the one in which this marginal aggregate payoff is minimum) in which agent j has $\frac{q+1}{\rho} - 1 \leq \sum_{l \in N_j} \frac{1}{\mu_l} < \frac{q+1}{\rho}$ and $o_k \geq 1$. In such a case:

$$\triangle \sum_{i \in N} \Pi_i = \frac{1}{\sum_{l \in N_k} \frac{1}{\mu_l} + 1}(\sum_{l \in N_k} cf(h_l + h_k) + g_{kk} f(h_k) + cf(h_i + h_k)) + \frac{1}{\sum_{l \in N_k} \frac{1}{\mu_l} + 1} cf(h_i + h_k) + \frac{\rho}{1+q} \sum_{l \in N_j \setminus i} cf(h_l + h_j) - \frac{1}{\sum_{l \in N_k} \frac{1}{\mu_l}}(\sum_{l \in N_k} cf(h_l + h_k) + g_{kk} f(h_k)) - \frac{\rho}{1+q} cf(h_i + h_j) - \frac{\rho}{1+q} \sum_{l \in N_j} cf(h_l + h_j)$$

After simple algebra $\triangle \sum_{i \in N} \Pi_i > 0$ if and only if:

$$\frac{1}{\sum_{l \in N_k} \frac{1}{\mu_l}} \left(\sum_{l \in N_k} cf(h_l + h_k) + g_{kk} f(h_k) \right) < 2(cf(h_i + h_k) - \frac{\rho(\sum_{l \in N_k} \frac{1}{\mu_l} + 1)}{1+q} cf(h_i + h_j))$$

Since $o_k \geq 1$, we can say that $\frac{q+1}{\rho} \leq \sum_{l \in N_k} \frac{1}{\mu_l} < \frac{q+1}{\rho} + 1$. Then $\frac{\rho(\sum_{l \in N_k} \frac{1}{\mu_l} + 1)}{1+q} \in [1 + \frac{\rho}{1+q}, 1 + \frac{2\rho}{1+q})$. Consequently, $lim_{\rho \to 0} \frac{\rho(\sum_{l \in N_k} \frac{1}{\mu_l} + 1)}{1+q} = 1$. Since $h_j < h_k$, we can say that $\exists \rho^{iv}$ such that the RHS will be higher than a certain $\varepsilon > 0$ for any $\rho < \rho^{iv}$. On the other hand, $lim_{\rho \to 0} \frac{1}{\sum_{l \in N_k} \frac{1}{\mu_l}} = 0$. Then, for the previous $\varepsilon > 0$, $\exists \rho^v$ such that the LHS will be lower than ε for any $\rho < \rho^v$. Then, we can always find a value of ρ (say ρ^{vi}) such that the last inequality will hold for any h and any $\rho < \rho^{vi}$.

Defining ρ^* as $min(\rho''', \rho^{vi})$ the statement of Proposition 13 is proved. ■