European University Institute

DEPARTMENT OF ECONOMICS

# Network Structure Matters:

## Applications to R&D Collaboration, Collusion, and Online Communication Networks

Gizem Korkmaz

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 20 December 2012

DOI: 10.2870/63947

European University Institute
**Department of Economics**

# Network Structure Matters:

## Applications to R&D Collaboration, Collusion,

## and Online Communication Networks

Gizem Korkmaz

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

**Examining Board**

Prof. Elena Carletti, EUI
Prof. Rense Corten, Tilburg University
Prof. Jose Luis Moraga-González, VU University of Amsterdam
Prof. Fernando Vega Redondo, Supervisor, EUI

# Abstract

This thesis studies the interplay between network structure and strategic decision making given the backdrop of economic and social networks. The first two chapters study how firms' incentives to invest in costly R&D are affected by the pattern of R&D collaborations in a certain industry. These two chapters propose formal models that build upon and enrich the previous literature, which abstracted from two crucial dimensions of the problem. The first chapter introduces the possibility that inter-firm links aiming at R&D collaboration could facilitate market collusion. The second chapter incorporates network-based externalities resulting from informational flows and congestion that are associated with R&D collaborations. These chapters suggest that the benefits of possible inter-firm collaboration must be reevaluated from the point of their welfare consequences.

The last chapter aims to improve our understanding of how collective action spreads in large and complex networks in which agents use online social networks as communication tools. To this end, we develop a dynamic game-theoretic model of the "on-set of revolutions" that focuses on the local spread of information in order to study how network structure, knowledge and information-sharing interact in facilitating coordination through online communication networks.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

# Acknowledgments

There are many people to whom I would like show my appreciation, and to take the opportunity to thank. I am sincerely grateful to my supervisor, Fernando Vega-Redondo, for supporting and believing in me during all these years. His guidance and encouragement made this thesis possible. His immeasurable communication skills always turned meetings, which in other circumstances would have been stressful, into a sincere pleasure. It is with the strongest conviction that he will always be present in my professional life, now that I finish PhD.

A very special thanks to Fikret Adaman for encouraging me to pursue this path, pushing me to start the PhD, and opening the doors to research.

I am grateful to José L. Moraga-González, Rense Corten, Massimo Morelli, Paolo Pin, Luis Izquierdo, Elena Carletti, for their help and comments on parts of the thesis.

I owe much to my co-authors Christopher Kuhlman, Achla Marathe and Madhav Marathe, in particular for the pleasure it is to work with them at Virginia Tech. I look forward to seeing where our current research will bring us.

I also want to thank the administrative staff at Villa San Paolo, especially Jessica, for facilitating life in general, and patiently supporting me with practical problems that otherwise could have been overwhelming.

Although a PhD requires a lot of devotion, work is not all. Florence and I was not love at first sight. Yet with time it grew to be my “sehonda hasa” (second home). Not only was I introduced to a new way of enjoying life, but also to a way of eating that was new to me:

I learned that first of all, there is antipasto. It is supposed to be light. But for us economists, the beginning was the heaviest. For constant support and assistance, I thus wish to thank my peers, Asen, Charles, Emma, Maren, Oege, Rong, as well as my

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

colleagues Christoph, Davie, David, Oskar, Sarah and Teodora for being there for me. During this initial period of the thesis, you were the zuppa di cozze of my life.

Primo piatto, they call it. It was the first time that I discovered that leading the life of a PhD student could also be fun. Eunate, Tamara, Teodora, Alex, Punky and Igor, with a special lasagna like you, there is no need for secondo, not even if they offer me tagliata con ruccola!

Dolce, I could not have had a better home-sweet-home in Florence. The gelato from heaven accompanied by great friends in Via dei Servi kitchen: Thanks to Leticia, Aslı, Bahar, Chantal and Alejandro for always offering support and love.

And comes the favorite part: coffee and amaro. Alex, Anders, Charles, Davie, Davud, Markie, Leo, Oege, the Brazilian Moustaches. Not only did you share my love for lucano, you were the slice of orange that flavored my negronis after dinner for all these years. I know that whatever happens, we will always be there to toast to each other's sorrows, each other's joys.

I wish to thank to the most wonderful friends and cousin in Istanbul and Adana for their love, inestimable support and encouragement. A very special thanks to Hilal, also for making my visit to the US, such a great experience.

All these friends, among many others, endured this long process with me. They showed me what happiness looks like, and I cannot see what a greater gift could be.

My beloved mother, my sister, my aunt and my cousin Neşide. You made me who I am, and all I ever will achieve I owe to you. If this thesis is ought to be dedicated to anyone, it is especially to my mom. Had it not been for you, I could never have sat by this marvelous table.

# Contents

**Introduction** . . . . . . . . . . **1**

**1 R&D Collaboration in Collusive Networks** . . . . . . . . . . **3**

1.1 Introduction . . . . . . . . . . 3

1.2 Related Literature . . . . . . . . . . 5

1.3 The Model . . . . . . . . . . 8

1.4 Benchmark I: R&D Collaboration Networks . . . . . . . . . . 14

1.5 Benchmark II: R&D in Collusive Networks . . . . . . . . . . 21

1.6 R&D Collaboration in Collusive Networks . . . . . . . . . . 28

1.6.1 Why does the Configuration Matter? . . . . . . . . . . 28

1.6.2 Regular Networks with Complete Components . . . . . . . . . . 34

1.6.3 Regular Bipartite Networks . . . . . . . . . . 37

1.6.4 Ring Lattices . . . . . . . . . . 42

1.6.5 Welfare and Stability . . . . . . . . . . 44

1.7 Conclusion . . . . . . . . . . 49

1.8 Appendix . . . . . . . . . . 52

1.9 References . . . . . . . . . . 62

**2 Congestion in R&D Collaboration Networks** . . . . . . . . . . **65**

2.1 Introduction . . . . . . . . . . 65

2.2 Related Literature . . . . . . . . . . 67

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

2.3 The Model ... 70

2.4 Independent markets ... 74

2.5 Homogeneous-product oligopoly ... 77

2.6 Conclusion ... 85

2.7 Appendix ... 87

2.8 References ... 92

**3 Online Communication Networks and Collective Action ... 94**

3.1 Introduction ... 94

3.2 Related Literature ... 97

3.3 The Model ... 100

3.4 Common Knowledge ... 103

3.4.1 Full Information on the Network Structure ... 105

3.4.2 Local Information on the Network Structure ... 108

3.5 Deviations and the Equilibrium ... 122

3.5.1 Static Framework ... 122

3.5.2 Dynamic Framework ... 125

3.5.3 Examples ... 127

3.6 Homophily ... 131

3.7 Conclusion ... 135

3.8 Appendix ... 138

3.9 References ... 139

# Introduction

This thesis studies the interplay between the network structure and strategic decision making given the backdrop of economic and social networks. The first two chapters study how firms' incentives to invest in costly R&D are affected by the pattern of R&D collaborations between firms in a certain industry. These two chapters propose formal models that build upon and enrich the previous literature, which abstracted from two crucial dimensions of the problem. The first chapter introduces the possibility that inter-firm links aiming at R&D collaboration could facilitate market collusion. The second chapter incorporates network-based externalities resulting from informational flows and congestion that are associated with R&D collaborations. We develop models to examine the architecture of strategically stable networks and the relation between individual incentives and social welfare. These chapters suggest that the benefits of possible inter-firm collaboration must be reevaluated from the point of their welfare consequences.

In particular, the first chapter studies the R&D incentives of oligopolistic firms that can form both R&D collaborations and reciprocal market sharing agreements, which restrain them from entering each other's markets. We find that in the absence of R&D collaboration links, collusive market sharing agreements result in increased industry profits at the expense of R&D investments and social welfare. When firms undertake collaborative R&D and collude in the market, we show that both the degree and the configuration of the network are crucial in determining equilibrium outcomes. We find that some network structures, in particular bipartite networks, could lead to higher R&D levels and profits than other networks with the same size and degree. Moreover, in the complete bipartite network, the level of social welfare obained is even higher than in the absence of collusion. We also show that the complete network is always stable but not socially optimal, while the stability of the empty network and the complete bipartite graph depend on the number of firms and the level of R&D spillovers. In addition, networks composed of complete components are also stable when R&D spillovers are low.

The second chapter studies the trade-off between the synergies arising in research collaborations among firms that compete in the product market and the negative externalities resulting from their indirect connections. We focus on the inefficiencies, referred to as the congestion effect, that arise in collaborative environments. We find that when firms collaborate in R&D and operate in independent markets, individual R&D effort decreases in the number of collaborators when the congestion effect is high. Industry profits and welfare are also decreasing in the level of collaborative activity. Introducing product competition, we observe that, with a high congestion effect, the incentives to invest in R&D increase with the connectivity of the network, while profits and welfare decrease. Finally, we show that when the linking costs are low and the congestion effect is high, the complete network is not pairwise stable. However, intermediate levels of collaboration can be stable.

The last chapter of this thesis is joint work with F. Vega-Redondo, M. Marathe, A. Marathe and C. Kuhlman. We aim to improve our understanding of how collective action spreads in large and complex networks in which agents use online social networks as communication tools. To this end, we develop a dynamic game-theoretic model of the "on-set of revolutions" that focuses on the local spread of information in order to study how network structure, knowledge and information-sharing interact in facilitating coordination.

In the proposed coordination game, people use communication mechanisms resembling online social networks (e.g. Facebook) to inform each other of their willingness - or *threshold* - to participate in collective action. This communication facilitates coordination by creating common knowledge among the agents. We find that for a given network, if there exists a set of agents who share common knowledge, then a sub-graph induced by this set must be complete bipartite. We then introduce dynamics by assuming that people communicate their past actions in addition to their thresholds. In this dynamic context, further information is revealed as agents learn about the actions of their neighbors and make inferences based on knowledge of their neighbors' types. Finally, we study the role of correlation of thresholds - also known as *homophily* - in ring networks. We find that higher *homophily* enables people with high thresholds to coordinate and revolt, at the expense of a slower spread of collective action through the population.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

# Chapter 1

# R&D Collaboration in Collusive Networks

## 1.1 Introduction

In recent years, the globalization of markets and the deregulation of industries that used to be regulated on a territorial basis (airlines, local telecommunication services and utilities) have increased the scope for explicit or implicit market sharing agreements (Belleflamme and Bloch, 2004). Antitrust authorities started to emphasize market sharing agreements as an alternative form of collusion that limits competition in the product market. It is stated in the rules for the assessment of horizontal cooperation agreements under EU competition law that, in order to maintain effective competition during the joint exploitation of R&D, none of the parties involved may be fully excluded from exploiting the results in the internal market. Article 5 (f) of Draft R&D Block Exemption Regulation (2010) prohibits R&D agreements *"which have their objective as the requirement to refuse to meet demand from customers in the parties' respective territories, or from customers otherwise allocated between the parties by way of specialization in exploitation ...*"

This motivates us to analyze theoretically the R&D incentives of firms in an industry where competition can occur both within and across markets. Firms may collaborate in R&D and may collude in the market in order to exclude each other from their respective territories. We study the socio-economic consequences of this practice in order to find whether R&D Block Exemption Regulation (2010) can be justified from a social welfare perspective. We improve the existing models of strategic networks in order to analyze the relationship between market competition, firms' incentives to invest in R&D, and the architecture of collusive networks. In particular, we address the following questions:


Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

(i) What are the effects of collusive activity and the level of competition on the R&D levels and the profits of competing firms?

(ii) What are the incentives of firms to collaborate, and what is the architecture of "incentive compatible" networks?

(iii) What is the architecture of collaborative networks that maximize social welfare, with and without market sharing agreements?

(iv) Can collusive activity be beneficial from a social welfare point of view when there are R&D collaborations?

The main contribution of this chapter is to study R&D collaborations among firms by taking into account the possibility that these inter-firm agreements can facilitate collusion. The previous literature abstracted from this dimension, thus, R&D collaborations and collusive arrangements were studied independently. When the links between firms are defined as R&D collaborations, as in Goyal and Moraga-Gonzalez (2001), or as market sharing agreements, as in Belleflamme and Bloch (2004), the only relevant structural network parameter is the degree (given the size) of the network. However, when links have the double and inseparable role as collaboration both in R&D and market collusion, the nature of the problem is dependent on the structure of interaction. Different network configurations (for the same connectivity) induce different equilibrium outcomes. Therefore, we suggest that policy makers whose objective is to maximize social welfare should not disregard these factors.

The rest of the chapter is organized as follows. We begin with a brief overview of the relevant literature. In Section 1.3, we present the model. In Section 1.4, we analyze the equilibrium and study efficiency and stability for the case in which firms collaborate on R&D in the absence of market sharing agreements. In Section 1.5, we analyze the case in which firms collude via market sharing agreements but undertake and benefit from R&D individually. Section 1.6 presents the results for the case when firms have both R&D collaborations and market sharing agreements. Section 1.7 concludes and discusses some possible extensions for further research.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

## 1.2 Related Literature

How competition affects incentives to invest in R&D is still debated in the industrial organization literature. Both the theoretical and empirical research on the link between market structure and innovation is not conclusive (Motta, 2004). Theories which suggest that a monopolist has less incentive to innovate than a competitive firm rely upon the Arrow's replacement effect. Competition pushes firms to invest in order to improve their competitive position relative to rivals. Schumpeter (1943), on the other hand, stressed the necessity of tolerating the creation of monopolies as a way to encourage innovation. Motta (2004) presents a model in which firms invest non-cooperatively in R&D and then compete in quantities. Taking the number of firms in the industry as a measure for the intensity of competition, he shows that a monopolistic market structure leads to lower R&D investment than a market structure where several firms co-exist and behave non-cooperatively. However, social welfare requires an intermediate level of competition. This result is also suggested by empirical studies such as Aghion et al (2002).

In the theoretical literature, the intensity of competition is measured either by the number of firms in the market, by the degree of product substitutability, or by the nature of competition (Bertrand vs. Cournot). We suggest that the number of firms may not fully capture the structure of the market and that it is necessary to also consider the collaborations and collusive agreements between them. Inter-firm agreements have the distinctive structural features of being *nonexclusive* (Milgrom and Roberts, 1992) and *bilateral* (Delapierre and Mytelka, 1998). While the coalition formation approach requires every player to belong to one group only, the non-exclusivity allows for a pattern of relations in which there are agreements between firms 1 and 2, and firms 2 and 3, but no such link between firms 1 and 3. The bilateral structure allows us to adopt a network approach to analyze these agreements with a strategic model of R&D with *pairwise* collusive links and to take the number and structure of collusive agreements (defined as the connectivity and configuration of the network) as an alternative measure for the degree of competition.

The literature on R&D cooperation as a way to internalize spillovers starts with d'Aspremont and Jacquemin (1988) who compare cooperative and non-cooperative schemes in oligopolistic industries, focusing on cost-reducing R&D. This strand of literature suggests that cooperation

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

increases the level of R&D, resulting in higher output and welfare. Non-cooperative incentives to undertake R&D are examined by a recent theoretical trend that adopts strategic models of network formation. Goyal and Moraga-Gonzalez (2001) contribute to this literature by providing a model in which the "quality" of the links is *endogenously* determined -via the choice of R&D efforts- by the players. They consider an oligopoly in which identical firms have the opportunity to form pairwise collaborative links in order to share nonexclusive R&D knowledge about a cost-reducing technology. Given this collaboration network, firms unilaterally choose a single (costly) level of R&D effort and operate (in independent markets and in a homogeneous-product market) by setting quantities. The authors show that when firms operate in independent markets, total R&D, industry profits and welfare are maximized in the complete network which is pairwise stable. However, when the collaborators compete in the homogeneous-product market, individual R&D effort declines in the level of collaborative activity. The complete network is the unique stable network but profits and welfare are maximized at intermediate levels of collaboration.

Collusion by means of market sharing agreements is studied by Belleflamme and Bloch (2004), who characterize stable collusive networks in oligopolistic markets with symmetric firms and identical markets. The market sharing agreements restrain firms from entering each other's markets, decreasing the number of firms in the markets and, at the same time reducing the number of foreign markets in which they can operate. The authors find that the socially efficient network is the empty network whereas the industry profits are maximized in the complete network which corresponds to a situation where firms are local monopolies in their home markets. They show that in a stable network, firms form complete components and they study the minimal size of the components. We contribute to this literature by extending the model to analyze the R&D incentives of firms that form market sharing agreements.

We study two benchmark models before introducing a generalized model with two types of agreements. The first benchmark model slightly extends the Goyal and Moraga-Gonzalez (2001) paper by introducing $n$ markets instead of one, and by allowing for imperfect spillovers between collaborators. We are therefore able to replicate their results. We show that there is an intermediate level of connectivity that maximizes industry profits and welfare when R&D

spillovers are high. When R&D collaborators are competitors in the homogeneous-product market, the empty network is not stable. The complete network is stable, but does not lead to a socially optimal outcome. Total R&D is highest in the empty network, but the outcome is not socially optimal. Efficiency requires an intermediate level of collaborative activity.

The second benchmark model is an extension of Belleflamme and Bloch (2004) to a setting where colluding firms undertake individual R&D that does not benefit other firms. We show that when firms invest in R&D and collude by forming market sharing agreements, the complete network, in which all firms are local monopolies, is pairwise stable. The complete network maximizes industry profits, but does not lead to a socially optimal outcome. Efficiency requires an intermediate level of collusive activity when the number of markets (firms) is sufficiently high. Total R&D is the highest under the empty network, which is also efficient and pairwise stable when there is a small number of competitors.

Finally, we develop a generalized model in which the links between firms are interpreted as both R&D collaboration and market sharing agreements. We show that in this case the results are dependent on the network structure. Networks with the same level of connectivity and same size can result in different equilibria, depending on the topology of the network. The equilibrium level of R&D effort is lowest in regular networks with complete components and is higher in ring lattices and bipartite graphs with the same size, $n$, and degree, $k$. The complete bipartite graph results in the highest equilibrium effort. Industry profits increase with respect to the degree in all of these configurations, but the highest social welfare is attained in the complete bipartite graph. Therefore, some intermediate level of collaboration is optimal for total R&D, profits and efficiency. Given the collaborative and collusive agreements, we also show that the regular networks with two complete components are stable only when R&D spillovers are very low. The complete network is always stable, while the stability of the empty network depends on the number of markets and the level of R&D spillovers.

## 1.3 The Model

We consider a three-stage game. In the first stage, firms form pairwise R&D collaboration links and reciprocal market sharing agreements. In the second stage, each firm chooses a level of costly effort in R&D. Along with the collaboration; the R&D effort defines the costs of the firms such that one firm's effort reduces the costs of its collaborators. In the last (market) stage, firms compete by setting quantities of the homogeneous goods. In the existence of market sharing agreements, firms compete only in the markets in which they have not formed any collusive agreements.[1]

In order to have a better understanding of the different forces generated by R&D collaboration and collusion, we start the analysis with two benchmark cases. In the first case, firms form R&D collaboration links and compete with all firms in the last stage of the game. In the second case, there are no collaboration links, but firms can form collusive agreements. Finally, we present a model with both type of inter-firm agreements and argue that the interaction between firms due to the R&D collaboration can facilitate the possibility of collusion.

### 1.3.1 Networks

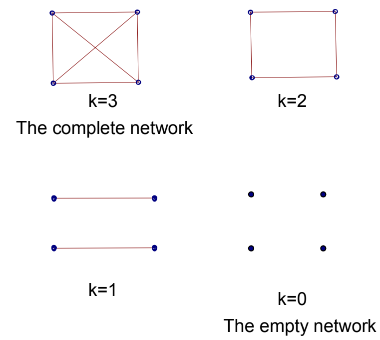
Let $N = \{1, 2, ..., n\}$ , $n \geq 2$ denote the set of firms. We associate a market to each firm such that the market of firm $i$ is interpreted as its *home* market and firm $i$ has access to all foreign markets unless it has market sharing agreements. For any pair of firms $i, j \in N$, the pairwise relationship between the two firms is represented by a binary variable $g_{ij} \in \{0, 1\}$. When $g_{ij} = 1$, two firms are linked; when $g_{ij} = 0$ there is no link between firm $i$ and $j$. Links are assumed to be costless. A network $g = \{g_{ij}\}_{i,j \in N}$ is a collection of pairwise collaboration and/or collusive links between the firms. Let $g + g_{ij}$ denote the network obtained by adding a new link between firms $i$ and $j$, e.g. by replacing $g_{ij} = 0$ in network $g$ with $g_{ij} = 1$. Similarly, $g - g_{ij}$ is the network obtained by severing an existing link between firms $i$ and $j$ from network $g$ by replacing $g_{ij} = 1$ with $g_{ij} = 0$. Let $N_i(g) = \{j \in N \setminus i : g_{ij} = 1\}$ be the set of firm $i$'s *neighbors* in network $g$, and let $\eta_i(g)$ be the cardinality of the set $N_i(g)$.

[1] We assume that these agreements are enforceable.


DOI: 10.2870/63947

We focus on regular networks where $\eta_i(g) = \eta_j(g) = k$ for any two firms $i$ and $j$, and where $k$ is the *degree* of the network, or the *level of collaborative and/or collusive activity,* depending on the model we study. Figure 1 below illustrates regular networks with 4 firms.

*Figure 1. Regular Networks for* $n = 4$




In the first benchmark model, as in Goyal and Moraga-Gonzalez (2001), we study the case in which links define only R&D collaborations between firms. Neighbors benefit from each other's R&D investments, and then compete with all firms in the homogenous-product market. When all the firms collaborate on R&D with each other, we obtain the complete network. In the empty network, firms do not benefit from other firms' R&D investments since there are no collaboration links.

In the second benchmark model, as in Belleflamme and Bloch (2004), the links represent only the market sharing agreements that limit competition in the product market. Here, there are no R&D spillovers between firms. In this case, the complete network corresponds to a situation in which the firms are local monopolies in their home markets. When the network is empty, all firms operate in all markets. The market sharing agreements are bilateral and only restrict firms from operating in their respective markets. Therefore, depending on the network formed, they can compete in a third market. For example, in the figure above, when we consider the regular network with $k = 1$, we see two pairs of linked firms. The firms on the top do not enter each others' markets due to their collusive agreement but they both enter the two remaining markets. Therefore, they are competitors in those markets. On the other hand, in the cycle network, i.e., $k = 2$, the colluding firms never compete because each firm $i$ has only one foreign market, say $j$, to enter but both of the neighbors of firm $i$ are linked to firm $j$ (thus, $i$'s neighbors do not enter to market $j$).

Finally, in the generalized model, links are interpreted as both market sharing agreements and R&D collaborations. This allows us to study the welfare implications of R&D collaborations under limited competition due to market sharing agreements.[2] As discussed above, firms that form market sharing agreements can compete in a third market. This will affect firms' R&D investment decisions since their neighbors will benefit from their R&D, becoming tougher competitors in those markets in which they both operate. Therefore, as opposed to the benchmark cases, both the degree and the configuration of the network play a crucial role in this model. In other words, given any two networks with same size and connectivity, firms can face more of their collaborators as competitors in one network than in the other, depending on the topology. Hence, one needs to distinguish between different families of graphs, and to study them in a systematic way in order to show that different configurations result in different equilibria.

Our model builds upon strongly regular graphs: A $k$-regular simple graph $G$ on $n$ nodes is *strongly $k$-regular* if there exist positive integers $k$, $\lambda$, and $\mu$ such that every vertex has $k$ neighbors (i.e., the graph is a regular graph), every adjacent pair of vertices has $\lambda$ common neighbors, and every nonadjacent pair has $\mu$ common neighbors. (West 2000, pp. 464-465).

In other words, all nodes have the same number of neighbors, all linked nodes (i.e., neighbours) have the same number of common neighbors, $\lambda$, and finally, all unlinked node pairs have the same number of common neighbors, $\mu$, in strongly regular graphs. As discussed above, firms in our model take into account whether they will compete with their collaborators in the foreign markets that they enter, i.e., whether linked firms (neighbors) have common "non-neighbors". Therefore, the first order conditions of our problem are characterized by the number of common "non-neighbors" that linked firms have in the network. This is closely related to the parameters $\lambda$ and $\mu$ in the definition of strongly regular graphs above.

[2] We are aware that not all collaborating firms form collusive agreements and that assuming that all collaborators collude in the product market is a strong assumption. One can construct a more complex model with two types of links (networks) to study whether firms have incentives to collude with non-collaborators or not. The underlying argument here is that when firms undertake R&D together, the interaction makes it easier for the collaborators to collude in the market stage. Our model is a particular case of the more general two-network model and it can be thought as the extreme (worst) case in which all collaborators collude.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

In order to study a wider set of graphs than strongly regular graphs, we relax the requirement that both $\lambda$ and $\mu$ are fixed in the network. We allow one of the two parameters, $\lambda$ or $\mu$, to take different values in the same network, keeping their distribution fixed for all nodes. Thus, let us define $\lambda_{ij}$ as the number of common neighbors of adjacent pairs $i$, $j$; and let $\mu_{i,m}$ denote the number of common neighbors of nonadjacent pairs $i$ and $m$ such that every node $i$ has the same distribution of $\lambda_{ij}$ and $\mu_{i,m}$ in the network. First of all, we observe that for any linked pair $ij$, there are $2 + 2(k-1) - \lambda_{ij}$ nodes in $N_i \cup N_j$ since both have $k-1$ remaining neighbors and $\lambda_{ij}$ of them are common. Thus, they have $n - [2 + 2(k-1) - \lambda_{ij}] = n - 2k + \lambda_{ij}$ common non-neighbors in the network, i.e., markets that they both enter. Second, for any unlinked pair $i$ and $m,$ the number of collaborators that firm $i$ faces in market $m$ can simply be calculated as $k - \mu_{i,m}$. Firm $i$ has $k$ neighbors, and by definition $\mu_{i,m}$ of them are also neighbors of firm $m.$ Therefore, only $k - \mu_{i,m}$ of $i$'s neighbors operate in market $m.$ If $\lambda_{ij} = \lambda$ for all adjacent pairs, then the values that $\mu_{i,j}$'s can take can be characterized by focusing on certain families of networks. In other words, since the parameters $\lambda$ and $\mu$ are not independent, if one is fixed in the network, the other can be varied, allowing different network configurations to be studied systematically.

To understand better, Figures 2 and 3 below illustrate different possible regular network configurations.

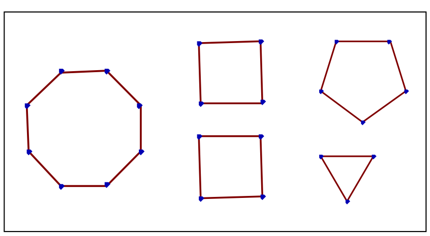


*Figure 2.* $n = 8$, $k = 2$

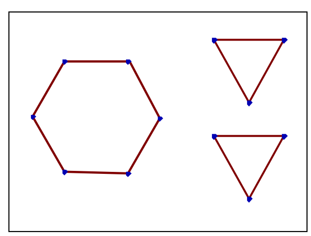


*Figure 3.* $n = 6$, $k = 2$

As we observe in Figure 2, for same size, $n = 8$, and same degree, $k = 2$, one can obtain three different network configurations, which are weakly regular. The first is a cycle graph with $(n, k, \lambda, \mu) = (8, 2, 0, \{0, 1\})$, the second, with $(n, k, \lambda, \mu) = (8, 2, 0, \{0, 2\})$ consists of two components with same sizes and the last, with $(n, k, \lambda, \mu) = (8, 2, \{0, 1\}, \{0, 1\})$, has two distinct components. We will see later in detail that the first-order conditions of our problem will be the same for every node in the cycle and in the two squares. However, the first-order conditions are different in the components of the last graph. In order to focus on symmetric equilibria, we restrict attention to graphs in which we keep the distribution of either $\lambda$ or $\mu$ constant in the network and allow the other to take different values.

Figure 3 illustrates two regular networks with $n = 6$ and $k = 2$. The first is a cycle graph with $(n, k, \lambda, \mu) = (6, 2, 0, \{0, 1\})$, and the second, with $(n, k, \lambda, \mu) = (6, 2, 1, 0)$, corresponds to two triangles. The latter (which is a strongly regular graph) consists of two components which are both complete. In this case, a firm competes with both of its neighbors in all three foreign markets since $k - \mu_{i,m} = k \quad \forall m \notin N_i$. On the other hand, in the cycle with $\mu_{i,m} \in \{0, 1\}$, say $\mu_{i,m} = (\mu^1_{i,m}, \mu^2_{i,m}) = (0, 1)$, there is only one market in which all neighbors compete ($k - \mu^1_{i,m} = k$) and two other markets where the firm competes with only one of its neighbor since $k - \mu^2_{i,m} = 1$.

### 1.3.2 R&D effort levels and spillovers

Given a network $g$, every firm unilaterally chooses an R&D effort level $e_i(g)$. This individual effort lowers the firm's own marginal cost and also has positive spillovers on the costs of other firms that have a collaboration link with this firm. We assume that there are no knowledge spillovers between unlinked firms.

We assume that firms are initially symmetric, with zero fixed costs and identical constant marginal costs $\bar{c}$. Given a network $g$ and the collection of effort levels $\{e_i(g)\}_{i \in N}$, firm $i$'s cost is given as follows:

$$c_i(\{e_i\}_{i \in N}) = \bar{c} - e_i - \beta \sum_{j \in N_i(g)} e_j$$

R&D effort is costly such that, given a level of effort $e_i \in [0, \bar{c}]$, the cost is $\Psi(e_i) = \gamma n e_i^2$, $\gamma > 0$. Under this specification, the cost of R&D effort is an increasing function exhibiting decreasing returns. Finally, $\beta \in [0, 1]$ measures the R&D spillovers between collaborating firms such that when $\beta = 1$, collaborators fully benefit from each others' R&D investments (as in Goyal and Moraga-Gonzalez, 2001), while $\beta = 0$ implies that there are no R&D spillovers between collaborators.

### 1.3.3 Payoffs

Given these costs, firms operate in the markets by choosing quantities $\{q_i^j(g)\}_{\substack{i,j\in N\\ g_{ij}=0}}$. Demand is assumed to be linear, and is given by $Q = a - p$ in all markets. Let $\pi_i^j(g)$ be the net profits attained by firm $i$ in market $j$ and network $g$. The total profits of firm $i$ are given by the sum of the profits firm $i$ collects in its home market and in all foreign markets for which it has not formed market sharing agreements:

$$\begin{aligned}\Pi_i(g) &= \pi_i^i(g) + \sum_{\substack{j\in N\\ g_{ij}=0}} \pi_i^j(g) - \gamma n e_i^2 \\ \text{where } \pi_i^j(g) &= \left[a - Q^j(g) - c_i(g)\right] q_i^j(g)\end{aligned}$$

### 1.3.4 Welfare

For any network $g$, social welfare is defined as the sum of consumer surplus and producers' profits in all markets. Aggregate welfare in network $g$, which is denoted by $W(g)$, is given by:

$$W(g) = \sum_{j=1}^{N} \frac{Q^j(g)^2}{2} + \sum_{i=1}^{N} \Pi_i(g)$$

where $Q^j(g) = \sum_{i\in N} q_i^j(g)$ is the aggregate output in market $j$. A network $g$ is *efficient* iff $W(g) \geq W(g')$ for all $g'$.

### 1.3.5 Stability

Following Jackson and Wolinsky (1996), a network $g$ is *stable* if and only if for all $i, j \in N$,

(i) if $g_{ij} = 1$, then $\Pi_i(g) \geq \Pi_i(g - g_{ij})$ and $\Pi_j(g) \geq \Pi_j(g - g_{ij})$

(ii) if $g_{ij} = 0$ and $\Pi_i(g + g_{ij}) > \Pi_i(g)$, then $\Pi_j(g + g_{ij}) < \Pi_j(g)$.

The idea is that, while a link can be severed unilaterally, a link can be formed if and only if the two firms involved agree to do so.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

## 1.4 Benchmark I: R&D Collaboration Networks

In this section, we analyze the model in which firms can form R&D collaboration agreements in the first stage and then unilaterally choose the level of R&D efforts that maximize individual profits. In the last stage, all firms compete in all markets. This is similar to Goyal and Moraga-Gonzalez (2001). The key difference is the existence of $n$ markets.[3] However, since we normalize the cost of effort by multiplying by $n$, we obtain the same equilibrium. In addition, we let $\beta \in [0, 1]$, while the authors assume perfect R&D spillovers among firms, i.e., $\beta = 1$.

### 1.4.1 Market Outcome

In the last stage of the game, given network $g$ and the R&D effort levels $\{e_i(g)\}_{i\in N}$, firms compete by choosing quantities

$$\underset{q_i^j}{Max}\ \Pi_i(g) = \sum_{j\in N}\left[a - q_i^j(g) - \sum_{m\neq i} q_m^j(g) - c_i(g)\right] q_i^j(g) - \gamma n e_i^2(g)$$

The (Cournot) equilibrium output in each market is given by

$$q_i^j = \frac{a - c_i(g) + \sum\limits_{m\neq i}(c_m(g) - c_i(g))}{(n+1)},$$

and the equilibrium profit of the Cournot competitors in the last stage is

$$\Pi_i(g) = \sum_{j\in N}\left[\frac{a - c_i(g) + \sum\limits_{m\neq i}(c_m(g) - c_i(g))}{(n+1)}\right]^2 - \gamma n e_i^2(g).$$

Now, let us denote the different groups of firms as follows:

(i) firm $i$

(ii) $k$ firms linked to firm $i$ denoted by $l$

(iii) $n - k - 1$ firms not linked to $i$, which we represent with $m$.

3 Goyal and Moraga Gonzalez (2001) also consider the independent market case in which firms collaborate in R&D and maximize monopoly profits in the last stage.

 DOI: 10.2870/63947

Therefore, we have

$$\begin{aligned} c_i(g) &= \bar{c} - e_i - \beta k e_l \\ c_l(g) &= \bar{c} - e_l - \beta e_i - \beta \sum_{\substack{j \neq i \\ j \in N_l}} e_j \\ c_m(g) &= \bar{c} - e_m - \beta \sum_{j \in N_m} e_j \end{aligned}$$

At the second stage of the game, firm $i$ will have the following profit function:

$$\begin{aligned} \Pi_i(g) &= n \frac{\left[ a - c_i + \sum_{l \in N_i} (c_l - c_i) + \sum_{m \notin N_i} (c_m - c_i) \right]^2}{(n+1)^2} - \gamma n e_i^2 \\ &= n \frac{[a - c_i + k(c_l - c_i) + (n-k-1)(c_m - c_i)]^2}{(n+1)^2} - \gamma n e_i^2 \\ &= n \frac{[a - c + (n - \beta k) e_i + \beta n k e_l - (1 + \beta(k-1)) k e_l - (1+\beta k)(n-k-1) e_m)]^2}{(n+1)^2} \\ &\quad -\gamma n e_i^2. \end{aligned}$$

The first order condition is:[4]

$$\begin{aligned} \frac{\partial \Pi_i(g)}{\partial e_i} &= \frac{n(n-\beta k)(a - \bar{c} + e_i)}{(n+1)^2} - \frac{n(n-\beta k)\beta k e_i}{(n+1)^2} \\ &\quad + \frac{n(n-\beta k) n \beta k e_l}{(n+1)^2} + \frac{n(n-\beta k) k (e_i - (1 + \beta(k-1)) e_l)}{(n+1)^2} \\ &\quad + \frac{n(n-\beta k)(n-k-1)(e_i - (1+\beta k) e_m)}{(n+1)^2} - \gamma n e_i = 0 \end{aligned}$$

The expression above reveals the different effects that govern the optimal R&D levels. The first one could be labeled (as it is in the literature) the *appropriability effect*: the larger the demand (or net demand, $a - \bar{c}$), the stronger is the incentive to invest in R&D. In Motta (2004), where R&D collaboration is not allowed, i.e., $k = 0$, this term decreases with $n$. In Goyal and

[4] The second order condition requires $\gamma > \frac{n^2}{(n+1)^2}$.

Moraga-Gonzalez (2001), this term becomes $\frac{(n-k)(a-\bar{c}+e_i)}{(n+1)^2}$ and it is also highest when $n = 1$, that is, when there is a monopoly.

The second term measures the spillovers received by the collaborators of firm $i$. Consequently, it is negative and disappears for $k = 0$. When choosing its R&D effort, firm $i$ realizes that it is making its collaborators tougher, an effect that is captured by this term.

The third term can be thought as the positive spillovers that the firm receives from its collaborators. When the firm has no links $(k = 0)$, as in Motta (2004), it disappears completely. This term causes the effort of the linked firms to be strategic complements since the second derivative with respect to $e_l$ is positive.

The fourth term captures the competition effect among collaborators. Together with the third term, it determines whether the collaborators' efforts are strategic complements or substitutes. We obtain

$$\frac{\partial^2 \Pi_i(g)}{\partial e_i \partial e_l} = \frac{2n(n-\beta k)k(\beta(n-k+1)-1)}{(n+1)^2}$$

In Goyal and Moraga-Gonzalez (2001), i.e., when $\beta = 1$, the efforts of linked firms are strategic complements. Here, the efforts can become strategic substitutes if $\beta < 1/(n-k+1)$. Hence, for low R&D spillovers, and in small and dense networks, the efforts of collaborators become strategic substitutes.

The fifth term is the pure competition effect among unlinked firms (thus, it will disappear for the complete network, i.e., when $k = n - 1$). When there is a monopoly ($n = 1$, and thus $k = 0$), the term also disappears. We observe that the efforts of unlinked firms are always strategic substitutes since

$$\frac{\partial^2 \Pi_i(g)}{\partial e_i \partial e_m} = -\frac{2n(n-\beta k)(n-k-1)(\beta k+1)}{(n+1)^2} < 0$$

Finally, the last term captures the marginal cost of R&D, and depends on the efficiency parameter $\gamma$, and $n$.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

### 1.4.2 R&D Efforts

Focusing on symmetric equilibrium, i.e., $e_i = e_l = e_m = e(g^k)$, we obtain the following equilibrium effort level for each firm:

$$e(g^k) = \frac{(a - \bar{c})(n - \beta k)}{\gamma(n+1)^2 - (n - \beta k)(1 + \beta k)}$$

The first thing to analyze is how the equilibrium level of R&D effort changes with the level of collaborative activity $k$.

**Proposition 1** *In R&D collaboration networks with $\beta > 0$, $e(g^k)$ decreases with $k$.*
(see the Appendix for the proofs.)

We observe that the marginal return of R&D effort is declining in the level of collaborative activity. The R&D effort of an individual firm decreases own production costs but also decreases neighbors', making them tougher competitors. In addition, an increase in $k$ implies that all firms have more collaborators operating at lower costs, which also reduces the returns to cost reduction for unlinked firms. This result is same as in Goyal and Moraga-Gonzalez (2001).

In addition, we can obtain the change in the total level of R&D in the industry with respect to the number of firms

$$\frac{\partial R(g^k)}{\partial n} = \frac{(a - \bar{c})[\gamma(n+1)[2n + \beta k(n-1)] - (n - \beta k)^2(1 + \beta k)]}{[\gamma(n+1)^2 - (n - \beta k)(1 + \beta k)]^2} > 0$$

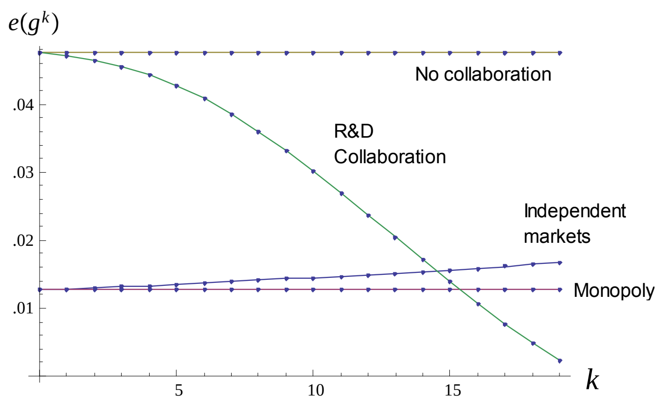
This expression shows that the amount of R&D investment is larger as the number of firms in the industry increases. However, the R&D level that is optimal for social welfare might be reached with a smaller number of firms. Therefore, we will study the welfare implications regarding the number of firms in the following sections.

Figure 4 below summarizes equilibrium R&D effort levels and the level of collaborative activity for $a - \bar{c} = 1$ and $\gamma = 1$, for different theoretical models. When firms invest in R&D in the absence of collaboration links, i.e., when $k = 0$, one can obtain "monopoly" or "no collaboration" equilibrium efforts for certain value of $n$. We observe that equilibrium effort levels are highest when firms compete in $n$ markets in the last stage (as in Motta 2004), when $n = 1$ (monopoly) individual R&D efforts are lower. When firms undertake R&D collaboratively and operate as monopolies in their home markets, we obtain the "independent market"

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

case studied by Goyal and Moraga-Gonzalez (2001). We observe that in this case the equilibrium effort is increasing in the level of collaborative activity since firms are not concerned about making their collaborators tougher competitors and their efforts are strategic complements. In our model, firms collaborate on R&D and operate as oligopolies in $n$ markets, which leads to declining equilibrium effort levels that can be even lower than the level of effort exerted by a monopoly for high levels of collaboration.

*Figure 4. Equilibrium Effort Levels for* $n = 20$, $\beta = 1$




Finally, we calculate the effect of the level of R&D spillovers on equilibrium effort

$$\frac{\partial e(g^k)}{\partial \beta} = -\frac{(a-\bar{c})k[\gamma(n+1)^2 - (n-\beta k)^2]}{[\gamma(n+1)^2 - (n-\beta k)(1+\beta k)]^2} < 0$$

We observe that equilibrium effort decreases in the level of spillovers. As $\beta$ goes to zero, the equilibrium effort approaches the "no collaboration" level as in Figure 4 above.

### 1.4.3 Cost Reduction

Substituting the equilibrium effort levels into the cost functions, we obtain the equilibrium level of production costs

$$c(g^k) = \frac{\bar{c}\gamma(n+1)^2 - a(n-\beta k)(1+\beta k)}{\gamma(n+1)^2 - (n-\beta k)(1+\beta k)}$$

Analyzing the change in costs with respect to the collaborative activity, $k$, we can state the following result:

**Proposition 2** *In R&D collaboration networks, $c(g^k)$ decreases in the number of collaborations for $k < \dfrac{n-\beta-1}{2\beta}$.*

For $\beta = 1$, as in Goyal and Moraga-Gonzalez (2001), costs are minimized when each firm is linked to almost half of the other firms. For lower levels of spillovers, more collaborations are required to attain the minimum cost level. When we substitute $k = n - 2$, we find that for $\beta < \frac{n-1}{2n-3}$, the equilibrium cost is always decreasing in the level of collaborative activity. The non-monotonic relationship between cost reduction and the collaborative activity stems from the fact that, as the network becomes relatively dense, the cost-reducing benefit from an additional collaboration link is offset by the detrimental effects arising from the induced decrease in R&D effort (Proposition 1). For a sufficiently low level of spillovers, we observe that the equilibrium level of R&D effort is higher, and that the latter effect disappears.

### 1.4.4 Equilibrium Profits

Finally, substituting the equilibrium level of effort and costs, we obtain the profit of a firm in a symmetric network

$$\Pi_i(g^k) = \frac{(a-\bar{c})^2\gamma n[\gamma(n+1)^2-(n-\beta k)^2]}{[\gamma(n+1)^2-(n-\beta k)(1+\beta k)]^2}$$

**Proposition 3** *In R&D collaboration networks, there exists $\bar{\beta} > 0.5$ such that if $\beta > \bar{\beta}$, there exists an intermediate level of collaborative activity, $0 < \hat{k} < n-1$, at which firms' profits are maximized.*

In order to analyze the effect of the level of spillovers, let us assume that $n$ and $k$ are continuous variables. The level of collaborative activity, $\hat{k}$, that maximizes $\Pi_i(g^k)$ satisfies:

$$\gamma(1+n)^2(2n-3\beta\hat{k}-1)-(n-\beta\hat{k})^3 = 0$$

In addition to the results that $d\hat{k}/dn > 0$ and $d\hat{k}/d\gamma > 0$ as in Goyal and Moraga-Gonzalez (2001), we find that $d\hat{k}/d\beta < 0$, which suggests that firms wish to collaborate less in environments where R&D spillovers are high.

 DOI: 10.2870/63947

### 1.4.5 Welfare

We now consider the social welfare aspects of R&D collaboration agreements. Using the optimal effort and the cost structure, social welfare can be written

$$\begin{aligned} W(g^k) &= \frac{1}{2}n\left[\frac{na-c_i-kc_l-(n-k-1)c_m}{n+1}\right]^2+n\Pi_i(g) \\ &= \frac{(a-\bar{c})^2\gamma n^2[\gamma(n+1)^2(n+2)-2(n-\beta k)^2]}{2[\gamma(n+1)^2-(n-\beta k)(1+\beta k)]^2} \end{aligned}$$

**Proposition 4** *In R&D collaboration networks, the empty network is not efficient. There exists a $\breve{\beta} > 0.5$ such that, if $\beta > \breve{\beta}$ and $\gamma = 1$, there exists an intermediate level of collaborative activity $\tilde{k}$ with $0 < \tilde{k} < \hat{k} < n-1$ for which the social welfare is maximized. When $\beta \leq 0.5$ and $\gamma = 1$, social welfare increases with the level of collaborative activity.*

We observe that profit maximizing level of collaborations is excessive from a welfare viewpoint, as was the case in Goyal and Moraga-Gonzalez (2001).
In order to analyze the effect of the level of spillovers, let us assume that $n$ and $k$ are continuous variables. The level of collaborative activity, $\tilde{k}$, that maximizes $W(g^k)$ satisfies:

$$\gamma(1+n)^2[(n+3)(n-2\beta\tilde{k})-2]-2(n-\beta\tilde{k})^3=0$$

We find that $d\hat{k}/dn > 0$ and $d\hat{k}/d\gamma > 0$ as in Goyal and Moraga-Gonzalez (2001). Moreover, we observe that $d\hat{k}/d\beta < 0$, which implies that in the environments where R&D spillovers are high, less collaborations are needed to maximize social welfare.

### 1.4.6 Stability

In order to analyze the pairwise stability of the complete and empty networks, we compute the deviation profits $\Pi(g^0+g_{ij})-\Pi(g^0)$ and $\Pi(g^{n-1}-g_{ij})-\Pi(g^{n-1})$. We can then state the following:

**Proposition 5** *Consider R&D collaboration networks; the empty network is not pairwise stable, and the complete network is stable.*


DOI: 10.2870/63947

Since, we have a similar framework, our results for R&D collaboration networks with high level of spillovers are in line with those of Goyal and Moraga-Gonzalez (2001). We have shown that when R&D collaborators are competitors in the homogeneous-product market, the empty network is not stable. The complete network is stable, but is not socially optimal. Total R&D is highest in the empty network, but efficiency requires an intermediate level of collaborative activity for a sufficiently high level of spillovers.

## 1.5 Benchmark II: R&D in Collusive Networks

In this section, we outline the second benchmark model in which firms have the opportunity to form collusive links agreeing that they do not enter each others' markets. Firms invest in R&D but, since they do not have any R&D collaboration agreements, firms' R&D efforts only reduce their own costs of production, and don't benefit their competitors. We analyze the incentives of firms to invest in R&D when they have the option to enter into collusive market sharing agreements.

### 1.5.1 Market Outcome

The total profits of firm $i$ are given by the sum of the profits in the home market and the foreign markets in which it operates.

$$\Pi_i(g) = \pi_i^i(g) + \sum_{\substack{j \in N \\ g_{ij} = 0}} \left[ \frac{a - (n - \eta_j(g))c_i + c_j + \sum\limits_{\substack{m \in N, m \neq i \\ g_{jm}=0}} c_m}{n - \eta_j(g) + 1} \right]^2 - \gamma n e_i^2(g)$$

$$\text{where } \pi_i^i(g) = \left[ \frac{a - (n - \eta_i(g))c_i + \sum\limits_{\substack{j \in N, j \neq i \\ g_{ij}=0}} c_j}{n - \eta_i(g) + 1} \right]^2$$

We focus on regular networks, i.e., $\eta_i(g) = \eta_j(g) = k$. Let $l$ denote the firms that are not linked to firm $i$ by collusive agreements, and let $m$ represent the firms that are not linked to $l$. The firms denoted by $m$ can be neighbors of firm $i$. We can then write the profit function as:

$$\begin{aligned}\underset{e_i}{Max}\ \Pi_i(g^k) &= \left[\frac{a-(n-k)c_i+(n-k-1)c_l}{n-k+1}\right]^2 \\ &+(n-k-1)\left[\frac{a-(n-k)c_i+c_m+(n-k-2)c_l}{n-k+1}\right]^2 - \gamma n e_i^2\end{aligned}$$

where (in the absence of spillovers)

$$\begin{aligned}c_i &= \bar{c} - e_i \\ c_l &= \bar{c} - e_l \\ c_m &= \bar{c} - e_m\end{aligned}$$

We observe that in the complete network, when $k = n-1$ and all firms form market sharing agreements, the profit function becomes identical to the one of a monopoly. In the empty network, when $k = 0$, firms are competing à la Cournot in $n$ markets.

### 1.5.2 R&D Efforts

Firms choose their R&D efforts $\{e_i(g)\}_{i\in N}$ to maximize $\Pi_i(g^k)$.

$$\begin{aligned}\Pi_i(g^k) &= \left[\frac{a-c+e_i+(n-k-1)(e_i-e_l)}{n-k+1}\right]^2 \\ &+(n-k-1)\left[\frac{a-c+e_i+(e_i-e_l)+(n-k-2)(e_i-e_m)}{n-k+1}\right]^2 - \gamma n e_i^2\end{aligned}$$

The first order condition is:

$$\begin{aligned}\frac{\partial \Pi_i(g^k)}{\partial e_i} &= \frac{(n-k)(a-\bar{c}+e_i)}{(n-k+1)^2} + \frac{(n-k)(n-k-1)(e_i-e_l)}{(n-k+1)^2} \\ &+\frac{(n-k)(n-k-1)(a-\bar{c}+e_i)}{(n-k+1)^2} + \frac{(n-k)(n-k-1)(e_i-e_l)}{(n-k+1)^2} \\ &+ \frac{(n-k)(n-k-1)(n-k-2)(e_i-e_m)}{(n-k+1)^2} - \gamma n e_i = 0\end{aligned}$$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

The expression above reveals the different effects in home and foreign markets that govern the optimal R&D levels. The first term represents the appropriability effect, and relates to the home market, and is decreasing in $n$. It is highest when $n = 1$ ($k = 0$ in this case) and whenever $k = n - 1$, that is, when the firm is a monopoly in its home market. The second term is the competition effect, and relates to the home market. When $k = n - 1$, this term disappears completely. The third term is the appropriability effect that relates to the foreign markets. It is increasing in $n$ (scale effect). The fourth term is the competition effect that relates to the foreign markets that contain firms with whom the firm competes in its home market. Notice that it is same as the competition effect in home market. The fifth term captures the effects from competition with firms with whom firm $i$ has a collusive agreement. When we take $e_l = e_m = \bar{e}$ and take the second derivative, we observe that

$$\frac{\partial^2 \Pi_i(g^k)}{\partial e_i \partial e_m} = -\frac{2(n-k)^2(n-k+1)}{(n-k+1)^2} < 0$$

so that the R&D efforts of firms are strategic substitutes.

### 1.5.3 Equilibrium Effort Level

Focusing on symmetric equilibrium, in which $e_i = e_l = e_m = e(g^k)$, we obtain the following equilibrium effort level for each firm:

$$e(g^k) = \frac{(a-\bar{c})(n-k)^2}{\gamma n(n-k+1)^2 - (n-k)^2}$$

The first thing to analyze is how the equilibrium level of R&D effort changes with the level of collusive activity $k$.

**Proposition 6** *Consider R&D in collusive networks. $e(g^k)$ decreases in $k$.*
(see the Appendix for the proofs.)

We have obtained the same effect as in the previous model, i.e., the equilibrium effort declines in the degree of the network in both models. In the previous model, where links define the R&D collaborations, this is due to the strategic behaviour of the firms, which take into account the spillovers and anticipate the market stage. Here, an increase in the connectivity enables firms to enjoy more market power, giving them less incentive to invest in R&D.

*Figure 5(a) Equilibrium Effort Levels for* $n = 20, \beta = 1$

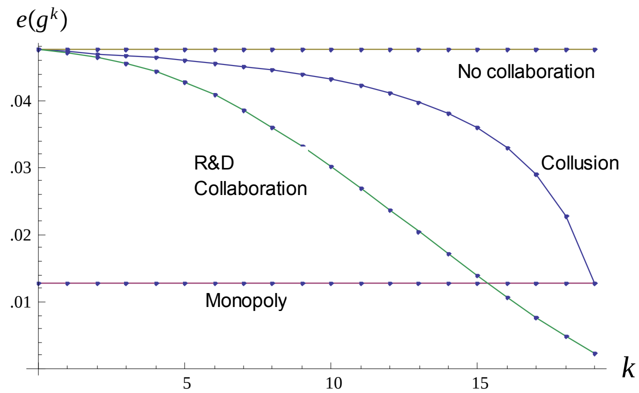



Figure 5(a) above illustrates the equilibrium R&D levels and the degree of the network for $a - \bar{c} = 1$ and $\gamma = 1$. Note that, when $k = 0$ (in the empty network), the equilibrium R&D effort is the same in all three models: R&D collaboration with $\beta = 1$, no collaboration ($\beta = 0$), and R&D in collusive networks. In addition, the model with collusive links approaches the monopoly case as the network gets denser. Under the complete network, in which $k = n - 1$, all firms are local monopolies in their home markets and the equilibrium R&D efforts in both models are the same.

In the two cases where the links are interpreted as R&D collaboration or as collusive market sharing agreements, equilibrium R&D effort is declining in the degree of the network. In the former model, the decline in the R&D effort with respect to the collaborative activity is due to a business stealing effect that causes firms to reduce their individual R&D efforts. The decreasing R&D in collusive networks is the result of increased monopoly power. Finally, we observe that for high levels of spillovers, the model with collaboration links generates lower R&D effort levels than the collusive networks model. As the level of spillovers gets smaller, the equilibrium effort level in the R&D collaboration network increases and approaches the no collaboration ($\beta = 0$) level. Depending on the level of spillovers, $\beta$, R&D efforts can be higher or lower in collusive networks.

### 1.5.4 Cost Reduction

Substituting the equilibrium effort levels into the cost functions, we obtain

$$c(g^k) = \frac{\bar{c}\gamma(n-k+1)^2 - a(n-k)^2}{\gamma n(n-k+1)^2 - (n-k)^2}$$

We analyze the change in costs with collusive activity, $k$, and show that $c(g^k) - c(g^{k+1}) < 0$. We state the following:

**Proposition 7** *Consider R&D in collusive networks.* $c(g^k)$ *increases in* $k$.

This result stems from the fact that an increase in number of market sharing agreements decreases the R&D effort of firms (Proposition 6) leading to higher costs of production.

### 1.5.5 Equilibrium Profits

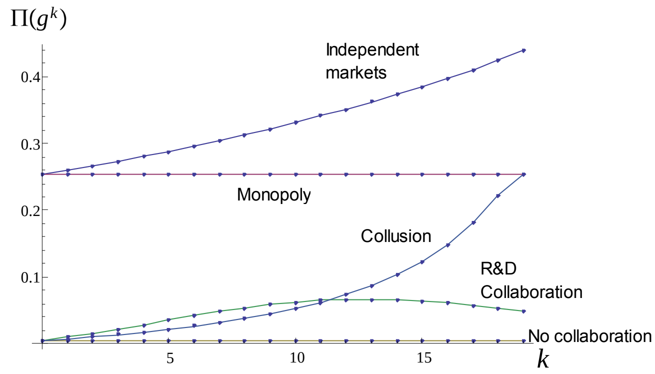
Finally, substituting in the equilibrium level of effort and costs, we obtain the profit of a firm in a symmetric network

$$\Pi(g^k) = \frac{(a-\bar{c})^2\gamma n(n-k)[\gamma n(n-k+1)^2 - (n-k)^3]}{[\gamma n(n-k+1)^2 - (n-k)^2]^2}$$

**Proposition 8** *Consider R&D in collusive networks.* $\Pi(g^k)$ *increases in* $k$.

Figure 5(b) below shows that the complete network maximizes the industry profits of firms undertaking individual R&D in collusive networks when there are no R&D collaborations. As the number of collusive links increases, firms decrease their R&D efforts, leading to increasing cost levels. However, this effect is offset by the increased monopoly power of firms, leading to increased profits. When firms collaborate on R&D and operate in independent markets, profits are also increasing in the level of collaborative activity. The resulting profits are higher than in the other models. In this case, firms benefit from their collaborators' R&D efforts and operate as monopolies in their home markets. This leads to even higher profits than the level in the monopoly case without R&D collaboration.

*Figure 5(b) Industry Profits for* $n = 20$, $\beta = 1$




### 1.5.6 Welfare

Now let us consider the social welfare aspects of market sharing agreements. Social welfare can be written

$$\begin{aligned} W(g^k) &= \sum_{j=1}^{N} \frac{Q^j(g^k)^2}{2} + \sum_{i=1}^{N} \Pi_i(g^k) \\ &= \frac{1}{2} n \left[ \frac{(n-k)a - c_i + (n-k-1)c_m}{n-k+1} \right]^2 + n\Pi_i(g^k) \end{aligned}$$

Using the optimal effort and the costs, we obtain equilibrium welfare

$$W(g^k) = \frac{(a-\bar{c})^2 \gamma n^2 (n-k)[\gamma n (n-k+1)^2 (n-k+2) - 2(n-k)^3]}{2[\gamma n (n-k+1)^2 - (n-k)^2]^2}$$

**Proposition 9** *Consider R&D in collusive networks. The complete network is never efficient and the empty network is efficient for* $\gamma = 1$, $n \leq 6$. *When* $n \geq 7$, *social welfare is maximized at intermediate levels of collusive activity.*

Goyal and Moraga-Gonzalez (2003) suggest that the empty network maximizes social welfare when firms collude by forming market sharing agreements, but don't form R&D collaborations. We have found that if the colluding firms also invest in R&D, an intermediate level of collusion is socially optimal when there is a sufficiently high number of firms (markets).

*Figure 5(c) Social Welfare for* $n = 20$, $\beta = 1$

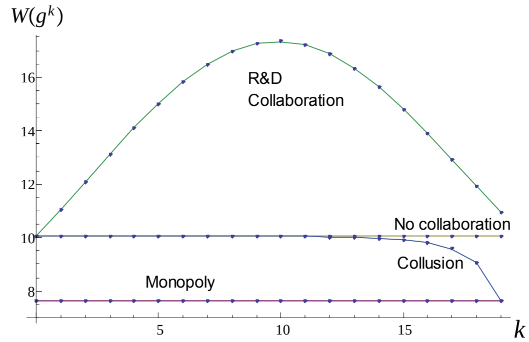



Figure 5(c) illustrates the social welfare results for the models studied so far. We can see that the highest social welfare is obtained when firms collaborate in R&D and compete in the homogeneous-product market. On the other hand, in absence of collaboration and competition, which is the "monopoly" case, social welfare is lowest. When firms invest in R&D in collusive networks, welfare is non-monotonic in the number of market sharing agreements, and it approaches the monopoly level as the network becomes complete. "No collaboration" is the case in which all firms compete but do not collaborate in R&D. This results in higher social welfare than in the monopoly case, but still not efficient compared to the R&D collaboration case because industry profits are very low when there is no R&D collaboration.

### 1.5.7 Stability

In order to analyze the pairwise stability of the complete and empty networks, we compute the deviation profits $\Pi(g^0 + g_{ij}) - \Pi(g^0)$ and $\Pi(g^{n-1} - g_{ij}) - \Pi(g^{n-1})$.

**Proposition 10** *Consider R&D in collusive networks. The complete network is pairwise stable and the empty network is stable for* $\gamma = 1$, $2 < n < 6$.

To sum up, we have shown that when firms undertake individual R&D and collude in the market stage by forming market sharing agreements, the complete network, which is pairwise stable, maximizes industry profits, but is not socially optimal. Efficiency requires an intermediate level of collusive activity when the number of markets (firms) is sufficiently high. Total R&D is the highest in the empty network, which is also efficient and pairwise stable for $2 < n < 6$.

## 1.6 R&D Collaboration in Collusive Networks

In this section, we present the model in which links are interpreted as both R&D collaborations and market sharing agreements. The market sharing agreement between two collaborators restrains these firms from entering each others' markets, however, they can compete in a third market when the local firm of that market is not their common neighbor. We analyze the incentives that firms have to reduce their costs with R&D research when there are both spillovers to other firms, and opportunities for market sharing agreements.

### 1.6.1 Why does the Configuration Matter?

As discussed before, we now show how different configurations of regular networks with same size and degree will have different equilibrium outcomes. To see this, we take the profit function coming from the last (competition) stage of the game:

$$\Pi_i(g) = \pi_i^i(g) + \sum_{\substack{j \in N \\ g_{ij}=0}} \left[ \frac{a-(n-k)c_i + \sum\limits_{\substack{m\in N, m\neq i \\ g_{jm}=0}} c_m}{n-k+1} \right]^2 - \gamma n e_i^2(g)$$

$$\text{where } \pi_i^i(g) = \left[ \frac{a-(n-k)c_i + \sum\limits_{\substack{j\in N, j\neq i \\ g_{ij}=0}} c_j}{n-k+1} \right]^2$$

Focusing on regular networks in which $\eta_i(g) = \eta_j(g) = k$, and letting $l$ denote the firms that are linked to firm $i$, and $(j, m)$ denote the firms that are not adjacent to $i$, we can write the profit function as:

$$\begin{aligned} \underset{e_i}{Max}\ \Pi_i(g) &= \left[\frac{a-(n-k)c_i+(n-k-1)c_m}{n-k+1}\right]^2 \\ &\quad + \sum_{\substack{r_{i,j} \\ j\in N \\ g_{ij}=0}} \left[\frac{a-(n-k)c_i + r_{i,j}c_l + (n-k-r_{i,j}-1)c_m}{n-k+1}\right]^2 - \gamma n e_i^2 \end{aligned}$$

The costs can be written as

$$\begin{aligned} c_i(g) &= \bar{c} - e_i - \beta k e_l \\ c_l(g) &= \bar{c} - e_l - \beta e_i - \beta \sum_{j \in N_l \setminus i} e_j \\ c_m(g) &= \bar{c} - e_m - \beta \sum_{j \in N_m} e_j \end{aligned}$$

where $r_{i,j}$ is the number of collaborators that firm $i$ faces in foreign market $j$. In other words, it is the number of neighbors of $i$ that are not common with firm $j$, where $g_{ij} = 0$. Using the definition of strongly regular graphs, we can write $r_{i,j} = k - \mu_{i,j}$, where $\mu_{i,j}$ is the number of common neighbors of firm $i$ and $j$. Intuitively, if $i$ has $k$ neighbors denoted by $l$, and $\mu_{i,j}$ of them cannot enter market $j$ due to the market sharing agreement between $j$ and $l$, firm $i$ will compete with $k - \mu_{i,j}$ neighbors in market $j$. Note that, depending on the configuration of the network, the $r_{i,j}$'s will take different values in different foreign markets. Since firm $i$ never competes with its collaborators in the home market, let us focus on the sum of profits in the foreign markets. The profit function can be written

$$\Pi_i(g) = \pi_i^i(g) + \sum_{\substack{j \in N \\ g_{ij}=0}} \left[ \frac{a - \bar{c} + (n-k)(e_i + \beta \sum_{l \in N_i} e_l) - (e_j + \beta \sum_{j' \in N_j} e_{j'}) - (\sum_{\substack{m \in N \\ g_{jm}=0}} e_m + \beta \sum_{m' \in N_m} e_{m'})}{(n - \eta_j + 1)} \right]^2$$

Hence, the firms operating in foreign market $j$ are: firm $i$, firm $j$ and the firms that are not linked to firm $j$, denoted by $m$. Since among those $m$, there might be neighbors of firm $i$, $e_i$ might appear in the the last term $\beta \sum_{m' \in N_m} e_{m'}$, which affects firm $i$'s R&D investment decision. We can write the partial derivative with respect to $e_i$ as follows:

$$\begin{aligned} \frac{\partial \Pi_i(g)}{\partial e_i} &= (A+1) \times \frac{2(n-k)[a - \bar{c} + (n-k)(e_i + \beta k e_l) - (n-k-1)(e_m + \beta k e_{m'}]}{(n-k+1)^2} \\ &\quad + B \times \frac{2(n-k-\beta k)[a - \bar{c} + (n-k)(e_i + \beta k e_l) - k(e_l + \beta e_i + \beta(k-1)e_{l'}) - (n-2k-1)(e_m + \beta k e_{m'}]}{(n-k+1)^2} \\ &\quad + C \times \frac{2(n-k-\beta r_{i,j}))[a - \bar{c} + (n-k)(e_i + \beta k e_l) - r_{i,j}(e_l + \beta e_i + \beta(k-1)e_{l'}) - (n-k-r_{i,j}-1)(e_m + \beta k e_{m'}]}{(n-k+1)^2} \end{aligned}$$

In the above expression, $A$ denotes the number of foreign markets for which $\mu_{i,j} = k$, meaning that firm $i$ does not face any of its collaborators. Since this is also the case for the home market, there will be $(A+1)$ markets with no collaborators. $B$ denotes the number of markets where $i$ competes with all of his collaborators, i.e., in which $\mu_{i,j} = 0$. Finally, $C$ denotes the number of markets in which a subset of firm $i$'s neighbors operate and $0 < \mu_{i,j} < k$. Here, for illustrative reasons, we assume that $\mu_{i,j}$ is the same in all of these foreign markets. As discussed before, networks can have different values of $\mu$ for different pairs of nodes, hence, we can have these three types of markets in a given network. Since firm $i$ operates is $n-k$ markets, we can see that $A+B+C = n-k-1$ for all networks.

Since we focus on symmetric equilibrium in which $e_i = e_l = e_m = e^*(g^k)$, the term becomes

$$\frac{\partial \Pi_i(g)}{\partial e_i} = \left[\frac{2\left[a-\bar{c}+(1+\beta k)e^*(g^k)\right]}{(n-k+1)^2}\right] \times [(A+1)(n-k)+B(n-k-\beta k)+C(n-k-\beta r_{i,j})]$$

This allows us to see how the configuration of the network is crucial in our model. Consider two different regular networks with same number of firms, $n$, and same degree $k$.

**Remark 1:** When we take $\beta = 0$, we obtain the expression for R&D in the collusive networks model (where links define the market sharing agreements). We observe that for $\beta = 0$, the term becomes $[(A+1)(n-k)+B(n-k-\beta k)+C(n-k-\beta r_{i,j})] = (A+B+C+1)(n-k) = (n-k)^2$. Therefore, the expression becomes independent of the parameters $(A, B, C)$. This implies that one obtains the same outcome for networks with different configurations as long as they have the same $(n, k)$ pair.

**Remark 2:** It is possible to return to the Goyal and Moraga-Gonzalez (2001) framework, where links define R&D collaborations, and collaborators compete in all markets, by substituting $A = 0$ and $C = 0$ into the expression. In this case, $B$ becomes the number of foreign markets in which firm $i$ operates, which equals $n-1$ for all possible networks when there are no market sharing agreements. Therefore, in this setup, it is also the case that different networks with the same $(n, k)$ result in the same equilibrium outcome.

In contrast, our model introduces an interaction between the configuration of the network and the different incentives generated under different regular networks with the same number of nodes and the same degree.

After simple calculations we can obtain the first order condition as follows

$$\frac{\partial \Pi_i(g)}{\partial e_i} = [(A+1)(n-k)+B(n-k-\beta k)+C(n-k-\beta r_{i,j})] \times \Phi(e^*(g^k)) - 2\gamma n e^*(g^k) = 0$$

where $\Phi(e^*(g^k)) = \left[\frac{2\left[a-\bar{c}+(1+\beta k)e^*(g^k)\right]}{(n-k+1)^2}\right]$.

$$\frac{\partial \Pi_i(g)}{\partial e_i} = \Phi(e^*(g^k)) \times [(A+B+C+1)(n-k) - \beta(Bk+Cr_{i,j})] - 2\gamma n e^*(g^k) = 0$$

$$\frac{\partial \Pi_i(g)}{\partial e_i} = \Phi(e^*(g^k)) \times [(n-k)^2 - \beta(Bk+Cr_{i,j})] - 2\gamma n e^*(g^k) = 0$$

Finally, we obtain the equilibrium effort

$$e^*(g^k) = \frac{(a-c)[(n-k)^2 - \beta(Bk+Cr_{i,j})]}{\gamma n(n-k+1)^2 - (1+\beta k)[(n-k)^2 - \beta(Bk+Cr_{i,j})]} \qquad (✠)$$

Roughly speaking, here we observe that as the number of markets in which a firm competes with all of its collaborators, i.e., $B$, increases, the equilibrium effort decreases. This is also true for $C$. In addition, we observe that as the number of collaborators, $r_{i,j}$, that a firm competes with in one market $j$ increases, the equilibrium effort level also declines.

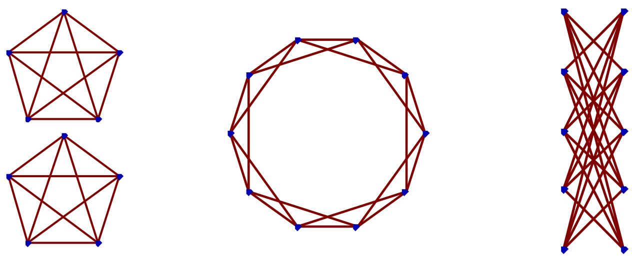
However, one cannot analyze the effect of connectivity on the equilibrium effort level by simply keeping other network parameters constant. First of all, different families of regular networks with the same $(n, k)$ have different values of $(\lambda, \mu)$ and $(A, B, C)$. More importantly, changes in $k$ (forming or severing links) will have different effects on these parameters depending on the particular network in question. Finally, even when we change $k$ in the same family of networks, the equilibrium outcomes will depend on the new configuration.

In order to make things more clear, Figure 6.1 below illustrates three different network configurations for regular graphs with $n = 10$, $k = 4$. First, note that in all these graphs, nodes are completely symmetric in terms of the number of neighbors that they have, the number of second degree neighbors they they have, and so on. This allows us to have the same first order conditions for all firms and to focus on symmetric equilibria. Second, remember that $A$ is the

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

number of markets for which $\mu_{i,j} = k$, meaning that firm $i$ does not face any of its collaborators; $B$ is the number of markets in which $i$ competes with all of its collaborators, i.e., where $\mu_{i,j} = 0$ and $C$ denotes the number of markets that a subset of firm $i$'s neighbors operate and $0 < \mu_{i,j} < k$ for these markets.

*Figure 6.1. Different Regular Network Configurations for* $n = 10$, $k = 4$




In the regular network with complete components, all adjacent (linked) pairs have $(k-1)$ common neighbors and all nonadjacent pairs have zero common neighbors. Therefore, $(n, k, \lambda, \mu) = (10, 4, 3, 0)$ and $(A, B, C) = (0, 5, 0)$. Thus, firms compete with all their collaborators in all of the foreign markets that they enter in the other component.

The second regular graph with the same $(n, k) = (10, 4)$, has different values of the parameters $(\lambda, \mu) = (\{1, 2\}, \{0, 1, 2\})$ and $(A, B, C) = (0, 1, 4)$. Note that $B = 1$ implies that there is one market, which is the furthest one (in graph distance terms), that has $\mu = 0$. Thus, the firm competes with all its collaborators in that market since $r = k - \mu = k$. In addition, there are two markets with $\mu = 2$, and two markets with $\mu = 1$, so it competes with $r = k - \mu = k - 1 = 3$ collaborators in two markets and $r = k - \mu = k - 2 = 2$ collaborators in the remaining two markets.

Finally, the crown graph, which is obtained by removing the horizontal links from the complete bipartite graph, has $(\lambda, \mu) = (0, \{0, 3\})$ and $(A, B, C) = (0, 1, 4)$. Adding back the horizontal links will generate $(n, k, \lambda, \mu) = (10, 5, 0, 5)$ and $(A, B, C) = (4, 0, 0)$, resulting in a very different outcome in which firms never compete with their collaborators in the foreign markets.


DOI: 10.2870/63947

As we have already discussed, the parameters of strongly regular graphs $(n, k, \lambda, \mu)$ are not independent. From graph theory, we know that $(n-k-1)\mu = k(k-\lambda-1)$. Here, we can see that $\lambda$ and $\mu$ are negatively correlated. When adjacent pairs have more common neighbors, nonadjacent pairs have less. We will use this fact to distinguish between some families of networks later on. In addition, the parameters of strong regularity $(\lambda, \mu)$ are also related to the parameters $(A, B, C)$ in our model.

For example, when $\mu_{i,j} < k \ \ \forall j$ with $g_{ij} = 0$, it must be that $A = 0$. If all of the markets that firm $i$ enters have some (less than $k$) common collaborators, then the number of market in which firm $i$ can operate without competing with any of its collaborators is zero. When $\mu_{i,j} = 0 \ \forall j$, it is again the case that $A = 0$, and since $B = n-k-1$ with $r = k$ and $\lambda = k-1$, $C = 0$. This gives us the regular graphs with complete components (see above) in which all firms in one component are linked, $\lambda = k-1$, and in which firms always complete with all their collaborators in the foreign markets since $r_{i,j} = k \ \ \forall j$. Finally, if $\mu_{i,j} = k \ \ \forall j$, then $r = 0$ (so that firms never compete with collaborators). It follows that $B = 0$, $C = 0$ and $A = n-k-1$. This is the well-known complete bipartite graph with $(\lambda, \mu) = (0, k)$. As we have seen before, bipartite graphs obtained by severing links from the complete bipartite graph result in different values of parameters $(\lambda, \mu)$ and $(A, B, C)$. This affects the equilibrium outcome.

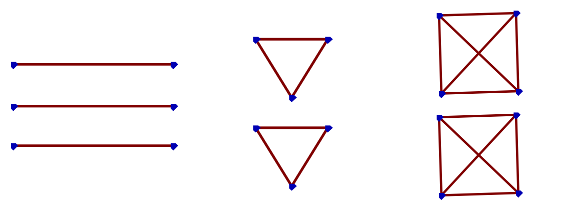
Therefore, in order to have comparative statics with respect to $k$, and to characterize a sufficient number of networks to illustrate how configuration matters, we will study a subset of these possible regular networks. In order to simplify the analysis and focus on symmetric equilibrium in a tractable, systematic way, we construct three different models for three different families of networks. We analyze regular networks with two or more complete components, ring lattices and regular bipartite graphs separately (see Figure 6.1 above). For each group of networks, we will calculate the equilibrium effort, industry profits, social welfare and in the last section, we will analyze the stability of these configurations.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

### 1.6.2 Regular Networks with Complete Components

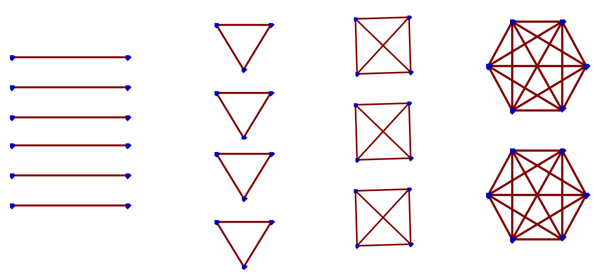
We model these graphs separately due to the fact that, since all firms in the same component are linked but have no links with the firms in the other components, all collaborators always compete with each other in all the foreign markets in the other components. Formally, we can state that regular networks with complete components are strongly regular with $(n, k, \lambda, \mu) = (n, k, k-1, 0)$ and $(A, B, C) = (0, n-k-1, 0)$. Figure 6.2(a) illustrates some examples of these networks for $n = 6$ and $n = 8$.

*Figure 6.2(a) Regular Networks with Complete Components for* $n = 6$, $n = 8$




Since we are analyzing graphs with complete components that are of the same size, note that the maximum size of the components is $n/2$, and the maximum degree of the network is $k = n/2 - 1$. Thus, we are dealing with relatively sparse networks in this section. In addition, using the degree, $k$, as the parameter for the comparative statics, an increase in degree (given the size of the network), implies a decrease in the number of the components while an increase in the size of each complete component. Below, Figure 6.2(b) illustrates the regular graphs with complete components for 12 firms as we alter the degree from $k = 1$ to $k = 5$.

*Figure 6.2(b) Regular Networks with Complete Components for* $n = 12$



**Market Outcome**

In this case, firm $i$ competes with all of its collaborators, $k$, in all of the foreign markets in which it operates, so that $r_{i,j} = k - \mu_{i,j} = k \ \ \forall j$ with $g_{ij} = 0$, and $B = n - k - 1$. Therefore, the profit function coming from the last (competition) stage of the game, imposing symmetry, can be written

$$\begin{aligned}\Pi_i(g^k) \quad &= \quad \left[\frac{a-(n-k)c_i+(n-k-1)c_m}{n-k+1}\right]^2 \\ &\quad +(n-k-1)\left[\frac{a-(n-k)c_i+kc_l+(n-2k-1)c_m}{n-k+1}\right]^2 - \gamma n e_i^2\end{aligned}$$

Here, an increase in $k$ will have the following mixed effects; a decrease in the level of competition in home and foreign markets due to market sharing agreements, a decrease in the number of foreign markets in which a firm operates (the scale effect), and an increase in the number of R&D collaborators for all firms. Since firms compete with their collaborators in all markets, they consider the possibility that they will make their competitors tougher when deciding whether to invest more in R&D. The number of markets in which firms compete with all collaborators also decreases with $k$.

**R&D Efforts**

Focusing on symmetric equilibria in which $e_i = e_j = e_m = e(g^k)$, we obtain the following equilibrium effort level for each firm:

$$e(g^k) = \frac{(a-c)[(n-k)^2 - \beta k(n-k-1)]}{\gamma n(n-k+1)^2 - (1+\beta k)[(n-k)^2 - \beta k(n-k-1)]}$$

Note that we can also obtain the expression for equilibrium effort by replacing $C = 0$, $B = n - k - 1$ and $r_{i,j} = k$ in the equation (✠) obtained before.

**Proposition 11** *Consider R&D collaboration in collusive networks and let $\beta = 1$. In regular networks with complete components, $e(g^k)$ decreases with $k$.*
(see the Appendix for proofs.)

In regular networks that consist of complete components, when links define both R&D collaborations and market sharing agreements, the equilibrium R&D effort decreases as the degree of the network increases. Given the size of the network, an increase in the degree and thus, a decrease in the number of complete components, has several effects in this model.

First of all, the number of firms in the home market and in each foreign market entered, and hence, the level of competition in each market, decreases. Second, the number of foreign markets in which an individual firm can operate decreases. As we have seen in the collusive networks model in the previous section, these forces together cause the equilibrium R&D effort level to decline as the degree of the network increases. In addition, with higher connectivity, the number of collaborators increases and, since firms in these networks always compete with all of their collaborators in the foreign markets, the number of collaborators that an individual firm competes also increases.

As analyzed in the first benchmark model of R&D collaboration, competition with collaborators reduces the equilibrium R&D level. Therefore, this model with complete components combines the parallel forces under the two models studied previously, resulting in declining equilibrium effort as suggested by both benchmark models. The equilibrium effort in this model is lower than the R&D effort obtained in either of the benchmark models.

**Equilibrium Profits**

Substituting in the equilibrium level of effort and costs, we obtain:

$$\Pi(g^k) = \frac{(a-c)^2\gamma n[\gamma n(n-k+1)^2(n-k) - [(n-k)^2 - \beta k(n-k-1)]^2]}{[\gamma n(n-k+1)^2 - (1+\beta k)[(n-k)^2 - \beta k(n-k-1)]]^2}$$

**Proposition 12** *Consider R&D collaboration in collusive networks and let $\beta = 1$. In regular networks with complete components, $\pi(g^k)$ increases with $k$ in the symmetric equilibrium.*

**Proof.** See the Appendix.

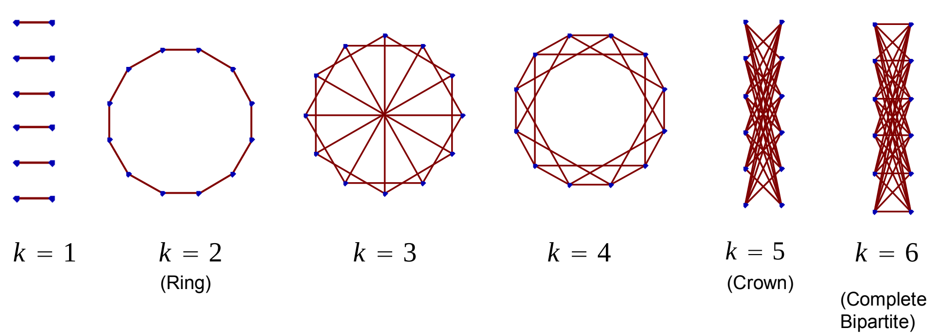
Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

### 1.6.3 Regular Bipartite Networks

In the previous section, we analyzed regular networks with complete components. These are an extreme case, in the sense that, in these networks, firms always compete with their collaborators. The other extreme case, in which collaborators never face each other in any market, occurs in a complete bipartite graph. A bipartite graph consists of a set of graph vertices decomposed into two disjoint sets such that no two graph vertices within the same set are adjacent. A *complete bipartite graph* is a bipartite graph such that every pair of graph vertices in these two sets are adjacent. It is also a strongly regular graph with $(n, k, \lambda, \mu) = (n, k, 0, k)$ and $(A, B, C) = (n - k - 1, 0, 0)$, where $k = n/2$.

In order to study the family of regular bipartite networks, we start with a complete-bipartite graph and decrease its degree by removing the edges. This leaves a $(k - 1)$-regular graph that is also bipartite. For example, removing all the horizontal edges in the first step generates the so-called *crown graph* (see Figure 6.3(a) below). We continue removing the edges one by one until we reach the empty graph. All bipartite graphs are undirected graphs in which no three vertices form a triangle of edges. In our framework, two firms with a collaboration and a collusive agreement do not have a common neighbor, i.e., $\lambda = 0$. Figure 6.3(a) illustrates some examples of these networks for $n = 12$.

*Figure 6.3(a) Regular Bipartite Networks for* $n = 12$

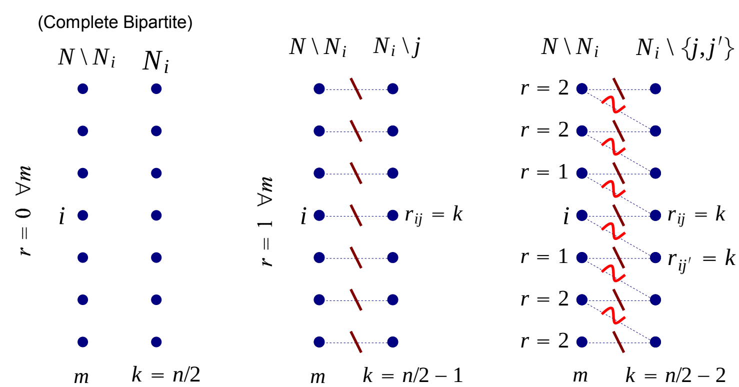



Note that one can construct triangle-free networks only for $k \leq n/2$. Consequently, we study relatively sparse graphs with $k \leq n/2$ as we did in the previous section. Moreover, in order to be able to obtain the complete bipartite graph as we increase $k$, we consider graphs with even number of vertices.

**Market Outcome**

In order to specify the profit function for regular bipartite networks, we again need to characterize the parameters $\lambda$, $\mu$, $A$, $B$ and $C$ for given values of $n$ and $k$. As we have mentioned before, $\lambda = 0$ for all bipartite graphs. However, as we decrease the degree by removing the edges of a complete bipartite graph, $\mu$ decreases. Therefore, $A$ and $C$ increase at the expense of $B$ since $A + B + C = n - k - 1$ must hold. We observe that the $\mu_{i,j}$'s, and consequently the $r_{i,j}$'s, change in a systematic way.

*Figure 6.3(b) Regular Bipartite Networks for* $n = 12$, $4 \leq k \leq 6$




For illustrative reasons, only the links which are being removed are shown in Figure 6.3(b). The first graph is a complete bipartite graph, which consists of two sets of nodes with $n/2$ firms in each set. There are no links between the firms in the same set, so firm $i$ enters all foreign markets in the set $N \backslash N_i$, which are denoted by $m$. However, firm $i$ is linked to all of the nodes in the other group. As we remove all of the horizontal edges ($k = n/2 - 1$), firm $i$ enters market $j$ in which all remaining neighbors of $i$ operate.In that market $j$, $r_{i,j} = k$. This will be the same for all markets in which firm $i$ severes its links. So there will always be $n/2 - k$ markets on the right-hand side, where firm $i$ competes with all of its remaining neighbors.

In addition, some of firm $i$'s neighbors also lose their links with firms in the same set as firm $i$. Therefore, firm $i$ starts to compete with those neighbors as well. In order to characterize the $r_{i,m}$'s, we check the links that any $m$ has lost, thus, those firms enter to market $m$. Then we check how many of those are neighbors of $i$, which will correspond to the $r_{i,m}$ for that market. As we go on removing the edges, some nodes in $N \backslash N_i$ (in this setup, the $m$'s that are

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

geometrically closer to firm $i$), end up severing links that $i$ has also lost. This implies that, as we move up from firm $i$, the $r_{i,m}$'s of the firms increase from $1$ to $n/2-k$. Due to symmetry, it is also the case for the firms below. The remaining firms (if they exist) have the maximum $r$, that is $r=n/2-k$. One can also construct bipartite networks starting from the empty graph and adding links, but this does not alter the results of the analysis. Finally, we can formulate the profit function of firm $i$ as follows:

$$\begin{aligned}
\Pi_i(g^k) &= \left[\frac{a-(n-k)c_i+(n-k-1)c_m}{n-k+1}\right]^2 \\
&+(n/2-k)\left[\frac{a-(n-k)c_i+kc_l+(n-2k-1)c_m}{n-k+1}\right]^2 \\
&+2\sum_{r=1}^{n/2-k}\left[\frac{a-(n-k)c_i+rc_l+(n-k-r-1)c_m}{n-k+1}\right]^2 \\
&+(4k-n-2)/2\left[\frac{a-(n-k)c_i+(n/2-k)c_l+(n/2-1)c_m}{n-k+1}\right]^2-\gamma n e_i^2
\end{aligned}$$

**R&D Efforts**

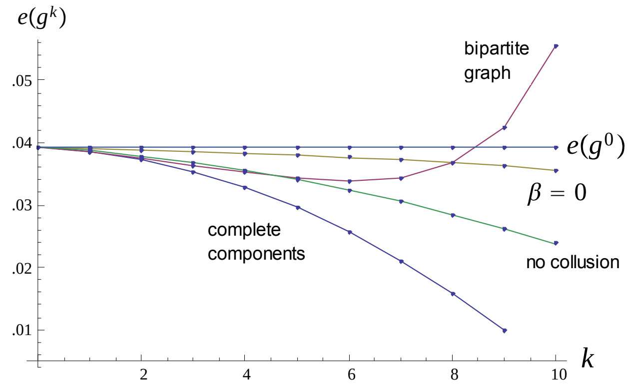
Focusing on symmetric equilibrium in which $e_i=e_j=e_m=e(g^k)$, we obtain the following equilibrium effort level for each firm:

$$e(g^k)=\frac{(a-c)[(n-k)^2-\beta k(n-2k)]}{\gamma n(n-k+1)^2-(1+\beta k)[(n-k)^2-\beta k(n-2k)]}$$

**Proposition 13** *Consider R&D collaboration in collusive networks and let $\beta=1$. In regular bipartite graphs with $k\leq n/2$, $e(g^k)$ is non-monotonic with respect to $k$. There exists a $\bar{k}$* with $0<\bar{k}<n/2$, *below which $e(g^k)$ decreases in $k$. For $k>\bar{k}$, $e(g^k)$ increases with $k$ and reaches a maximum at $k=n/2$, i.e., under the complete bipartite-graph.*

As we have seen in equation (✠), the configuration of the network (negatively) affects the equilibrium effort level by the term $\beta k(Bk+Cr_{i,j})$. In the regular networks with complete components, since $B=n-k-1$ and $C=0$, we have $\beta k(Bk+Cr_{i,j})=\beta k(n-k-1)$. The marginal effect of a change in $k$ on this term is $\beta(n-2k-1)$, which is positive since $k\leq n/2-1$ for these networks. On the other hand, in regular bipartite networks we have the term $\beta k(n-2k)$ which changes at the rate $\beta(n-4k)$ and is negative for $k>n/4$. This explains why equilibrium effort is non-monotonic in $k$, increasing after an initial decrease, in the regular bipartite graphs as opposed to the regular networks with complete components, where the effort declines as degree increases. In addition, when $k = n/2$ (in the complete bipartite network), the term disappears completely, resulting in the highest equilibrium effort level.

*Figure 6.3(c) Equilibrium Efforts for* $n = 20$, $\beta = 1$




As an example, Figure 6.3(c) illustrates the equilibrium effort levels for $k \leq n/2$ for all the models we have studied so far. We observe that when links are defined only by market sharing agreements, i.e., $\beta = 0$, (as in Belleflamme and Bloch, 2004) or only by R&D collaborations, i.e., *no collusion*, (as in Goyal and Moraga-Gonzalez, 2001), equilibrium effort is decreasing in the level of connectivity of the network. In these benchmark models, in which the degree is the only network parameter that changes the outcome, the highest equilibrium effort is attained in the empty network, i.e., $e(g^0)$. However, when links are interpreted as both R&D collaborations and market sharing agreements, different network configurations with the same degree have different equilibrium outcomes. In the networks with complete components, equilibrium effort decreases to a level that is lower than in all other other models. In the regular bipartite graphs it is nonmonotoic in $k$. The complete bipartite graph, $k = n/2$, results in the highest equilibrium effort, which is even higher than the level in the empty network. Intuitively, the driving force behind the non-monotonic relationship is that although competition with collaborators tends to decrease R&D effort, this effect is offset by the benefits that firms gain in cost reduction. This is because there are both fewer markets in which collaborators operate together, and fewer collaborators in each market. The complete bipartite graph is the extreme

 DOI: 10.2870/63947

case in which firms never compete with their collaborators. This corresponds to the independent market case studied by Goyal and Moraga-Gonzalez (2001).

**Equilibrium Profits**

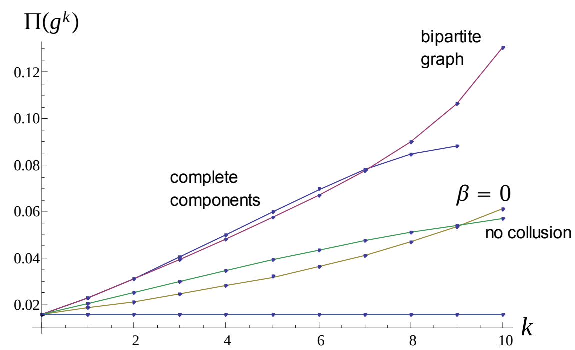
Finally, substituting the equilibrium level of effort and costs, we can obtain the profit of a firm

$$\Pi(g^k) = \frac{(a-c)^2\gamma n[\gamma(n-k+1)^2(n-k) - [(n-k)^2 - \beta k(n-2k)]^2]}{[\gamma(n-k+1)^2 - (1+\beta k)[(n-k)^2 - \beta k(n-2k)]]^2}$$

**Proposition 14** *Consider R&D collaboration in collusive networks. In regular bipartite graphs with $k \leq n/2$, $\Pi(g^k)$ increases in $k$.*

Figure 6.3(d) illustrates the equilibrium level of industry profits for the same parameter values as in the previous figures.

*Figure 6.3(d) Equilibrium Profits for $n = 20$, $\beta = 1$*

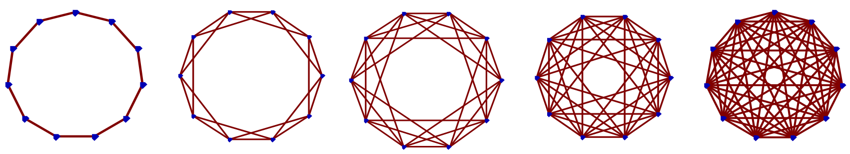



We observe that in all the models studied, industry profits are increasing in the level of connectivity for $k \leq n/2$. When links are defined only by R&D collaborations (*no collusion*), profits are non-monotonic in the level of collaborative activity and are maximized at an intermediate level of connectivity. When $\beta = 0$ (collusive networks), industry profits increase monotonically, which is also the case for regular bipartite graphs with both types of agreements. Equilibrium profits in the regular graphs with complete components follow the same pattern as the *no collusion* case since collaborators always compete with each other in both cases. However, it is higher in the former because market sharing agreements lead to increased market power in all markets. Finally, note that the complete bipartite graph results in the highest industry profits, while profits are lowest in the empty network.

### 1.6.4 Ring Lattices

In this section, we will analyze another family of graphs, named ring lattices, which exhibit high level of clustering. As we have discussed before, $k(k-\lambda-1) = \mu(n-k-1)$ implies that $\lambda$ and $\mu$, and consequently $r$ are negatively correlated. Hence, for networks with a high $\lambda$, the nonadjacent pairs have low $r_{i,j}$'s, which implies that a firm competes with more collaborators in the foreign markets than it did in the regular bipartite networks,

*Figure 6.4(a) Ring Lattices for* $n = 11$

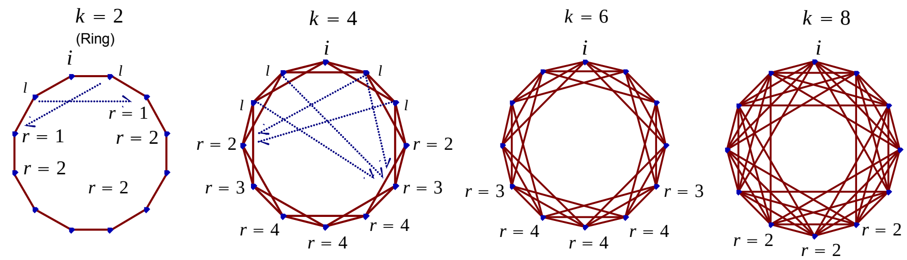



As we show in Figure 6.4(a), we construct these networks starting from a ring ($k = 2$) and increasing the degree of each node by adding links to the neighbors at distance-2. Because the new neighbors and the previous neighbors of $i$ become its common neighbors, adding links increases $\lambda_{ij}$. We keep on adding links until we obtain the complete network.

**Market Outcome**

Figure 6.4(b) illustrates how we characterize the $r_{i,j}$'s for ring lattices to obtain firm $i$'s profit function. Note that, in order to analyze these graphs systematically, we focus on networks with even degrees. In other words, in order to increase the degree of the network while keeping the network symmetrical, we add direct links to the neighbors at distance-2 in both sides. Second, when the network becomes dense, in particular when $k \geq \frac{2n-4}{3}$, all non-neighbors of firm $i$ will be linked. Therefore firm $i$ competes with only one non-neighbor which is the local firm of the foreign market it enters. To see that, remember there are $n-k-1$ foreign markets that firm $i$ enters. When $k/2 \geq n-k-2$, which becomes $k \geq \frac{2n-4}{3}$, all of these firms will be connected. In the example below, for $n = 12$ when $k \geq 8$, firm $i$ will always compete with $n-k-2$ neighbors in the foreign markets. Thus, for dense ring lattices, there is no need to restrict the analysis to networks with even degrees. However, in order to compare the results with the previous networks, we restrict attention to sparse ring lattices.

*Figure 6.4(b) Ring Lattices for* $n = 12, 2 \le k \le 8$




We can write the profit function for sparse ring lattices ($k < \frac{2n-4}{3}$) as follows:

$$\Pi_i^{sparse}(g^k) = \left[\frac{a-(n-k)c_i+(n-k-1)c_m}{n-k+1}\right]^2$$

$$+2\sum_{r=k/2}^{k-1}\left[\frac{a-(n-k)c_i+rc_l+(n-k-r-1)c_m}{n-k+1}\right]^2$$

$$+(n-2k-1)\left[\frac{a-(n-k)c_i+kc_l+(n-2k-1)c_m}{n-k+1}\right]^2-\gamma n e_i^2$$

**R&D Efforts**

Focusing on symmetric equilibrium in which $e_i = e_j = e_m = e(g^k)$, we obtain the following equilibrium effort levels:

$$e^s(g^k) = \frac{(a-c)[4(n-k)^2-\beta k(4n-5k-6)]}{4\gamma n(n-k+1)^2-(1+\beta k)[4(n-k)^2-\beta k(4n-5k-6)]}$$

**Proposition 15** *Consider R&D collaboration in collusive networks and let* $\beta = 1$*. In sparse ring lattices with* $k < \frac{2n-4}{3}$*,* $e(g^k)$ *decreases in* $k$ *in the symmetric equilibrium.*

For sparse ring lattices, as the degree of the network increases, the number of firms in the home market and in each foreign market entered decreases. The number of foreign markets that an individual firm can operate also decreases. As we have pointed out for networks with complete components, these forces together cause the equilibrium R&D effort level to decline as the degree of the network increases. However, in addition to the fact that higher connectivity implies more collaborators, firms in this model do not compete with all of their collaborators in the markets in which they operate. As a result of high level of clustering, firm $i$ competes with


DOI: 10.2870/63947

less of its collaborators in markets that are close (in graph distance terms) to its home market. Hence, firm $i$ competes with fewer of its collaborators in ring lattices than in regular graphs with complete components, where firm $i$ faces all of its collaborators in all markets it enters. This results in higher R&D levels. However, the equilibrium R&D effort is still lower than the levels that were found in the two benchmark models.

**Equilibrium Profits**

Substituting the equilibrium levels of effort and costs, we obtain the profit of a firm in the sparse ring lattices

$$\Pi^s(g^k) = \frac{(a-c)^2\gamma n[\gamma n(n-k+1)^2(n-k) - [4(n-k)^2 - \beta k(4n-5k-6)]^2]}{[4\gamma n(n-k+1)^2 - (1+\beta k)[4(n-k)^2 - \beta k(4n-5k-6)]]^2}$$

**Proposition 16** *Consider R&D collaboration in collusive networks and let $\beta = 1$. In sparse regular ring lattices with $k < \frac{2n-4}{3}$, $\pi(g^k)$ increases in $k$.*

In line with the previous analysis, we observe that the profit levels in sparse ring lattices are higher than the levels in networks with complete components, but lower than the equilibrium profits obtained in regular bipartite networks.

### 1.6.5 Welfare and Stability

We first consider the social welfare aspects of R&D collaboration and market sharing agreements for the families of networks studied. Social welfare can be written as:

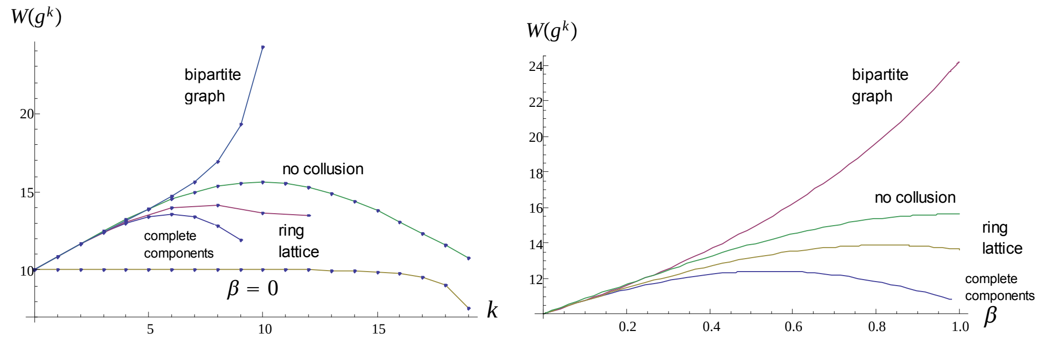
$$W(g) = \frac{1}{2}n\left[\frac{(n-k)a - c_i + (n-k-1)c_m}{n-k+1}\right]^2 + n\Pi_i(g)$$

Using the optimal effort and cost structures for complete components, bipartite graphs and ring lattices, we can obtain a value for equilibrium welfare and state the following result:

**Proposition 17** *Consider R&D collaboration in collusive networks and let $\beta = 1$. $W(g^k)$ increases in $k$ in regular bipartite graphs and is highest in the complete bipartite networks. In ring lattices and in networks with complete components, $W(g^k)$ is maximized at an intermediate level of connectivity, i.e., $k^* \approx n/3 - 1$.*

Figures 6.5(a) and 6.5(b) below summarizes the welfare results for all the models we have studied.

*Figure 6.5. Social Welfare for* $n = 20$ *(a)* $\beta = 1$ *(b)* $k = 10$




We observe that social welfare is lowest when $\beta = 0$, i.e., when links define only the market sharing agreements. When links imply both R&D collaborations and market sharing agreements, social welfare in regular bipartite graphs can be higher than in the R&D collaborations model. This is due both to the market power that firms gain in all markets, and also to a higher consumer surplus. As we have seen so far, firms in these networks invest more in R&D because they compete with less of their collaborators. In this way, they fully benefit from the R&D investments of their neighbors. It is very similar to the idea in the independent market case studied by Goyal and Moraga-Gonzalez (2001). They find that when firms collaborate on R&D, but operate in independent markets, the complete network is the efficient network. Thus, we can say that both firms and consumers benefit more from R&D in bipartite networks.

Finally, in Figure 6.5(b), we observe that the effect of a change in the level of spillovers is not always positive. Social welfare is, with the exception of the bipartite graphs, concave with respect to $\beta$ in all networks. In sectors where technological spillovers are high, the gap between bipartite graphs and other networks, in terms of social welfare and the consumer surplus is even higher. This is also the case for industries where firms undertake R&D collaborations, but do not collude.

### Stability

We will use *pairwise stability*, introduced by Jackson and Wolinsky (1996), to analyze the stability of the complete network, the empty network, the complete components and the complete bipartite network, in turn.

#### Stability of the Complete Network

As we have discussed before, firms operate as local monopolies in the complete network since there are market sharing agreements together with R&D collaborations among all firms. When a firm deviates by deleting its link to another firm, the two firms stop benefiting from each other's R&D efforts and they compete in their respective markets, thus become duopolies. In order to check whether the complete network is stable, we need to calculate the difference in the equilibrium profits of the monopoly and the duopoly cases.

Calculating $\Pi_i(g^{n-1}) - \Pi_i(g^{n-1} - g_{ij})$, we can state the following result:

**Proposition 18** *Consider R&D collaboration in collusive networks. The complete network is always pairwise stable.*

*Figure 6.6(a) Stability of Complete Network*

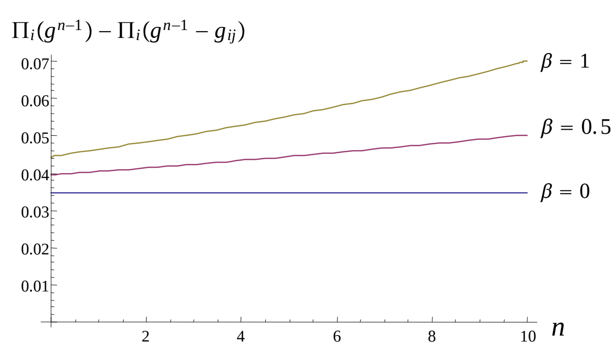



Figure 6.6(a) illustrates that the monopoly always gives higher profits than in the deviation. As the level of spillovers increases, the difference becomes even greater. $\beta = 1$ can be thought of as the independent market case in Goyal and Moraga-Gonzalez (2001), who find the same result for the stability of the complete network.

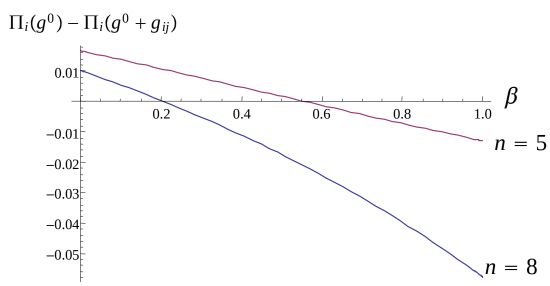
Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

### Stability of the Empty Network

In order to analyze the stability of the empty network, we need to check whether any two firms have incentives to form a link. We will calculate the equilibrium profits when two firms deviate by creating a link and compare the outcome to the equilibrium profits in the empty network. Calculating $\Pi_i(g^0) - \Pi_i(g^0 + g_{ij})$, we can state the following result:

**Proposition 19** *Consider R&D collaboration in collusive networks. The empty network is not pairwise stable when* $\beta = 1$. *There exists a* $\beta^* > 0$ *such that if* $\beta < \beta^*$, *then the empty network is stable.*

*Figure 6.6(b) Stability of Empty Network*




We observe that the stability of the empty network depends on the level of spillovers and the number of firms (markets). When spillovers are sufficiently high, firms have a greater incentive to create links, since the cost reducing benefits of R&D collaboration offsets the competition effect. The deviation profit is higher when the number of firms is large. This is because the negative effect of competition with a firm's one and only collaborator becomes negligible when there are many foreign markets that the firm operates. Therefore, the empty network is pairwise stable only when the level of spillovers is sufficiently low and the number of firms (markets) is sufficiently small.

### Stability of Complete Components

In order to analyze the stability of regular networks with complete components, we check whether any two firms in distinct components have incentives to deviate by forming a link. We can state the following result:

**Proposition 20** *Consider R&D collaboration in collusive networks and let $\beta = 1$. Regular networks with complete components with degree $k = n/2 - 1$ are never stable.*

Intuitively, when two firms in distinct components form a link, the number of collaborators increases by one, and the two firms benefit from each others' R&D efforts. At the same time, the number of competitors decreases by one. In addition, these two deviating firms never compete with each other since they operate in different markets, i.e., in different components. Therefore, in regular networks with complete components, firms always have an incentive to form links.

Goyal and Moraga-Gonzalez (2002) find that if an individual firm's profits are decreasing and log-convex in the number of firms in a market, then the stable network consists of a group of isolated firms and distinct complete components of different sizes. They show that when the components have the same size, firms have an incentive to form links. Our results suggest that firms in different components of the same size also have an incentive to form links as long as R&D spillovers are not very small.

**Stability of the Complete Bipartite Graph**

We have shown that the complete bipartite graph results in higher levels of R&D, industry profits and social welfare than the levels obtained with other configurations. It is therefore important to analyze whether it is pairwise stable. In order to do so, we first calculate the equilibrium profits of firms in the complete bipartite graphs by simply substituting $k = n/2$ into the equilibrium profit function of regular bipartite networks. We then compute the equilibrium profits when two firms deviate by forming a link, i.e., when they start to collaborate in R&D and they agree not to enter each others' markets. Comparing equilibrium profits in both cases, we can state the following result:

**Proposition 21** *Consider R&D collaboration in collusive networks. The complete bipartite graph is not stable for $\beta = 1$. If $\beta = 0$, it is stable for $n \geq 6$. Moreover, there exists a $\bar{\beta} > 0$ such that if $\beta < \bar{\beta}$, the complete bipartite graph is stable for sufficiently high $n$.*

Remember that in complete bipartite graphs, collaborating firms do not compete with each other because all of their neighbors are directly linked to their non-neighbors (they do not enter to the same markets). When two competing firms form a link, there will be less competition in their respective home markets. In addition, they will benefit from each others' R&D efforts. On the other hand, firms will take into account the fact that, by investing in R&D, they will be lowering the costs of a competitor in the foreign markets and this will affect their incentives to invest in R&D.

The intuition for obtaining pairwise stability for a large number of firms is that the deviating firm competes with the new collaborator in all of the foreign markets it enters. If the spillovers are sufficiently low, the competition effect offsets the benefits from R&D collaboration. Therefore, firms do not have an incentive to create more links in the complete bipartite network when they compete in a sufficiently high number of markets.

## 1.7 Conclusion

This chapter aims to analyze the R&D collaborations between oligopolistic firms producing homogeneous products, by introducing the possibility that these inter-firm agreements can facilitate collusion. Market sharing agreements are an alternative form of collusion that limit competition in the product market. These agreements restrain firms from entering each others' markets, decreasing the number of firms in each market and, at the same time, reducing the number of foreign markets that firms can enter. We analyze the effects of collusive agreements on the R&D incentives of firms in the case where firms form R&D collaborations, and in the case where they do not. We develop a model to examine the architecture of strategically stable networks and the relation between individual incentives and social welfare in networks.

The first benchmark model, which introduces a slight modification to Goyal and Moraga Gonzalez (2001), analyzes the case in which firms collaborate in R&D and compete in quantities in all markets. We find that R&D effort declines in the level of collaborative activity, and industry profits and social welfare are maximized at an intermediate level. Moreover, we find that the complete network is pairwise stable and that the empty network is not.

The second benchmark extends the collusive network model of Belleflamme and Bloch (2004) by allowing firms to invest in cost-reducing R&D. There are no R&D collaborations in this case, and hence firms do not benefit from each others' R&D efforts. We have shown that the equilibrium R&D effort level decreases as the number of collusive agreements increases. As the connectivity of the network increases, profits rise and consumer surplus drops, resulting in lower social welfare. When there are R&D collaborations between firms, but no market sharing agreements, we show that an intermediate level of competition is optimal for industry profits and efficiency.

When links are interpreted as possibilities for both R&D collaboration and collusion, we show that both the degree and the configuration of the network play a crucial role. Different regular networks of the same size and the same degree result in different equilibrium levels of R&D, profits and welfare. Therefore, in order to show that the configuration matters, we study three different families of regular networks: networks with complete components, bipartite graphs, and ring lattices.

The first family of graphs captures the case in which collaborators always compete with each other in the foreign markets in which they operate. Firms in the same component jointly collaborate and collude. These firms operate in the markets in the other component because they are not linked to the firms in that component. Since collaborators benefit from each others' R&D efforts, firms have less incentives to invest in R&D in order to avoid making their competitors tougher in the foreign product markets. This family of graphs results in the lowest level of R&D effort, compared to other configurations.

In bipartite networks, where collaborators are less likely to compete with each other in the foreign markets, connectivity increases the equilibrium level of R&D, industry profits and social welfare. The equilibrium level of R&D effort is higher than the level obtained in the complete components and the ring lattices. In the latter, collaborators are more likely to compete with each other in the foreign markets due to high level of clustering. Industry profits increase with respect to the degree of the network in all these configurations, but the highest social welfare is attained under the complete bipartite graph in which collaborators never compete with each other. Moreover, consumers also benefit as a result of the high levels of cost-reducing R&D. The social welfare attained under complete bipartite graphs can be even higher than levels of welfare with no collusion. This result emphasizes the importance of taking network structure into account when considering policy questions.

As for the stability, we have shown that the complete network is pairwise stable. Stability of the empty network, the complete components and complete bipartite graph depend on the level of spillovers and the number of firms (markets) in the industry. When the level of R&D spillovers is high, firms always have an incentive to form links in all of these graphs, which means that they are not stable. The empty network becomes stable when there is a sufficiently small number of firms (markets) because the effect of competition with collaborators becomes considerable if there are only a few markets. In the complete bipartite graph, firms have less incentives to form links if there are many markets in which they will compete with the new collaborator. Therefore, the complete bipartite graph is stable when there is a high number of firms (markets).

One possibility for further research would be to study the incentives of firms to form collusive agreements and R&D collaborations when the decisions are taken independently. A model with two types of links networks, if tractable, would allow for this kind of analysis.

We have implemented the pairwise stability notion of Jackson and Wolinsky (1996), in which firms are allowed to add or sever links one at a time. This results in a weak stability criterion since firms' deviations are constrained. In order to allow firms to change an arbitrary number of links, one can implement the strong pairwise stability notion proposed in Belleflamme and Bloch (2004), which uses the simultaneous linking game introduced by Myerson (1991). We can then analyze how it will alter the results of this paper.

An important extension would be to model an alternative form of collusion in which firms choose link-specific quantities for each market and maximize joint profits with their neighbors (similar to the model with link-specific efforts as in *Hybrid R&D* paper by Goyal, Konovalov and Moraga-Gonzalez (2005)). This will allow us to see the robustness of the current results, and complement the existing theoretical models on mergers and R&D joint ventures with a richer framework.

In the literature it is still debated whether large firms, because of their monopoly power, have higher or fewer incentives to undertake R&D. One could introduce firm heterogeneity

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

into this model in order to study the effects of firm size on R&D incentives. Moreover, one could extend the model to account for product innovations in differentiated markets, instead of cost-reducing R&D in homogenous-product markets. This would allow us to analyze R&D as a stochastic process and to study the effect of uncertainty about the product innovation on the stable network configurations. The underlying idea would be that when uncertainty is high, networks with isolated nodes might turn out to be stable.

Last but not least, an important contribution to the empirical literature could be made by using R&D collaboration network data (as in Hagedoorn, 2002) and to analyze the markets in which these firms operate. One could assume that there is an implicit 'coordination' taking place between collaborating firms that operate in distinct markets. This would allow us to generate a collusive network, to study its structure and to analyze whether firms have higher levels of R&D investments in these networks.

## 1.8 Appendix

Proof of Proposition 1. We can calculate

$$e(g^k) - e(g^{k+1}) = \frac{(a-\bar{c})\beta[\gamma(n+1)^2 - (n-\beta k)(n-\beta(k+1))]}{[\gamma(n+1)^2 - (n-\beta k)(1+\beta k)]\,[\gamma(n+1)^2 - (n-\beta(k+1))(1+\beta(k+1))]} > 0$$

It follows from the second order condition, i.e., $\gamma > \frac{n^2}{(n+1)^2}$, that $\gamma(n+1)^2 > (n-\beta k)(n-\beta(k+1))$. Hence, the numerator is positive. Since $n \geq 1+\beta k$, using the second order condition we can write $\gamma(n+1)^2 > (n-\beta k)(1+\beta k)$. Hence, the denominator is also positive.

Proof of Proposition 2. It follows from

$$c(g^k) - c(g^{k+1}) = \frac{(a-\bar{c})\gamma(n+1)^2[n-\beta(2k+1)-1]}{[\gamma(n+1)^2 - (n-\beta k)(1+\beta k)][\gamma(n+1)^2 - [(1+\beta+\beta k)(n-\beta(1+k))]]}$$

We know from the proof of Proposition 1 that the denominator is positive. The numerator is positive as long as $n > \beta(2k+1)-1$.

Proof of Proposition 3. It is enough to show that $\Pi_i(g^0) < \Pi_i(g^{n-1}) < \Pi_i(g^{n-2})$.

First we show that

$$\begin{array}{rcl} \Pi_i(g^0) & < & \Pi_i(g^{n-1}) \\ \dfrac{(a-\bar{c})^2\gamma n[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)^2 - n]^2} & < & \dfrac{(a-\bar{c})^2\gamma n[\gamma(n+1)^2 - (n-\beta(n-1))^2]}{[\gamma(n+1)^2 - (n-\beta(n-1))(1+\beta(n-1))]^2} \end{array}$$

It is easy to see that the numerator of $\Pi_i(g^0)$ is smaller since $n^2 > (n-\beta(n-1))^2$, and that the denominator is larger since $n < (n-\beta(n-1))(1+\beta(n-1))$.

Second,

$$\Pi_i(g^{n-2}) - \Pi_i(g^{n-1}) = \frac{(a-\bar{c})^2\gamma n\Theta}{[\gamma(n+1)^2-(n-\beta(n-1))(1+\beta(n-1))]^2[\gamma(n+1)^2-(n-\beta(n-2))(1+\beta(n-2))]^2}$$

where

$$\begin{aligned} \Theta &= \gamma^2(n+1)^4\beta(6\beta n-4n-9\beta+2) && (1)\\ &+\gamma(n+1)^2\beta(2n(\beta-1)-3\beta)(1-n(2+n)-6\beta+4n\beta+(5+2(-3+n)n)\beta^2) && (2)\\ &-\beta(n-\beta(n-2))^2(n-\beta(n-1))^2(2+2\beta n-3\beta) \end{aligned}$$

For any value of $n$, both (1) and (2) are increasing in $\gamma$; (1) increases at the rate $2\gamma(n+1)^4\beta(6\beta n-4n-9\beta+2)$, and (2) increases at rate $(n+1)^2\beta(2n(\beta-1)-3\beta)(1-n(2+n)-6\beta+4n\beta+(5+2(-3+n)n)\beta^2)$. We know that $\gamma \geq n^2/(n+1)^2$ must hold. And the inequality must hold for the lowest possible $\gamma$, which is $\gamma = n^2/(n+1)^2$. After the substitution, we have $2n^2(6\beta n-4n-9\beta+2) > (2n(\beta-1)-3\beta)(1-n(2+n)-6\beta+4n\beta+(5+2(-3+n)n)\beta^2)$. With the help of Mathematica 7.0, we observe that the inequality holds for $n \geq 4$ and $0.946254 < \bar{\beta} \leq 1$. As $n$ increases, it is also satisfied with $\beta < \bar{\beta}$. When $\beta < 0.68$, it never holds for any $n$.

Proof of Proposition 4.

Part 1. In order to show that the empty network is not efficient we calculate

$$W(g^0) - W(g^1) = \frac{1}{2}(a-\bar{c})^2\gamma n^2\left[\frac{\gamma(n+1)^2(n+2)-2n^2}{[\gamma(n+1)^2-n]^2} - \frac{\gamma(n+1)^2(n+2)-2(n-\beta)^2}{[\gamma(n+1)^2-(n-\beta)(1+\beta)]^2}\right].$$

We can easily see that the numerator of the first fraction is smaller since $n^2 > (n-\beta)^2$. The denominator is greater if $n < (n-\beta)(1+\beta)$, which becomes $n > 1+\beta$. Therefore, we find that $W(g^0) - W(g^1) < 0$ for $0 < \beta \leq 1$ and $n \geq 2$.

Part 2. In order to prove that for high level of spillovers the social welfare is maximized at an intermediate level, we show that $W(g^{n-1}) - W(g^{n-2}) < 0$. For $\gamma = 1$ and $\beta = 1$, we can compute

$$W(g^{n-1}) - W(g^{n-2}) = \frac{1}{2}(a-\bar{c})^2 n^2\left[\frac{n(5+n(4+n))}{(1+n+n^2)^2} - \frac{n(5+n(4+n))-6}{(3+n^2)^2}\right].$$

$W(g^{n-1}) - W(g^{n-2}) < 0$ when

$$(n-2)n(n^2+4n+5)(2n^2+n+4) > 6(1+n+n^2)^2$$

which is true for $n > 2$.

We have calculated in Mathematica that for $\gamma = 1$ and $\beta \leq 0.5$, $W(g^{n-1}) > W(g^{n-2})$. Hence, welfare is increasing in the level of collaborative activity for low levels of spillovers. When $\beta > 0.5$, there exists an $\bar{n}$ such that if $n > \bar{n}$, we have $W(g^{n-2}) > W(g^{n-1})$. As $\beta \to 1$, $\bar{n}$ becomes smaller.

Part 3. In order to prove that $\tilde{k} < \hat{k}$, we need to show that $dW(g^k)/dk \mid_{k=\hat{k}} < 0$. Assuming that $k$ is a continuous variable, the $\tilde{k}$ that maximizes $W(g^k)$ satisfies:

$$\gamma(n+1)^2[(n+3)(n-2\beta\tilde{k})-2] - 2(n-\beta\tilde{k})^3 = 0$$

We have found that the $\hat{k}$ that maximizes $\Pi(g^k)$ satisfies:

$$\gamma(n+1)^2(2n-3\beta\hat{k}-1) - (n-\beta\hat{k})^3 = 0$$

We need to show

$$\gamma(n+1)^2[(n+3)(n-2\beta\hat{k})-2] > 2\gamma(n+1)^2(2n-3\beta\hat{k}-1)$$

After simplifications, the condition becomes $n - 2\beta\hat{k} > 1$, which holds for $n > 3$.

Proof of Proposition 5.

In order to analyze the pairwise stability of the complete network, we need to check whether a firm has an incentive to break a link. We will calculate the equilibrium profits when a firm deviates by breaking a link and compare the outcome with the equilibrium profits in the complete network.

We can obtain the complete network profits by simply substituting $k = n - 1$ into the equilibrium profit function derived before. Therefore, we get

$$\Pi(g^{n-1}) = \frac{(a-\bar{c})^2\gamma n[\gamma(n+1)^2 - (n-\beta(n-1))^2]}{[\gamma(n+1)^2 - (n-\beta(n-1))(1+\beta(n-1))]^2}$$

Consider that some firm $i$ deletes its link with some firm $j$. There are two types of firms: on one hand, firms $i$ and $j$, which have $n-2$ links; and on the other hand, the rest of the firms, which have $n-1$ links. We will look for symmetric solutions. Let us denote the effort of the latter $n-2$ firms that maintain $n-1$ collaborations as $e_m$, and the representative firm in this group as firm $l$. Since all the firms compete with each other in the last stage, the profit functions of the two types will be the same. The costs will be:

$$\begin{aligned}
c_i &= c - e_i - \beta e_l - \beta(n-3)e_m \\
c_j &= c - e_j - \beta e_l - \beta(n-3)e_m \\
c_l &= c - e_l - \beta e_i - \beta e_j - \beta(n-3)e_m
\end{aligned}$$

Using these cost expressions we can write the second stage profit functions

$$\Pi_i(g^{n-1} - g_{ij}) = n\frac{\left[a - c + (n-\beta(n-2))e_i - (1+\beta(n-2))e_j + (2\beta-1)e_l + (n-3)(2\beta-1)e_m\right]^2}{(n+1)^2} - \gamma n e_i^2$$

The first order condition becomes

$2n(n-\beta(n-2))\left[a - c + (n-\beta(n-2))e_i - (1+\beta(n-2))e_j + (2\beta-1)e_l + (n-3)(2\beta-1)e_m\right] - \gamma n(n+1)^2 e_i = 0$

Note that the efforts of $i$ and $j$ become strategic substitutes.

$$\Pi_l(g^{n-1} - g_{ij}) = n\frac{\left[a - c + (n-\beta(n-1))e_l + (3\beta-1)e_i + (3\beta-1)e_j + (n-3)(2\beta-1)e_m\right]^2}{(n+1)^2} - \gamma n e_l^2$$

The first order condition is

$2n(n-\beta(n-1))\left[a - c + (n-\beta(n-1))e_l + (3\beta-1)e_i + (3\beta-1)e_j + (n-3)(2\beta-1)e_m\right] - \gamma n(n+1)^2 e_l = 0$

Invoking symmetry, i.e., $e_i = e_j$ and $e_l = e_m$, we find the equilibrium efforts

$$\begin{aligned}
e_i(g^{n-1} - g_{ij}) &= \frac{-((a-c)(n(-1+\beta) - 2\beta)((n(-1+\beta) - \beta)(-1+\beta) - (1+n)\gamma))}{\Lambda} \\
e_l(g^{n-1} - g_{ij}) &= \frac{-((a-c)(n(-1+\beta) - \beta)((n(-1+\beta) - 2\beta)(-1+2\beta) - (1+n)\gamma))}{\Lambda}
\end{aligned}$$

$$\begin{aligned}
\text{where } \Lambda \;=\; & (6\beta^3 - 8\beta^4 + \beta\gamma - \gamma^2 + n^3(1+\beta(-3+2\beta) - \gamma)((-1+\beta)\beta + \gamma) \\
& + \beta^2(-2+3\gamma) - n^2(1+\beta(-5+2\beta(7+5(-2+\beta)\beta)) - 3\gamma \\
& + \beta(3+\beta)\gamma + 3\gamma^2) + n(-3\beta - 23\beta^3 + 16\beta^4 + (2-3\gamma)\gamma + \beta^2(12+\gamma)
\end{aligned}$$

Substituting these into the profit function, we get

$$\Pi_i(g^{n-1} - g_{ij}) = \frac{(a-c)^2 n\gamma(-(n+2\beta-n\beta)^2 + (1+n)^2\gamma)(((-1+\beta)\beta + \gamma + n(-(-1+\beta)^2 + \gamma))^2)}{\Lambda^2}$$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Setting $\gamma = 1$ and $\beta = 1$, we can compute

$$\Pi(g^{n-1}) - \Pi_i(g^{n-1} - g_{ij}) = \frac{(a-c)^2 n\Phi}{[n^3+4n^2-2n+1]^2[(n+1)^2-n]^2}$$

where

$$\Phi = 4n^7 + 15n^6 - 4n^5 - 20n^4 + 30n^3 + 8n^2 + 12n + 3 > 0$$

We also observe in Mathematica that $\Pi(g^{n-1}) - \Pi_i(g^{n-1} - g_{ij}) > 0$ for $0 \leq \beta \leq 1$.

Stability of the Empty Network

In order to analyze the stability of the empty network, we need to check whether any two firms have incentives to form a link. We will calculate the equilibrium profits when two firms deviate by forming a link and compare the outcome with the equilibrium profits in the empty network.

We can obtain the empty network profits by simply substituting $k = 0$ into the equilibrium profit function. Therefore, we get

$$\Pi(g^0) = \frac{(a-\bar{c})^2\gamma n[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)^2-n]^2}$$

Consider that firm $i$ and $j$ form a link. There are two types of firms: on one hand, firms $i$ and $j$, which have a link; and on the other hand, the rest of the firms, which do not have any links. Let us denote the effort of the firms without collaborations as $e_m$, and the representative firm in this group as firm $l$. Since all of the firms compete with each other in the last stage, the profit functions of the two types will be the same. The costs will be:

$$\begin{aligned} c_i &= c - e_i - \beta e_j \\ c_j &= c - e_j - \beta e_i \\ c_l &= c - e_l \end{aligned}$$

Using these cost expressions, we can write the second stage profit functions as

$$\Pi_i(g^0 + g_{ij}) = n\frac{\left[a - c + (n-\beta)e_i + (n\beta - 1)e_j - e_l - (n-3)e_m)\right]^2}{(n+1)^2} - \gamma n e_i^2$$

The first order condition becomes

$$2n(n-\beta)\left[a - c + (n-\beta)e_i + (n\beta-1)e_j - e_l - (n-3)e_m)\right] - \gamma n(n+1)^2 e_i = 0$$

Note that the efforts of $i$ and $j$ become strategic complements for $\beta > 1/n$.

$$\Pi_l(g^0 + g_{ij}) = n\frac{\left[a - c + ne_l - (1+\beta)e_i - (1+\beta)e_j - (n-3)e_m)\right]^2}{(n+1)^2} - \gamma n e_l^2$$

The first order condition is

$$2n^2\left[a - c + ne_l - (1+\beta)e_i - (1+\beta)e_j - (n-3)e_m)\right] - \gamma n(n+1)^2 e_l = 0$$

Invoking symmetry, i.e., $e_i = e_j$ and $e_l = e_m$, we can write the reaction functions of firm $i$ and $l$

$$\begin{aligned} e_i(g^0 + g_{ij}) &= \frac{(n-\beta)(a-c-(n-2)e_l)}{\gamma(n+1)^2 - (n-1)(n-\beta)(1+\beta)} \\ e_l(g^0 + g_{ij}) &= \frac{n(a-c-2(1+\beta)e_i)}{\gamma(n+1)^2 - 3n} \end{aligned}$$

For $\gamma = 1$, there is a corner solution with $e_l(g^0 + g_{ij}) = 0$ and $e_i(g^0 + g_{ij}) = \dfrac{(n-\beta)(a-c)}{\gamma(n+1)^2 - (n-1)(n-\beta)(1+\beta)}$ if $n \geq 5$.

Plugging these into the profit function, we obtain

$$\Pi_i(g^0 + g_{ij}) = \frac{(a-c)^2 n(2n-\beta+1)(1+\beta)}{[\beta(n-1)(n-\beta-1)-(3n+1)]^2}$$

We can then write the difference as

$$\Pi(g^0) - \Pi_i(g^0+g_{ij}) = \frac{(a-\bar{c})^2 n(2n+1)}{[(n+1)^2-n]^2} - \frac{(a-c)^2 n(2n-\beta+1)(1+\beta)}{[\beta(n-1)(n-\beta-1)-(3n+1)]^2} < 0$$

The numerator of $\Pi(g^0)$ is smaller than the numerator of $\Pi_i(g^0+g_{ij})$ if $(2n+1) < (2n+1-\beta)(1+\beta)$. After simplifications, it becomes $\beta < 2n$, which holds for $n > 0$. When we compare the denominators, in order to have the first fraction smaller, it must be that $(n+1)^2 - 1 + 2n > \beta(n-1)^2 - \beta^2(n-1)$. Since $(n+1)^2 > \beta(n-1)^2$, we can focus on the second terms. It must be that $1 - 2n < \beta^2(n-1)$. Rearranging the terms, we get $1 + \beta^2 < 2n + \beta^2 n$. Since $1 < 2n$ and $\beta^2 < \beta^2 n$, we can say that the denominator of the first fraction is greater. We can conclude that $\Pi(g^0) - \Pi_i(g^0 + g_{ij}) < 0$, thus, the empty network is not pairwise stable.

Proof of Proposition 6. We calculate

$$e(g^k) - e(g^{k+1}) = \frac{(a-\bar{c})\gamma n[2(n-k)^2-1]}{[\gamma n(n-k+1)^2-(n-k)^2]\,[\gamma n(n-k)^2-(n-k-1)^2)]} > 0$$

We observe that the numerator is positive since $(n-k)^2 > 1/2$ and the denominator is positive due to the second order condition, i.e., $\gamma n(n-k+1)^2 > (n-k)^3$.

Proof of Proposition 7. It follows from

$$c(g^k) - c(g^{k+1}) = \frac{(a-\bar{c})\gamma[1-2(n-k)^2]}{[\gamma n(n-k+1)^2-(n-k)^2][\gamma n(n-k)^2-(n-k-1)^2]} < 0$$

We observe that the numerator is negative since $(n-k)^2 > 1/2$ and the denominator is positive because the second order condition is $\gamma n(n-k+1)^2 > (n-k)^3$.

Proof of Proposition 8.

Suppose that $n$ and $k$ are continuous variables. The level of collusive activity, $\check{k}$, that maximizes $\Pi(g^k)$ satisfies

$$\frac{\partial \Pi(g^k)}{\partial k} = \frac{(a-\bar{c})^2\gamma^2 n^2[(n-k)^2-1][\gamma n(n-k+1)^2-3(n-k)^3]}{[\gamma n(n-k+1)^2-(n-k)^2]^3} = 0$$

We observe that the equality is satisfied when $(n-\check{k})^2 - 1 = 0$, which holds only for $\check{k} = n-1$, i.e., the complete network.

Proof of Proposition 9.

First we show that $W(g^{n-2}) > W(g^{n-1})$. Replacing $k = n-2$ and $k = n-1$ in $W(g^k)$, we obtain

$$W(g^{n-2}) - W(g^{n-1}) = \left[\frac{(a-\bar{c})^2\gamma^2 n^3(10\gamma n+1)}{(4\gamma n-1)^2(9\gamma n-4)}\right] > 0$$

since $\gamma(n+1)^2 > n^2$ from the second order condition.

We then show when the empty network is not efficient by computing $W(g^1) - W(g^0)$ :

$$W(g^1) - W(g^0) = \frac{1}{2}(a-\bar{c})^2\gamma^2 n\left[\frac{\gamma n^3(n^2-1)-2(n-1)^4}{[\gamma n^3-(n-1)^2]^2} - \frac{\gamma(n+1)^2(n+2)-2n^2}{[\gamma(n+1)^2-n]^2}\right]$$

For $\gamma = 1$, this becomes

$$W(g^1) - W(g^0) = \frac{1}{2}(a-\bar{c})^2 n^2\left[\frac{(-4+n(7+n(4+n(-8+n(2+n(-13+2n))))))}{(-1+n(1+n^2(2+n^2)))^2}\right]$$

which is positive for $n \geq 7$. (For any value of $\gamma > n^2/(n+1)^2$, $W(g^1) > W(g^0)$ when $n \geq 5$.)

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Using Mathematica 7.0, if we solve for $W(g^k) - W(g^{k+1}) < 0$ when $\gamma = 1$ and $a > c$, the output we obtain is as follows:

$(k = 0$ & $n \geq 7)$ || $(k = 1$ & $n \geq 9)$ || $(k = 2$ & $n \geq 11)$ || $(k = 3$ & $n \geq 13)$ || $(k = 4$ & $n \geq 16)$ || $(k = 5$ & $n \geq 18)$ || $(k = 6$ & $n \geq 20)$ || $(k = 7$ & $n \geq 22)$ || $(k = 8$ & $n \geq 24)$ || $(k = 9$ & $n \geq 26)$ || $(k \geq 10$ & $n > f(k))$ where $f(k)$ is a quadratic function of $k$.

Therefore, we can conclude that the empty network is efficient for $n \leq 6$ when $\gamma = 1$.
In addition, suppose that $n$ and $k$ are continuous variables, we calculate:

$$\frac{\partial W(g^k)}{\partial k} = \frac{(a-\bar{c})^2 n^3 (n-k+1)[-n + ((n-k)(2k^2 - 3k(n-1) + n(n-5))]}{[n + (n-k)((n(n+1) - k(n-1))]^3}$$

Solving for $\dfrac{\partial W(g^k)}{\partial k} = 0$ when $\gamma = 1$, the Mathematica output we obtain is as follows:

$(k = 1$ & $7 \leq n < 10)$ || $(k = 2$ & $10 \leq n < 11)$ || $(k = 3$ & $12 \leq n < 13)$ ||
$(k = 4$ & $14 \leq n < 17)$ ..... || $(k = 8$ & $23 \leq n < 25)$.... || $(k = 16$ & $39 \leq n < 41)$, and so on.

Therefore, we can conclude that when $n \geq 7$, intermediate levels of collusive activity maximize social welfare.

Proof of Proposition 10.
We obtain the complete network profits by simply substituting $k = n - 1$ into the equilibrium profit function derived before. Therefore, we get

$$\Pi(g^{n-1}) = \frac{(a-\bar{c})^2 \gamma n}{(4\gamma n - 1)^2}$$

Suppose that some firm $i$ deletes its link with some firm $j$. There are two types of firms: on the one hand, firms $i$ and $j$, which have $n-2$ links, are duopoly in their respective markets. On the other hand, the rest of the firms, denoted by $l$, which have $n-1$ links, are monopolies in their home markets. We can write the profit functions of the firms as:

$$\begin{aligned}
\Pi_i(g^{n-1} - g_{ij}) &= 2\frac{(a - 2c_i + c_j)^2}{9} - \gamma n e_i^2 \\
\Pi_j(g^{n-1} - g_{ij}) &= 2\frac{(a - 2c_j + c_i)^2}{9} - \gamma n e_j^2 \\
\Pi_l(g^{n-1} - g_{ij}) &= 2\frac{(a - c_l)^2}{4} - \gamma n e_l^2 \\
c_i &= c - e_i \\
c_j &= c - e_j
\end{aligned}$$

Since there are no collaboration links among firms, the costs of firms $i$ and $j$ only depend on their own effort. We will look for symmetric solutions. Using the cost expressions we can write the second stage profit functions as

$$\begin{aligned}
\Pi_i(g^{n-1} - g_{ij}) &= 2\frac{[a - c + 2e_i - e_j]^2}{9} - \gamma n e_i^2 \\
\Pi_j(g^{n-1} - g_{ij}) &= 2\frac{[a - c + 2e_j - e_i]^2}{9} - \gamma n e_j^2
\end{aligned}$$

The first order conditions become

$$\begin{aligned}
8\left[a - c + 2e_i - e_j\right] - 18\gamma n e_i &= 0 \\
8\left[a - c + 2e_j - e_i\right] - 18\gamma n e_j &= 0
\end{aligned}$$

Note that the efforts of $i$ and $j$ become strategic substitutes.

 DOI: 10.2870/63947

Invoking symmetry, i.e., $e_i = e_j$, we find the equilibrium effort, profit and the deviation profit:

$$\begin{aligned} e_i(g^{n-1} - g_{ij}) &= \frac{4(a-c)}{9\gamma n - 4} \\ \Pi_i(g^{n-1} - g_{ij}) &= \frac{2(a-c)^2(9\gamma n - 8)}{(9\gamma n - 4)^2} \\ \Pi(g^{n-1}) - \Pi_i(g^{n-1} - g_{ij}) &= \frac{(a-c)^2 n^2 \gamma^2 (9\gamma n + 10)}{(9\gamma n - 4)^2(4\gamma n - 1)} > 0 \end{aligned}$$

Therefore, the complete network is pairwise stable.

Stability of the Empty Network

In order to analyze the stability of the empty network, we need to check whether any two firms have incentives to form a link. We will calculate the equilibrium profits when two firms deviate by forming a link and compare the outcome with the equilibrium profits under the empty network.

We can obtain the empty network profits by simply substituting $k = 0$ into the equilibrium profit function. Therefore, we get

$$\Pi(g^0) = \frac{(a-\bar{c})^2\gamma n[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)^2 - n]^2}$$

Suppose that firm $i$ and $j$ form a link. There are two types of firms: on one hand, firms $i$ and $j$, which have a link, thus do not enter each other's markets. They compete with $n-1$ firms in their home markets and with $n$ firms in all the other foreign markets. On the other hand, the rest of the firms, which do not have any links, operate in all the markets. Let us denote the effort of the firms without collaborations as $e_m$, and the representative firm in this group as firm $l$. The profit functions of the two types are

$$\begin{aligned} \Pi_i(g^0 + g_{ij}) &= \frac{(a - (n-1)c_i + c_l + (n-3)c_m)^2}{n^2} + (n-2)\frac{(a - nc_i + c_j + c_l + (n-3)c_m)^2}{(n+1)^2} - \gamma n e_i^2 \\ \Pi_l(g^0 + g_{ij}) &= \frac{(a - nc_l + c_i + c_j + (n-3)c_m)^2}{(n+1)^2} + \frac{(a - (n-1)c_l + c_i + (n-3)c_m)^2}{n^2} \\ &\quad + \frac{(a - (n-1)c_l + c_j + (n-3)c_m)^2}{n^2} + (n-3)\frac{(a - nc_l + c_i + c_j + (n-3)c_m)^2}{(n+1)^2} - \gamma n e_l^2 \\ c_i &= c - e_i \\ c_j &= c - e_j \\ c_l &= c - e_l \end{aligned}$$

Using the cost expressions we can write the profit functions of the second stage as:

$$\Pi_i(g^0 + g_{ij}) = \frac{\left[a - c + (n-1)e_i - e_l - (n-3)e_m)\right]^2}{n^2} + (n-2)\frac{\left[a - c + e_i - e_j - e_l - (n-3)e_m)\right]^2}{(n+1)^2} - \gamma n e_i^2$$

The first order condition becomes

$$\frac{2(n-1)\left[a - c + (n-1)e_i - e_l - (n-3)e_m)\right]}{n^2} + \frac{2(n-2)\left[a - c + e_i - e_j - e_l - (n-3)e_m)\right]}{(n+1)^2} - 2\gamma n e_i = 0$$

Note that the efforts of $i$ and $j$ are strategic substitutes.

$$\begin{aligned} \Pi_l(g^0 + g_{ij}) &= \frac{\left[a - c + (n-1)e_l - e_i - (n-3)e_m)\right]^2}{n^2} + \frac{\left[a - c + (n-1)e_l - e_j - (n-3)e_m)\right]^2}{n^2} \\ &\quad + (n-2)\frac{\left[a - c + ne_l - e_i - e_j - (n-3)e_m)\right]^2}{(n+1)^2} - \gamma n e_i^2 \end{aligned}$$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

The first order condition is

$$\frac{4(n-1)\left[a-c+(n-1)e_l-(n-3)e_m\right]}{n^2}-\frac{2(n-1)(e_i+e_j)}{n^2}$$
$$+\frac{2n(n-2)\left[a-c+ne_l-e_i-e_j-(n-3)e_m\right]}{(n+1)^2}-2\gamma ne_l=0$$

Invoking symmetry, i.e., $e_i=e_j$ and $e_l=e_m$, we can write the reaction functions of firm $i$ and $l$ as:

$$e_i(g^0+g_{ij}) = \frac{(a-c-(n-2)e_l)((n^2+1)n(n-1)-1))}{2n^2-n^3(2-\gamma)-n^5(1-\gamma)+2n^4(1+\gamma)-1}$$

$$e_l(g^0+g_{ij}) = \frac{(a-c)(n^4-2n^2-2n-2)-2((n^2+1)n(n-1)-1)e_i}{\gamma(n+1)^2-3n}$$

For $\gamma=1$, there is a corner solution with $e_l(g^0+g_{ij})=0$ and $e_i(g^0+g_{ij})=\dfrac{(a-c)((n^2+1)n(n-1)-1)}{2n^2-n^3+4n^4-1}$ when $n\geq 6$.

Plugging these into the profit function, we obtain

$$\Pi_i(g^0+g_{ij})=\frac{(a-c)^2[n^3(n+1)^2(((n(n-1)+2)n+1)-((n^2+1)n(n-1)-1))^2]}{[2n^2-n^3+4n^4-1]^2}$$

We can calculate

$$\Pi(g^0)-\Pi_i(g^0+g_{ij})=\frac{(a-\bar{c})^2n(2n+1)}{[(n+1)^2-n]^2}-\frac{(a-c)^2[n^3(n+1)^2(((n(n-1)+2)n+1)-((n^2+1)n(n-1)-1))^2]}{[2n^2-n^3+4n^4-1]^2}$$

Using Mathematica, we find that $\Pi(g^0)-\Pi_i(g^0+g_{ij})<0$. Hence, the empty network is not stable for $n\geq 6$ and $\gamma=1$.

Finally, for $n<6$, we calculate

$$\Pi(g^0)-\Pi_i(g^0+g_{ij})=\frac{(a-\bar{c})^2n(2n+1)}{[(n+1)^2-n]^2}-\frac{(a-c)^2[(2+n^2)^2(-1+n(-2+n(1+n(1+n(1+3n)(3+n(n-1))))))]}{(2+n(2+n+8n^2-n^3+n^5+2n^6))^2}$$

and find that $\Pi(g^0)-\Pi_i(g^0+g_{ij})\geq 0$ for $n\geq 3$. Therefore, we can conclude that the empty network is stable for $3\leq n<6$ and $\gamma=1$.

Proof of Proposition 13. We prove this by showing first that $e(g^{n-1})<e(g^{n-2})$

$e(g^{n-1})=\dfrac{(a-\bar{c})}{4\gamma-\beta(n-1)-1}$ and $e(g^{n-2})=\dfrac{4(a-\bar{c})}{9\gamma-4\beta(n-2)-4}$

$e(g^{n-1})<e(g^{n-2})$ reduces to $4\beta<7\gamma$.

This is always true since $\gamma>8/9$ from the second order condition $\gamma>\dfrac{(n-k)^3}{(n-k+1)^2}$ for $k=n-2$ and $\beta\leq 1$.

Then, we can calculate

$$e(g^k)-e(g^{k+1})=\frac{(a-\bar{c})[(n-k)^2[2\gamma-\beta(n-k-1)^2]-\gamma]}{[\gamma(n-k+1)^2-(n-k)^2(1+\beta k)]\left[\gamma(n-k)^2-(n-k-1)^2(1+\beta(k+1))\right]}$$

Since we know that $e(g^{n-1})<e(g^{n-2})$ for all $n$, we need to show that $e(g^k)-e(g^{k+1})<0$ for some $k<n-2$.

We see that $e(g^k)-e(g^{k+1})<0$ for $(n-k)^2[2\gamma-\beta(n-k-1)^2]<\gamma$. From the second order condition, we know that $\gamma>\dfrac{(n-k)^3}{(n-k+1)^2}$. Thus, it is enough to show that $\dfrac{(n-k)^3}{(n-k+1)^2}>(n-k)^2[2\gamma-\beta(n-k-1)^2]$.

Substituting $n-k=d$, we have $\dfrac{d^3}{(d+1)^2}<\gamma<\dfrac{1+\beta(d-1)^2(d+1)^2}{2(d+1)^2}$.

Assume that, towards a contradiction, $\dfrac{d^3}{(d+1)^2}>\dfrac{1+\beta(d-1)^2(d+1)^2}{2(d+1)^2}$. This never holds for $\beta=1$ and $d>2$.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Proof of Proposition 17.

Using the optimal effort and cost structures, we obtain the equilibrium welfare for networks with complete components

$$W^{cc}(g) = \frac{(a-c)^2\gamma n^2[\gamma n(n-k+1)^2(n-k)(n-k+2) - 2((n-k)^2 - \beta k(n-k-1))^2]}{2[\gamma n(n-k+1)^2 - (1+\beta k)[(n-k)^2 - \beta k(n-k-1)]]^2}$$

In regular bipartite graphs, we get

$$W^{b}(g) = \frac{(a-c)^2\gamma n^2[(\gamma n(n-k+1)^2(n-k)(n-k+2) - 2((n-k)^2 - \beta k(n-2k))^2]}{2[\gamma n(n-k+1)^2 - (1+\beta k)[(n-k)^2 - \beta k(n-2k)]]^2}$$

Finally, social welfare in dense and sparse ring lattices becomes

$$\begin{aligned} W^{sr}(g^k) &= \frac{(a-c)^2\gamma n^2[(8\gamma n(n-k+1)^2(n-k)(n-k+2) - (4(n-k)^2 - \beta k(4n-5k-6))^2]}{[4\gamma n(n-k+1)^2 - (1+\beta k)[4(n-k)^2 - \beta k(4n-5k-6)]]^2} \\ W^{dr}(g^k) &= \frac{(a-c)^2\gamma n^2[(\gamma n(n-k+1)^2(n-k)(n-k+2) - 2((n-k)^2 - \beta k(n-k-1)(n-k-2))^2]}{[\gamma n(n-k+1)^2 - (1+\beta k)[(n-k)^2(1-\beta) + \beta(3(n-k)-2)]]^2} \end{aligned}$$

In Mathematica, we observe that $W^b(g^k)$ increases in $k$ and that it is highest in the complete bipartite networks, i.e., when $k = n/2$. In addition, $W^{cc}(g^k)$ and $W^{sr}(g^k)$ are maximized at intermediate levels of connectivity, which is close to $n/3 - 1$. The equations are omitted here.

Proof of Proposition 18.

In order to analyze the stability of the complete network, we first obtain the complete network (monopoly) profits by simply substituting $k = n-1$ into the equilibrium profit function derived before:

$$\Pi_i(g^{n-1}) = \frac{(a-c)^2 n\gamma[4n\gamma - 1]}{[4n\gamma - \beta(n-1) - 1]^2}$$

When two firms break a link, they compete only with each other in the market stage, and thus become duopolies in two markets, while the remaining firms remain as monopolies in their home markets. Hence, we can write

$$\begin{aligned} \Pi_i(g^{n-1} - g_{ij}) &= 2\left[\frac{a - 2c_i + c_j}{3}\right]^2 - \gamma n e_i^2 \\ \Pi_j(g^{n-1} - g_{ij}) &= 2\left[\frac{a - 2c_j + c_i}{3}\right]^2 - \gamma n e_j^2 \\ \Pi_m(g^{n-1} - g_{ij}) &= \left[\frac{a - c_m}{2}\right]^2 - \gamma n e_m^2 \end{aligned}$$

Taking into account the R&D collaborations with the other firms,

$$\begin{aligned} c_i &= \bar{c} - e_i - \beta k e_m \\ c_j &= \bar{c} - e_j - \beta k e_m \\ c_m &= \bar{c} - e_m - \beta e_i - \beta e_j - \beta(k-2)e_m \end{aligned}$$

Substituting the costs into the profit functions, we can calculate the equilibrium efforts:

$$\begin{aligned} e_i(g^{n-1} - g_{ij}) &= e_j(g^{n-1} - g_{ij}) \\ &= \frac{4(a-\bar{c})(4n\gamma + \beta - 1)}{(36n\gamma - 25)n\gamma - \beta(9n\gamma - 4)(n-3) - 8\beta^2(n-2) + 4} \\ e_m(g^{n-1} - g_{ij}) &= \frac{4(a-\bar{c})(9n\gamma + 8\beta - 4)}{(36n\gamma - 25)n\gamma - \beta(9n\gamma - 4)(n-3) - 8\beta^2(n-2) + 4} \end{aligned}$$

Plugging the equilibrium efforts into the profit functions above, we obtain the following deviation profit for firms $i$ and $j$.

$$\Pi_i(g^{n-1} - g_{ij}) = \frac{2(a-c)^2\gamma n(4n\gamma + \beta - 1)^2(9n\gamma - 8)}{[(36n\gamma - 25)n\gamma - \beta(9n\gamma - 4)(n-3) - 8\beta^2(n-2) + 4]^2}$$

Calculating $\Pi_i(g^{n-1}) - \Pi_i(g^{n-1} - g_{ij})$, and analyzing the sign using Mathematica, we find that the complete network is stable.

Proof of Proposition 19.

In order to analyze the stability of the empty network, we first obtain the empty network profits by simply substituting $k = 0$ into the equilibrium profit function:

$$\Pi(g^0) = \frac{(a-\bar{c})^2 \gamma n[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)^2 - n]^2}$$

Consider that firm $i$ and $j$ form a link. There are two types of firms: on one hand, firms $i$ and $j$, which have a link, and thus do not enter each other's markets. They compete with $n-1$ firms in their home markets and with $n$ firms in all the other foreign markets. On the other hand, the rest of the firms, which do not have any links, operate in all the markets. Let us denote the effort of the firms without collaborations as $e_m$, and the representative firm in this group as firm $l$. The profit functions of the two types will be

$$\begin{aligned}
\Pi_i(g^0 + g_{ij}) &= \frac{(a-(n-1)c_i + c_l + (n-3)c_m)^2}{n^2} + (n-2)\frac{(a-nc_i + c_j + c_l + (n-3)c_m)^2}{(n+1)^2} - \gamma n e_i^2 \\
\Pi_l(g^0 + g_{ij}) &= \frac{(a-nc_l + c_i + c_j + (n-3)c_m)^2}{(n+1)^2} + \frac{(a-(n-1)c_l + c_i + (n-3)c_m)^2}{n^2} \\
&\quad + \frac{(a-(n-1)c_l + c_j + (n-3)c_m)^2}{n^2} + (n-3)\frac{(a-nc_l + c_i + c_j + (n-3)c_m)^2}{(n+1)^2} - \gamma n e_l^2 \\
c_i &= \bar{c} - e_i - \beta e_j \\
c_j &= \bar{c} - e_j - \beta e_i \\
c_l &= \bar{c} - e_l \\
c_m &= \bar{c} - e_m
\end{aligned}$$

Using the cost expressions we can write the second stage profit functions as

$$\Pi_i(g^0 + g_{ij}) = \frac{\left[a - c + (n-1)e_i - e_l - (n-3)e_m)\right]^2}{n^2} + (n-2)\frac{\left[a - c + e_i - e_j - e_l - (n-3)e_m)\right]^2}{(n+1)^2} - \gamma n e_i^2$$

The first order condition becomes

$$\frac{2(n-1)\left[a - c + (n-1)e_i - e_l - (n-3)e_m)\right]}{n^2} + \frac{2(n-2)\left[a - c + e_i - e_j - e_l - (n-3)e_m)\right]}{(n+1)^2} - 2\gamma n e_i = 0$$

Note that the efforts of $i$ and $j$ are strategic substitutes.

$$\begin{aligned}
\Pi_l(g^0 + g_{ij}) &= \frac{\left[a - c + (n-1)e_l - e_i - (n-3)e_m)\right]^2}{n^2} + \frac{\left[a - c + (n-1)e_l - e_j - (n-3)e_m)\right]^2}{n^2} \\
&\quad + (n-2)\frac{\left[a - c + ne_l - e_i - e_j - (n-3)e_m)\right]^2}{(n+1)^2} - \gamma n e_i^2
\end{aligned}$$

The first order condition is

$$\begin{aligned}
&\frac{4(n-1)\left[a - c + (n-1)e_l - (n-3)e_m)\right]}{n^2} - \frac{2(n-1)(e_i + e_j)}{n^2} \\
&+ \frac{2n(n-2)\left[a - c + ne_l - e_i - e_j - (n-3)e_m)\right]}{(n+1)^2} - 2\gamma n e_l = 0
\end{aligned}$$

Invoking symmetry, i.e., $e_i = e_j$ and $e_l = e_m$, we can write the reaction functions of firm $i$ and $l$ as

$$e_i(g^0 + g_{ij}) = \frac{(a-c-(n-2)e_l)((n^2+1)n(n-1)-1))}{2n^2 - n^3(2-\gamma) - n^5(1-\gamma) + 2n^4(1+\gamma) - 1}$$

$$e_l(g^0 + g_{ij}) = \frac{(a-c)(n^4 - 2n^2 - 2n - 2) - 2((n^2+1)n(n-1)-1)e_i}{\gamma(n+1)^2 - 3n}$$

For $\gamma = 1$, there is a corner solution with $e_l(g^0 + g_{ij}) = 0$ and $e_i(g^0 + g_{ij}) = \frac{(a-c)((n^2+1)n(n-1)-1)}{2n^2 - n^3 + 4n^4 - 1}$ if $n \geq 6$.

Plugging these into the profit function, we obtain

$$\Pi_i(g^0 + g_{ij}) = \frac{(a-c)^2[n^3(n+1)^2(((n(n-1)+2)n+1) - ((n^2+1)n(n-1)-1))^2]}{[2n^2 - n^3 + 4n^4 - 1]^2}$$

Using Mathematica, we can see that

$$\Pi(g^0) - \Pi_i(g^0 + g_{ij}) = \frac{(a-\bar{c})^2 n(2n+1)}{[(n+1)^2 - n]^2} - \frac{(a-c)^2[n^3(n+1)^2(((n(n-1)+2)n+1) - ((n^2+1)n(n-1)-1))^2]}{[2n^2 - n^3 + 4n^4 - 1]^2} < 0$$

Hence, the empty network is not stable for $\beta = 1$.

Remark: Parts of some proofs are omitted here. The proofs for the stability of complete components and the complete bipartite graph, which are obtained using Mathematica 7.0, will be included once simplified. Ask the author for the codes and the outputs if required.

## 1.9 References


Aghion, P., Bloom, D., Blundell R., Griffith, R. and Howitt, P. (2002). "Competition and innovation: an inverted U relationship". *NBER Working Paper,* 9269.

Belleflamme, P. and Bloch, F. (2004). "Market Sharing Agreements and Collusive Networks." *International Economic Review, 45, 387-411.*

Belleflamme P. and Peitz, M. (2010). *Industrial Organization: Markets and Strategies* (Cambridge University Press).

D'Aspremont, C. and Jacquemin, A. (1988). "Cooperative and Noncooperative R&D in Duopoly with Spillovers." *American Economic Review*, 78, 1133-1137.

Deroian, F. and Gannon, F. (2006). "Quality-Improving Alliances in Differentiated Oligopoly." *International Journal of Industrial Organization*, 24, 629-637.

Delapierre, M. and Mytelka, L. K. (1998). "Blurring Boundaries: New Inter-firm Relationships and the Emergence of Networked, Knowledge-based Oligopolies." in M. Colombo (ed.) *The Changing Boundaries of the Firm, Explaining Evolving Inter-Firm Relations*, (London, Routledge Press), 4, 72-94.

European Commission. (2010). "Draft R&D Block Exemption Regulation." Available at http://ec.europa.eu/competition/consultations/2010_horizontals/draft_rd_ber_en.pdf

Hagedoorn, J. (2002). "Inter-firm R&D Partnerships: An Overview of Major Trends and Patterns since 1960." *Research Policy* 31, 477-492.

Goyal, S. and Moraga-Gonzalez, J. L. (2001). "R&D Networks." *RAND Journal of Economics*, 32, 4, 686-707.

Goyal, S. and Moraga-Gonzalez, J. L. (2003). "Firms, Networks and Markets: A Survey of Recent Research." *Revue d'Economie Industrielle*.

Goyal, S., Konovalov, A. and Moraga-Gonzalez, J. L. (2003). "Hybrid R&D." *Tinbergen Institute discussion paper* TI 2003-041/1.

Goyal, S., Konovalov, A. and Moraga-Gonzalez, J. L. (2002). "Individual Research, Joint Work and Networks of Collaboration." *mimeo*, Erasmus University and Queen Mary, University of London.

Goyal, S. and Joshi, S. (2003). "Networks of Collaboration in Oligopoly." *Games and Economic Behavior*, 43, 57-85.

Jackson, M. O. and Wolinksy, A. (1996). "A Strategic Model of Social and Economics Networks." *Journal of Economic Theory*, 71, 44-74.

Kamien, M. I., Muller, E. and Zang, I. (1992). "Research Joint Ventures and R&D Cartels." *American Economic Review*, 85, 1293-1306.

Kranton, R. E. and Minehart, D. F. (2001). "A Theory of Buyer-Seller Networks." *American Economic Review*, 91, 485-508.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Milgrom, P. and Roberts, J. (1992). *Economics, Organization and Management*. (Englewood Cliffs, N.J.: Prentice Hall).

Motta, M. (1992). "Cooperative R&D and Vertical Product Differentiation." *International Journal of Industrial Organization*, 10, 643-661.

Motta, M. (2004). *Competition Policy: Theory and Practice* (Cambridge University Press).

Myerson, R. B. (1991). *Game Theory: Analysis of Conflict* (Cambridge, MA: Harvard University Press).

West, D. B. (2000). *Introduction to Graph Theory*, 2nd ed. (Englewood Cliffs, NJ: Prentice-Hall).

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

# Chapter 2

# Congestion in R&D Collaboration Networks with Degree Externalities

## 2.1 Introduction

There are many ways in which rival firms can collaborate in the market. These collaboration activities could involve the sharing of information about market conditions and new technologies, or even the sharing of facilities for distribution purposes. In some cases, firms share their information in exchange of advertising benefits. In an empirical survey, Hagedoorn (2002) finds that there has been a significant increase in the number of R&D collaboration between firms since 1960's, with the sharpest increase happening in 1990's. High tech sectors, such as pharmaceuticals, information technology, aerospace and defense account for 80% of the total number of collaborations. The most common forms of research collaboration between firms are joint research corporations and joint R&D and technology exchange agreements.

When we consider gaining access to information or additional resources by forming R&D collaborations, it becomes clear that connected firms might be more valuable partners due to increased knowledge spillovers. Undertaking distinct, but in some sense related, projects can create synergies in which the total return from R&D can be higher than the sum of the R&D invested in each project. These could benefit the firm and thus its collaborators. Becker and Dietz (2003) find a positive relationship between the number of partners and R&D intensity in the German manufacturing industry. They argue that the mix of heterogeneous parties cooperating in R&D releases synergies, with stimulating effects on firms' innovation input.

Deeds and Hill (1996), on the other hand, use data from biotechnology firms in the United States and find an inverted U-shaped relationship between the number of alliances a firm is engaged in and the rate of new product development. They suggest that initially, strategic alliances


Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

represent a viable way for firms to gain access to complementary assets, and to increase their rate of new product development. However, negative returns may ultimately set in due to the complexity and specific nature of the knowledge required for R&D, and burdens on management. In addition, learning from collaborators is also likely to exhibit diminishing returns, which might explain the reductions in performance.

In addition, there is evidence that a large proportion of collaboration agreements between firms fail (Podolny and Page, 1998). Dodgson (1992), Veugelers (1998) and Hemphill and Vonortas (2003) studied the motivations of firms that chose to collaborate, and the problems that they had in forming alliances. They point out the importance of factors such as communication, competence, and moral hazard for the success of research collaborations. Moreover, due to the limited supply of resources that can be devoted to R&D (such as time, space and workers), collaborating with a firm has an opportunity cost of not working with others. Therefore, the number of collaborators of a firm can negatively affect its contribution to each particular collaboration, due to time constraints and limited supplies of labor. In addition, inefficiencies might arise as firms with a large number of projects might produce spillovers that are too small to justify their R&D spending. These issues motivate us to introduce an instiututional aspect to existing models of collaboration in order to study the stability of R&D collaborations among firms that are also competitors in the product market.

Non-cooperative incentives to form R&D agreements are examined by recent theoretical models. Goyal and Moraga-Gonzalez (2001) contribute to the strategic models of network formation by developing a model in which the "quality" of links is *endogenously* determined by the players' choices of R&D effort. The authors show that when firms operate in independent markets, total R&D, industry profits and welfare are maximized in the complete network that is also pairwise stable. When the collaborators compete in the homogeneous-product market, the authors find that profits and welfare are maximized at intermediate levels of collaboration and that the complete network is still the unique stable network.

We argue that the inverted U-shaped relationship between the number of collaborators and the intensity of R&D can also be explained by the limited resource argument mentioned above. The discrepancy between the theoretical results about the stability of collaboration


DOI: 10.2870/63947

networks and the empirical evidence motivates us to introduce network-based externalities associated to R&D collaborations. We analyze theoretically how the connectivity (number of projects/collaborators) of a single firm affects its incentives to invest in R&D. When a firm chooses how much to invest in R&D, it has to take into account how best to allocate resources between its collaborators. At the same time, when deciding whether to form a collaboration link, the firm has to take into account how many collaborators the firm in consideration has, since the returns from this R&D collaboration will depend on its partner's connectivity.

We analyze the relationship between market competition, firms' incentives to invest in R&D, the effects of the number of R&D partners and the architecture of collaboration networks. In particular, we address the following questions:

(i) What are the effects of collaborative activity and the level of connectivity of R&D partners on the R&D levels and the profits of competing firms?

(ii) What are the strategically stable networks?

(iii) What is the architecture of networks that maximize social welfare in the presence of positive and negative externalities due to connectivity?

The rest of the chapter is organized as follows. We begin with a brief overview of the relevant literature. In Section 2.3, we present the model. Section 2.4 focuses on the equilibrium and welfare properties when collaborating firms operate in independent markets. We introduce competition into the model in Section 2.5. In Section 2.6, we conclude and discuss the possible extensions for further research.

## 2.2 Related Literature

The literature on R&D cooperation as a way to internalize spillovers starts with d'Aspremont and Jacquemin (1988), who compare cooperative and non-cooperative schemes in oligopolistic industries, focusing on cost-reducing R&D. The literature suggests that cooperation increases the level of R&D, resulting in more output and higher welfare. Motta (1992) draws the same conclusion by extending d'Aspremont and Jacquemin's cooperative framework to product innovations.

Collaboration agreements have the distinctive structural feature of being *nonexclusive* (Milgrom and Roberts, 1992) and *bilateral* (Delapierre and Mytelka, 1998). While the coalition formation approach requires every player to belong to one group only, non-exclusivity allows for a pattern of relations in which firms 1 and 2, and firms 2 and 3 have joint ventures, but there is no link between firms 1 and 3. The bilateral structure allows us to analyze these collaborations in a strategic model of R&D networks with *pairwise* links.

As mentioned before, Goyal and Moraga-Gonzalez (2001) consider an oligopoly with identical firms that have opportunities to form pairwise collaborative links through which they can share nonexclusive R&D knowledge about a cost-reducing technology. Given this collaboration network, firms unilaterally choose a single (costly) level of effort in R&D and operate (in independent markets and in a homogeneous-product market) by setting quantities. The authors show that when firms operate in independent markets, total R&D, industry profits and welfare are maximized in the complete network, which is pairwise stable. However, when collaborators compete in the homogeneous-product market, individual R&D effort declines in the level of collaborative activity. Levels of cost reduction and profits are non-monotonic in the number of collaborations (decreasing after an initial increase). The complete network is the unique stable network but profits and welfare are maximized at intermediate levels of collaboration. The authors argue that these results provide an explanation for why a large number of strategic alliances are unstable. They also argue that the complete network is consistent with individual incentives, finding it always to be strategically stable.

Goyal, Konovalov and Moraga-Gonzalez (2005) interpret R&D collaboration as research projects. The authors consider a hybrid form of R&D that combines in-house R&D and collaborative R&D with distinct partners. The authors show that the R&D efforts of linked firms are strategic complements, and that the efforts of unlinked firms are strategic substitutes. They find that the degree of collaborative activity increases firm R&D investment both for in-house research and for joint projects, and that profits are decreasing in the number of partners.

R&D collaborations with costly link formation is studied by Goyal and Joshi (2003) who consider an oligopoly in which firms form pairwise collaborative links that involve a commitment of resources, and yield lower costs of production for the collaborators. They analyze two cases in which the costs of forming links are either small or significant. They assume that a collaboration link between two firms involves a fixed cost and leads to an exogenously specified reduction in the marginal cost of production. They find that the complete network is the unique stable and socially efficient network under quantity competition. On the other hand, with price competition, the empty network is the unique stable network, while the efficient network is an inter-linked star, with two central firms.

In the *Co-Author Model,* Jackson and Wolinsky (1996) study the network structures that reward direct links while penalizing indirect links under a condition that requires sharing of limited resources. Focusing on the behavior of research collaborations, the authors suggest that if an individual's collaborator increases the time spent on other projects, the additional synergy arising from the time spent working together on a common project decreases. In this setup, each player has a fixed amount of time to spend on projects and the more projects a researcher is involved in, the lower the synergy that is obtained per project. They show the existence of a tension between efficiency and stability and suggest that the inefficiency of stable networks stems from the fact that (up to a point) a given player sees more benefit from adding a new link than harm in terms of dilution of the time he spends with his original partners, while those existing partners see only harm.

Taking the strategic models of network formation, R&D collaboration and the Co-Author Model as a point of departure, we study the incentives of firms to invest in R&D while operating under the conflicting forces of cooperation and competition. When firms collaborate, there are both knowledge spillovers between partners about cost-reducing technologies, and changes in the competitive positions of firms. These can lead to changes to both the market structure, and the performance of firms. Firms in our model create bilateral R&D collaborations and decide how much to invest in R&D in a given period and how to allocate these resources among the distinct projects. The information created in each project depends on the R&D investment devoted by the firm and its collaborator's R&D investment, which is divided between a number of projects. The institutional aspect that motivates our work is that there are a finite amount of resources a firm can invest on R&D and that synergies and inefficiencies might arise in collaborative environments. Therefore, we introduce network-based externalities and linking

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

costs to the model studied by Goyal and Moraga-Gonzalez (2001). We examine the architecture of the resulting network and its properties regarding stability and efficiency.

We find that when firms collaborate on R&D and operate in independent markets, individual R&D effort decreases in the number of collaborators when the congestion effect is high. Industry profits and welfare also decrease in the level of collaborative activity. Introducing product competition, we observe that, with a high congestion effect, firms' incentives to invest in R&D increase along with the connectivity in the network. If we consider the creation of synergies through working on multiple projects, firms choose to invest less in R&D as the number of partners increases. However, if the negative externalities are high, profits decrease as collaborative activity increases since the competition effect outweighs the benefits received from each collaboration. When the level of negative externalities is low, profits and social welfare increase in the degree of the network and are highest in the complete network. However, when the congestion effect is high, the complete network turns out to be inefficient. Finally, we show that under quantity competition, together with small linking costs, the congestion effect may give incentives to firms to deviate from the complete network by breaking links that would be stable with no negative externalities (as in Goyal and Moraga-Gonzalez (2001)), even with the same level of linking costs.

## 2.3 The Model

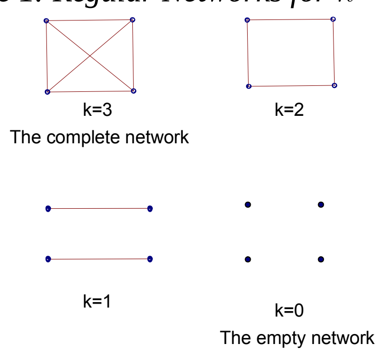
We adopt the three-stage game introduced by Goyal and Moraga-Gonzalez (2001). In the first stage, firms form costly pairwise collaboration links. In the second stage, each firm chooses a level of costly effort in R&D. The costs of firms are defined by their R&D efforts and by their collaboration links. Firms benefit from the R&D efforts of their collaborative partners. However, a firm's effort is shared between all of the projects on which he collaborates. In the last stage, firms compete in the market by setting quantities of the homogeneous goods.

### 2.3.1 *Collaboration Networks*

Let $N = \{1, 2, ..., n\}$ , $n \geq 3$ denote the set of firms. Each firm produces a homogeneous good. For any pair of firms $i, j \in N$, the pairwise relationship between the two firms is represented by a binary variable $g_{ij} \in \{0, 1\}$. When $g_{ij} = 1$, two firms are linked, while when $g_{ij} = 0$ there is no link between the firms. Links are assumed to be costly. A network $g = \{g_{ij}\}_{i,j \in N}$ is a collection of pairwise collaboration links between the firms. Let $g + g_{ij}$ denote the network obtained by adding a new link between firms $i$ and $j$, by replacing $g_{ij} = 0$ in network $g$ with $g_{ij} = 1$. Similarly, $g - g_{ij}$ is the network obtained by severing an existing link between firms $i$ and $j$ in network $g$ by replacing $g_{ij} = 1$ with $g_{ij} = 0$. Let $N_i(g)$ be the set of firms with which firm $i$ has a collaboration link in network $g$, which will be referred to as the set of $i$'s *neighbors*: $N_i = \{j \in N \setminus i : g_{ij} = 1\}$ and let $\eta_i(g)$ be the cardinality of the set $N_i(g)$.

We focus on regular networks in which $\eta_i(g) = \eta_j(g) = k$ for any two firms $i$ and $j$, and in which $k$ is the *degree* of the network or the *level of collaborative activity*. Figure 1 below illustrates some examples of regular networks with 4 firms.

*Figure 1. Regular Networks for* $n = 4$




In the complete network, all firms collaborate with each other. They benefit from each others' R&D efforts and compete in the homogenous-product market. The empty network captures the absence of collaboration links between the oligopolistic firms.

### 2.3.2 *The nature of cost-reducing R&D effort levels and spillovers*

Given a network $g$, each firm $i$ unilaterally chooses an R&D effort level $e_i(g)$. This individual effort lowers the firm's marginal cost and also has positive spillovers on the costs of the firm's collaborative partners.

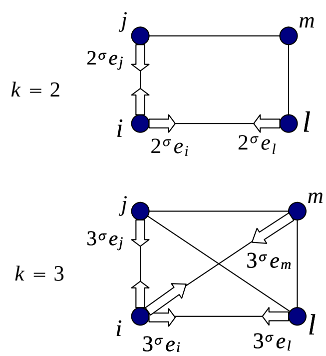
We assume that firms are initially symmetric, with zero fixed costs and identical constant marginal costs $c$. Given a network $g$ and the collection of effort levels $\{e_i(g)\}_{i \in N}$, the cost of firm $i$ is

$$c_i(\{e_i\}_{i \in N}) = c - e_i - \sum_{j \in N_i(g)} (\eta_j(g))^{\sigma} e_j$$

where $\sigma$ is a measure for the externalities arising from the degree of a firm's neighbor. When $\sigma = -1$, the benefits arising from a firm's effort are shared equally between all of its neighbors. $\sigma > 0$ can be thought of as the case in which there is synergy from collaboration. In this case, as the number of a firm's collaborators increases the return from its effort is higher than the actual effort. In particular, $\sigma = 0$ is the case in which there are no network-based externalities (as in Goyal and Moraga-Gonzalez (2001)).

In this chapter, we focus mainly on the equilibrium in which negative externalities arise due to the connectivity of a firm and $\sigma < 0$. In this case, the degree of a firm negatively affects its contribution to each particular collaboration. Figure 2 illustrates the total spillovers to collaborators when the individual effort is $\bar{e}$, and $k = 2, 3$.

*Figure 2. Total Spillovers to Collaborators*




| $\sigma = -2$ | $\sigma = -1$ | $\sigma = -0.5$ | $\sigma = 0$ |
|---|---|---|---|
| $\frac{\bar{e}}{2}$ | $\bar{e}$ | $\sqrt{2}\bar{e}$ | $2\bar{e}$ |
| $\frac{\bar{e}}{3}$ | $\bar{e}$ | $\sqrt{3}\bar{e}$ | $3\bar{e}$ |

We observe that when firms choose the R&D effort $\bar{e}$, the spillovers to each particular collaborator are decreasing in the degree of the firm for $\sigma < 0$. When $\sigma = 0$, all the collaborators of a firm benefit as much as the effort exerted by that firm, $\bar{e}$, and hence the total spillovers

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

to the collaborators increase as the degree of the firm increases, i.e., $2\bar{e}$ for $k = 2$ and $3\bar{e}$ for $k = 3$. When $\sigma = -1$, each collaborator of a firm benefits by $\bar{e}/2$ and $\bar{e}/3$ for $k = 2$ and $k = 3$, respectively, and the total level of spillovers remains constant. When $\sigma = -2$, inefficiencies arise due to degree of a firm. In this case, not only the benefit to each particular collaborator decreases, but also the total spillovers decrease as the number of collaborations increases.

We can also observe that the cost of production is decreasing in $\sigma$ and we assume that there are no knowledge spillovers between non-collaborating firms.

R&D effort is costly such that, given a level of effort $e_i \in [0, c]$, the cost is $Z(e_i) = \gamma e_i^2$, $\gamma > 0$. Under this specification, the cost of R&D effort is an increasing function that exhibits decreasing returns.

### 2.3.3 *Payoffs*

Given these costs, firms operate in the market by choosing quantities $\{q_i(g)\}_{i\in N}$. The demand is assumed to be linear and is given by $Q = a - p$, $a > c$. Finally, we introduce the cost of forming links, which is measured by the parameter $\beta$. Let $\pi_i(g)$ be the net profits attained by firm $i$ in network $g$. Therefore,

$$\pi_i(g) = [a - Q(g) - c_i(g)]\, q_i(g) - \gamma e_i^2 - \beta \eta_i(g)$$

### 2.3.4 *Welfare*

For any network $g$, social welfare is defined as the sum of consumer surplus and producers' profits. Let $W(g)$ denote the aggregate welfare in network $g$. Therefore, social welfare is:

$$W(g) = \frac{Q(g)^2}{2} + \sum_{i=1}^{N} \pi_i(g)$$

where $Q(g) = \sum\limits_{i\in N} q_i(g)$ is the aggregate output in network $g$ for the homogeneous-product oligopoly. A network $g$ is *efficient* if and only if $W(g) \geq W(g')$ for all $g'$.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

### 2.3.5 *Stability*

Following Jackson and Wolinsky (1996), a network $g$ is *stable* if and only if for all $i, j \in N$,

(i) if $g_{ij} = 1$, then $\pi_i(g) \geq \pi_i(g - g_{ij})$ and $\pi_j(g) \geq \pi_j(g - g_{ij})$

(ii) if $g_{ij} = 0$ and $\pi_i(g + g_{ij}) > \pi_i(g)$, then $\pi_j(g + g_{ij}) < \pi_j(g)$.

The idea is that, while a link can be severed unilaterally, a link can be formed if and only if the two firms involved agree to do so.

Given the setup of the model, in this chapter we focus on regular (symmetric) networks in which all firms have the same number of collaborators. In order to find the equilibrium levels of R&D effort and profits, we start by solving the last stage of the game, in which firms compete in quantities. We consider two cases as in the Goyal and Moraga-Gonzalez (2001) paper. In the first, firms operate in independent markets. In the second case, they compete in the homogeneous-product market. In the second stage of the game, firms choose their R&D efforts to maximize individual profits. We analyze how the parameters $\sigma$ and $\beta,$ which measure the congestion effect and the cost of linking, respectively, alter the equilibrium levels of R&D, profits, welfare, and the stability of the given networks.

## 2.4 Independent markets

In the market stage, given network $g$ and the R&D effort levels $\{e_i(g)\}_{i\in N}$, firms choose quantities to maximize their monopoly profits.

$$\max_{q_i} \pi_i(g) = [a - q_i(g) - c_i(g)]\, q_i(g) - \gamma e_i^2 - \beta \eta_i(g)$$

The quantity and profits coming from the last stage are

$$q_i(g) = [a - c_i(g)]/2$$

$$\pi_i(g) = \left[\frac{a - c_i(g)}{2}\right]^2 - \gamma e_i^2 - \beta \eta_i(g)$$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

In the second stage, firms choose their R&D efforts $\{e_i(g)\}_{i \in N}$.

$$\max_{e_i} \pi_i(g^k) = \frac{\left[a - c + e_i + \sum_{j \in N_i(g)} (\eta_j(g))^\sigma e_j\right]^2}{4} - \gamma e_i^2 - \beta \eta_i(g)$$

We first observe that the efforts of linked firms operating in independent markets are strategic complements since $\dfrac{\partial^2 \pi_i(g^k)}{\partial e_i \partial e_j} > 0$.

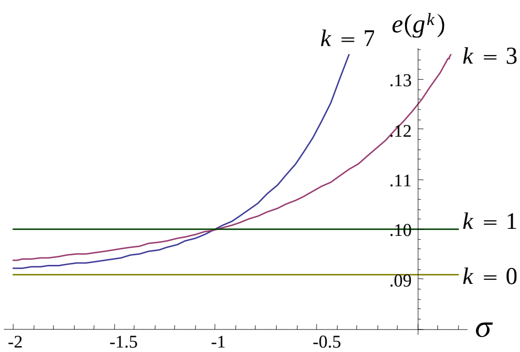
Focusing on symmetric networks, i.e., $\eta_i(g) = \eta_j(g) = k$, and solving for symmetric equilibrium, we obtain the equilibrium level of R&D effort and the cost[5]

$$e(g^k) = \frac{(a-c)}{(4\gamma - k^{1+\sigma} - 1)}; \qquad c(g^k) = c - \frac{(a-c)(k^{1+\sigma}+1)}{(4\gamma - k^{1+\sigma} - 1)}.$$

**Proposition 1** *When collaborating firms operate in independent markets, $e(g^k)$ is increasing (decreasing) and $c(g^k)$ is decreasing (increasing) in $k$ for $\sigma > -1$ ($\sigma < -1$).*
(see the Appendix for the proofs).

Figure 3 illustrates how the equilibrium level of effort changes with respect to $\sigma$ for $a = c + 1$ and $\gamma = 3$.

*Figure 3. Equilibrium Effort Level*




We observe that the level of the equilibrium R&D effort is lowest in the empty network, when $k = 0$ and that it increases (decreases resp.) with the level of connectivity for $\sigma > -1$ ($\sigma < -1$). Note that when $\sigma < -1$, the congestion effect offsets the cost reducing effect

[5] We assume that $\gamma > \max\{\frac{a(k^{1+\sigma}+1)}{4c}, \frac{k^{1+\sigma}+1}{4}\}$ so that the equilibrium levels of effort and costs are positive. This ensures that the second order condition, i.e., $\gamma > 1/4$ is satisfied as well.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

of R&D spillovers received from neighbors. As the degree of a firm increases, the spillovers received from neighbors' R&D efforts are lower. We know that the effort levels of linked firms are strategic complements, which leads to a reduction of the effort level of an individual firm. Therefore, the equilibrium level of R&D effort is lower for higher levels of collaboration. When $\sigma = -1$, the degree of the network does not play a role since the individual efforts are shared equally between all neighbors. The benefit from each collaborator is $\frac{\bar{e}}{k}$, and the total level of spillovers received from all neighbors remains constant in all regular networks (because $k \times \frac{\bar{e}}{k} = \bar{e}$), as we saw in Figure 2.

Finally, the equilibrium level of effort increases with $\sigma$ for all graphs, except for the empty network and the dyad, in which $k = 1$. In these graphs, the congestion effect does not play a role.

By substituting the equilibrium levels of effort and costs into the profit function obtained above, the equilibrium profits

$$\pi_i(g^k) = \frac{\gamma(4\gamma - 1)(a - c)^2}{(4\gamma - k^{1+\sigma} - 1)^2} - \beta k$$

**Proposition 2** *When collaborating firms operate in independent markets, given $\sigma > -1$, there exists a $\bar{\beta} > 0$ such that if $\beta \leq \bar{\beta}$, then $\pi(g^k)$ is increasing in $k$. $\pi(g^k)$ is decreasing in $k$ for $\sigma \leq -1$ and $\beta > 0$.*

**Stability and Welfare**

**Proposition 3** *The empty network is never stable for $\beta < \frac{\gamma(a-c)^2(8\gamma-3)}{4(2\gamma-1)^2(4\gamma-1)}$. Given $\sigma > -1$, there exists a $\bar{\beta} > 0$ such that, if $\beta \leq \bar{\beta}$, then the complete network is stable.*

Now, we can analyze the social welfare aspects of the equilibrium. Substituting the equilibrium levels of effort and cost, we obtain the equilibrium social welfare in a market

$$W(g^k) = \frac{(a - c)^2\gamma(6\gamma - 1)}{(4\gamma - k^{1+\sigma} - 1)^2} - \beta k$$

**Proposition 4** *When collaborating firms operate in independent markets, given $\sigma > -1$, there exists a $\breve{\beta} > 0$ such that if $\beta \leq \breve{\beta}$, then $W(g^k)$ is increasing in $k$. $W(g^k)$ is decreasing in $k$ for $\sigma \leq -1$ and $\beta > 0$.*

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Note that $\breve{\beta} - \bar{\beta} > 0$ for $\sigma > -1$, and for any $k$. This implies that there is a range of $\beta$ in which an increased level of collaborative activity results in both increased social welfare and decreased industry profits. We also observe that the complete network is not pairwise stable in this range.

## 2.5 Homogeneous-product oligopoly

In this section, we study R&D collaborations between firms that operate in the same market and analyze how competition affects the R&D incentives of firms and social welfare. In the last stage of the game, given network $g$ and the R&D effort levels $\{e_i(g)\}_{i \in N}$, firms compete by choosing quantities.

$$\underset{q_i}{Max}\ \pi_i(g) = \left[a - q_i(g) - \sum_{j \neq i} q_j(g) - c_i(g)\right] q_i(g) - \gamma e_i^2 - \beta \eta_i(g)$$

The (Cournot) equilibrium output is given by

$$q_i = \frac{a - c_i + \sum_{j \neq i}(c_j - c_i)}{(n+1)},$$

and the profits of the Cournot competitors are given by

$$\pi_i(g) = \left[\frac{a - c_i(g) + \sum_{j \neq i}(c_j - c_i)}{(n+1)}\right]^2 - \gamma e_i^2(g) - \beta \eta_i(g).$$

Now, we distinguish between three groups of firms:

(i) firm $i$

(ii) $k$ firms linked to firm $i$, denoted by $l$

(iii) $n - k - 1$ firms not linked to $i$, which we represent with $m$

Therefore, we have

$$\begin{aligned} c_i(g) &= c - e_i - k^{1+\sigma} e_l \\ c_l(g) &= c - e_l - k^{\sigma} e_i - (k-1)k^{\sigma} e_l \\ c_m(g) &= c - e_m - k^{1+\sigma} e_m \end{aligned}$$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

In the second stage of the game, firm $i$ will have the following profit function:

$$\pi_i(g) = \frac{\left[a - c_i + \sum_{l \in N_i}(c_l - c_i) + \sum_{m \notin N_i}(c_m - c_i)\right]^2}{(n+1)^2} - \gamma e_i^2 - \beta\eta_i$$

$$= \frac{[a - c_i + k(c_l - c_i) + (n-k-1)(c_m - c_i)]^2}{(n+1)^2} - \gamma e_i^2 - \beta k$$

Plugging in the costs and taking the derivative, we obtain first order condition:[6]

$$\frac{\partial \pi_i(g)}{\partial e_i} = \frac{(n - k^{1+\sigma})(a - c + e_i)}{(n+1)^2} + \frac{(n - k^{1+\sigma})(n-k)k^{1+\sigma}e_l}{(n+1)^2}$$

$$+\frac{(n - k^{1+\sigma})k(1 - k^{\sigma})(e_i - e_l)}{(n+1)^2}$$

$$+\frac{(n - k^{1+\sigma})(n-k-1)(e_i - (1 + k^{1+\sigma})e_m)}{(n+1)^2} - \gamma e_i = 0$$

This expression is useful because it reveals the different effects that govern the optimal R&D levels. The first one is called (as it is mentioned in the literature) *appropriability effect*. The larger is demand (or net demand, $a - c$), the stronger is the incentive to do R&D. This term decreases with $n$, and is highest when $n = 1$ ($k = 0$ in this case), that is, when there is a monopoly.

The second term can be thought of as the positive spillovers that the firm receives from its R&D collaborators. When the firm has no links $(k = 0)$, this term disappears completely. This term shows that the effort of the linked firms are strategic complements when the second derivative with respect to $e_l$ is positive. This happens when $n > k^{1+\sigma}$.[7]

The third term, which cancels out when $\sigma = 0$, captures the competition effect between collaborators. Together with the second term, it determines whether the collaborators' efforts are strategic complements or substitutes. In Goyal and Moraga-Gonzalez (2001), the absence

6 The second order condition requires $\gamma > \dfrac{(n - k^{1+\sigma})^2}{(n+1)^2}$.

7 For $\sigma < -1$, this condition is always satisfied, but for $\sigma > -1$, the network should be sparse enough not to violate this condition.

of this term means that the efforts of linked firms are always strategic complements. Here, we obtain

$$\frac{\partial^2 \pi_i(g)}{\partial e_i \partial e_l} = \frac{2(n-k^{1+\sigma})k^{1+\sigma}(n+1-k-k^{-\sigma})}{(n+1)^2}$$

which implies that the efforts become strategic substitutes in dense networks, i.e., when $n+1 < k + k^{-\sigma}$.

The fourth term is the pure competition effect between unlinked firms. Consequently, it will disappear when $k = n-1$ and when there is a monopoly ($n = 1$). We observe that

$$\frac{\partial^2 \pi_i(g)}{\partial e_i \partial e_m} = -\frac{2(n-k^{1+\sigma})(n-k-1)(k^{1+\sigma}+1)}{(n+1)^2} < 0$$

This tells us that the efforts of unlinked firms are strategic substitutes. Their efforts do not play a role in the independent market case.

Finally, the last term captures the marginal cost of R&D, and only depends on the efficiency parameter $\gamma$.

### Equilibrium Effort Level

Focusing on symmetric equilibrium, i.e., $e_i = e_l = e_m = e(g^k)$, we obtain the following equilibrium effort level for each firm:

$$e(g^k) = \frac{(a-c)(n-k^{1+\sigma})}{\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)}$$

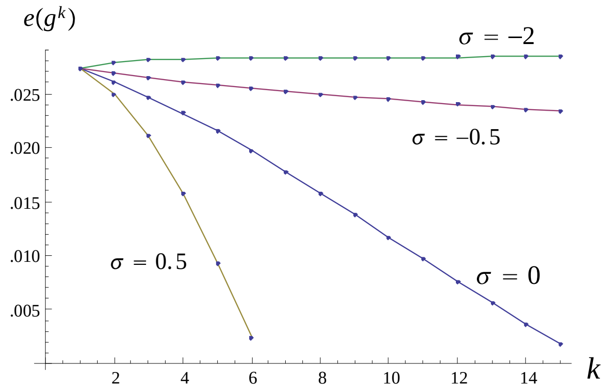
The first thing to analyze is how the equilibrium level of R&D effort changes in the level of collaborative activity $k$ :

**Proposition 5** *When collaborating firms compete in the homogeneous-product market, $e(g^k)$ is decreasing in the level of collaborative activity for $\sigma > -1$ and is increasing in $k$ for $k \geq 1$ when $\sigma < -1$. The highest $e(g^k)$ is attained in the empty network.*

**Proof** See the Appendix for proofs.

Figure 4 below illustrates the individual effort level in an industry with 16 firms for symmetric networks with different degrees of collaboration, $k$ (parameters are $a - c = 1$, $\gamma = 1$).

*Figure 4. Equilibrium Effort Level for n=16*




We observe that equilibrium effort level is decreasing in the level of collaborative activity for $\sigma > -1$. A particular case is when the congestion effect is absent, i.e., $\sigma = 0$ as in Goyal and Moraga-Gonzalez (2001). As the authors argue, a firm's effort reduces its own production costs, but also lowers the costs of collaborators, making them tougher competitors. As $k$ increases, more firms in the market benefit from a firm's effort. In addition, an increase in $k$ also implies lower costs for all firms. Therefore, the overall effect is of decreasing equilibrium effort levels.

As the graph suggests, when $\sigma > 0$, the decline is steeper since the aforementioned competition effects are stronger, and is weaker for low levels of negative externalities, i.e., $\sigma \in [-1, 0]$.

On the other hand, when $\sigma < -1$, an increase in $k$ reduces the benefits that a firm receives from its collaborators, thus giving the firm more incentives to invest in R&D. In addition, the firm's R&D effort creates negative spillovers to its neighbors which imply higher costs of production for all firms. Therefore, the overall effect is that R&D effort increases with the degree of the network.

The total R&D level in the industry is

$$R(g^k) = ne(g^k) = \frac{n(a-c)(n-k^{1+\sigma})}{\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)}$$

Note that

$$\frac{\partial R(g^k)}{\partial k} = -\frac{(a-c)nk^{\sigma}[\gamma(n+1)^2 - (n-k^{1+\sigma})^2](1+\sigma)}{[\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)]^2}$$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

This expression tells us that when $\sigma > -1$ ($\sigma < -1$ resp.), the more connected is the industry, the smaller is (larger is) the amount of R&D carried out. Calculations show that $R(g^k)$ decreases with $\sigma$ and that the total R&D level increases with $n$ (that is, with more competition in the market). However, this does not necessarily mean that welfare increases with $n$. In the following sections we analyze the optimal amount of resources to be devoted to R&D.

**Cost Reduction**

Substituting the equilibrium effort levels into the cost functions, we obtain the equilibrium level of costs

$$c(g^k) = \frac{c\gamma(n+1)^2 - a(n-k^{1+\sigma})(k^{1+\sigma}+1)}{\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)}$$

We calculate $c(g^k) - c(g^{k+1})$, and we can state the following result.

**Proposition 6** *When collaborating firms compete in the homogeneous-product market, costs are increasing in the level of collaborative activity for $k \geq 1$, when $\sigma < -1$. The relationship between cost reduction and the level of collaborative activity is non-monotonic for $\sigma > -1$. $c(g^k)$ decreases in $k$ for $(k+1)^{1+\sigma} + k^{1+\sigma} < n-1$. Production costs are highest in the empty network and in the complete network when $\sigma = 0$.*

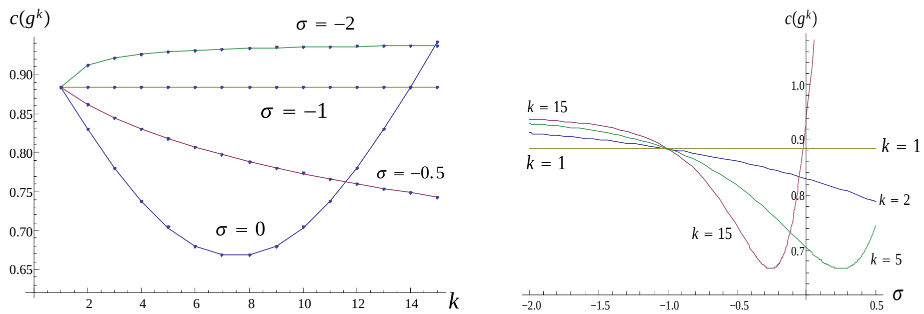
Substituting $\sigma = 0$ as in Goyal and Moraga-Gonzalez (2001), we obtain:

$$c(g^k) - c(g^{k+1}) = \frac{(a-c)\gamma(n+1)^2(n-2k-2)}{[\gamma(n+1)^2 - (n-k)(k+1)][\gamma(n+1)^2 - (n-k-1)(k+2)]}$$

and this suggests that an increase in the level of collaborations reduces the cost of firms if and only if $k < n/2 - 1$. However, when we introduce the congestion effect, which is measured by $\sigma$, the condition becomes $(k+1)^{1+\sigma} + k^{1+\sigma} < n-1$ and we observe that the cost reduction is highest at a greater connectivity level for $\sigma \in (0, -1)$.

Moreover, we do not find the non-monotonic relationship between the cost reduction and the level of collaborative activity for $\sigma < -1$. In this case, costs are always increasing with connectivity. We obtain the following figures for the parameter values $a - c = 1$, $\gamma = 1$.

*Figure 5. Equilibrium Cost Level for n=16*




**Equilibrium Profits**

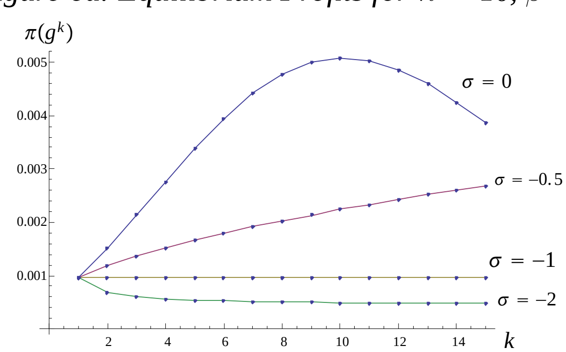
Finally, substituting the equilibrium levels of effort and costs, we obtain the profit of a firm in a symmetric network

$$\pi_i(g^k) = \frac{(a-c)^2\gamma[\gamma(n+1)^2 - (n-k^{1+\sigma})^2]}{[\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)]^2} - \beta k$$

**Proposition 7** *When collaborating firms compete in the homogeneous-product market, profits are decreasing in the level of collaborative activity when* $\sigma < -1$.

The following figure depicts firms' profits in an industry with 16 firms with different degrees of collaboration $k$ and with no linking costs, i.e., $\beta = 0$.

*Figure 6a. Equilibrium Profits for* $n = 16$, $\beta = 0$



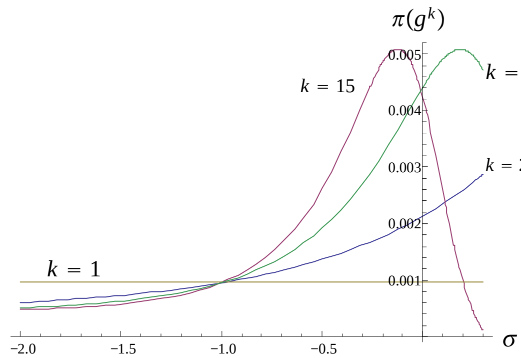

As discussed in Goyal and Moraga-Gonzalez (2001), when $\sigma = 0$, profits are maximized at an intermediate level of collaborative activity. The reason is that, as the number of links rises, firm performance improves because the effects of lower research efforts are offset by lower costs that result from the sharing of R&D. For high values of $k$, the impact of lower

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

efforts dominates and profits decline. In our model, as seen in the figures above, for levels of $\sigma$ that are close to $0$, we obtain the same result, but the connectivity level that maximizes profits is higher, since the decline in effort is slower. Profits decline at higher levels of congestion because both the effort levels and costs are higher.

On the other hand, as we see in Figure 6a, when $\sigma < -1$, profits are strictly decreasing as the number of connections increases. This result follows from the previous propositions. As the degree of the symmetric network increases, equilibrium effort levels increase and cost savings from R&D sharing are offset by the negative externalities. As a result, industry profits decline.

*Figure 6b. Equilibrium Profits for* $n = 16$, $\beta = 0$




Intuitively, as a firm's neighbor form larger number of collaborative links, the neighbor's research effort will be divided between a greater number of partners. The firm will receive a smaller proportion of its neighbor's cost reducing R&D, which will result in lower profits. This result is robust to the level of $\gamma$, the cost of effort.

This result follows from the finding that an increase in the level of $\sigma < -1$ increases the effort level for networks for all possible collaborative activity levels. In addition, we have shown that the cost level also increases with the level of the "co-author effect" for all but the complete network. Therefore, we observe a decline in profits as $\sigma < -1$ increases in absolute terms. For small values of $\sigma$, we have shown that the effort level and the cost level are decreasing, which suggests that profits are increasing as $\sigma$ increases, which is consistent with the previous findings.

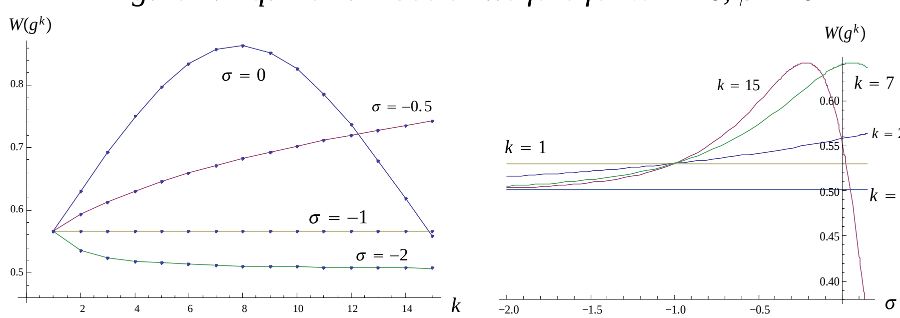
### Welfare

We can now analyze the social welfare aspects of the equilibrium. Substituting the equilibrium levels of effort and cost, we obtain the equilibrium social welfare

$$W(g^k) = \frac{(a-c)^2 n\gamma[\gamma(n+1)^2(n+2) - 2(n-k^{1+\sigma})^2]}{2[\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)]^2} - \beta k$$

**Proposition 8** *Suppose that collaborating firms compete in quantities in the homogeneous-product market. Welfare decreases with the degree of the network for* $\sigma < -1$.

*Figure 7. Equilibrium Social Welfare for* $n = 16$, $\beta = 0$




We observe that for low levels of $\sigma$, the relationship between social welfare and the number of collaborators is non-monotonic. In addition, optimal social welfare requires a level of collaboration that is less than the level that maximizes industry profits. Thus, for low levels of sigma the discrepancy between the optimal levels of R&D for social welfare and for industry profits exists.

### Stability

We can now analyze the stable configurations given this framework.

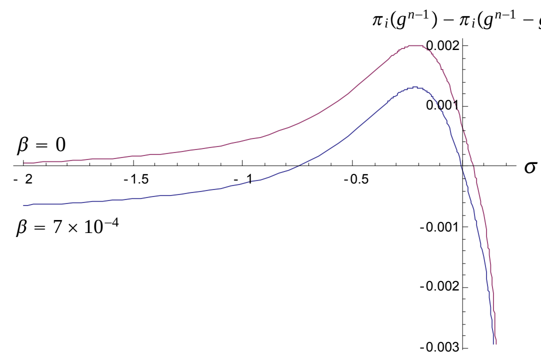
**Proposition 9** *Suppose that collaborating firms compete in the homogeneous-product market. There exists a* $\tilde{\beta} > 0$ *such that, if* $\beta \leq \tilde{\beta}$, *then the empty network is never pairwise stable. The complete network is always pairwise stable when* $\beta = 0$. *When creating links is costly, for* $\sigma < -1$, *there exist pairwise stable networks with* $0 < k < n-1$.

As an example, if we take $a = c + 1$ and $n = 16$, we find that that for $\beta < 6.84 \times 10^{-4}$, the complete network is stable. Using the same values, we observe that if $\sigma < -0.75$, then the

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

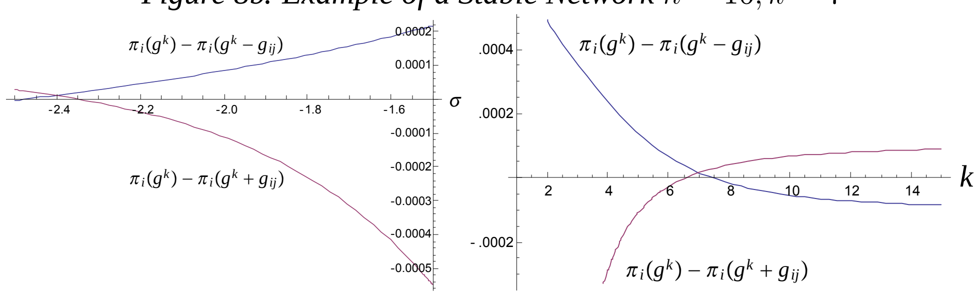
deviation profits are higher than the profits in the complete network. The figure below illustrates the difference in profits $\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij})$ for $\beta = 0$ and $\beta = 7 \times 10^{-4}$ when $n = 16$.

*Figure 8a. Stability of the Complete Network* $n = 16$




For $\beta = 1 \times 10^{-4}$, we find that the complete network is pairwise stable when $\sigma > -1.75$ and $n = 16$. From the calculations given in the appendix, we obtain the following figure for $k = 7$ in which $\pi_i(g^k) - \pi_i(g^k - g_{ij}) \geq 0$ and $\pi_i(g^k) - \pi_i(g^k + g_{ij}) \geq 0$. This implies that there exist networks with intermediate levels of collaborations that are pairwise stable.

*Figure 8b. Example of a Stable Network* $n = 16, k = 7$




We observe that for $\sigma = -2.4$, the equilibrium profits are higher than the deviation profits (profits after breaking a link and after adding a link). Therefore the network with $k = 8$ is pairwise stable. This example shows that an intermediate level of collaboration can be pairwise stable, and that the complete network is not stable when there is high congestion effect.

## 2.6 Conclusion

This chapter builds upon the existing strategic models of network formation and R&D collaboration. We introduce the institutional aspect that has been disregarded in the previous

 DOI: 10.2870/63947

literature. We introduce network-based externalities in order to account for empirical evidence, which suggests an inverted-U relationship between the number of collaborators and the R&D intensity of firms. We analyze the trade-off between the positive spillovers arising from research collaborations between firms that compete in product markets and the negative externalities resulting from their indirect connections. We take the Co-Author Model as a point of departure, and focus on the inefficiencies, referred to as congestion effects, that arise in collaborative environments and study their impact on R&D investments and industry profits. We examine the architecture of the resulting network, and describe its stability and efficiency properties.

We find that when firms collaborate on R&D and operate in independent markets, individual R&D effort decreases in the number of collaborators when the congestion effect is high. Industry profits and welfare also decrease in the level of collaborative activity. Introducing product competition, we observe that with high congestion effect, firms' incentives to invest in R&D increase with the connectivity of the network. Industry profits decrease as the collaborative activity increases since the competition effect outweighs the benefits received from each collaboration. With a very low level of negative externalities, profits and social welfare increase in the degree of the network and are highest in the complete network. However, when the congestion effect is high, the complete network turns out to be inefficient. Finally, we show that, when we introduce small linking costs, the congestion effect may give incentives to firms to deviate from the complete network by breaking links that, if there were no externality, would otherwise be viable. In this case, we also show that networks with intermediate levels of collaboration can be sustained as pairwise stable.

For further research, one could analyze the relationship between market competition, firms' incentives to invest in R&D, and the architecture of collaboration networks when the market is characterized by differentiated products, and study how the degree of horizontal product differentiation alters the outcome. It might also be interesting to check the results when firms have price compete as opposed to quantity competition that is considered in this model.

In addition, following Deroian and Gannon (2006), one could study rival firms' incentives when R&D is quality improving rather than cost reducing and when the synergies and the congestion effect still exist.

Last but not least, complementing the theoretical analysis in this chapter with empirical evidence by collecting firm-level R&D collaboration data would be an important contribution to the empirical literature.

## 2.7 Appendix

Proof of Proposition 1. We can calculate

$$e(g^{k+1}) - e(g^k) = \frac{(a-c)[(k+1)^{1+\sigma} - k^{1+\sigma}]}{(4\gamma - k^{1+\sigma} - 1)(4\gamma - (k+1)^{1+\sigma} - 1)}$$

$$c(g^{k+1}) - c(g^k) = -\frac{(a-c)4\gamma[(k+1)^{1+\sigma} - k^{1+\sigma}]}{(4\gamma - k^{1+\sigma} - 1)(4\gamma - (k+1)^{1+\sigma} - 1)}$$

We observe that for $\sigma > -1$, $k > 0$, $(k+1)^{1+\sigma} > k^{1+\sigma}$ and the equilibrium R&D effort increases and costs decrease in the level of collaborative activity.

Proof of Proposition 2. We can calculate

$$\pi(g^{k+1}) - \pi(g^k) = \frac{\gamma(4\gamma-1)(a-c)^2[(k+1)^{1+\sigma} - k^{1+\sigma}][2(4\gamma-1) - (k+1)^{1+\sigma} - k^{1+\sigma}]}{[(4\gamma - k^{1+\sigma} - 1)(4\gamma - (k+1)^{1+\sigma} - 1)]^2} - \beta$$

In order to have a non-zero denominator and non-negative effort levels it must be that $4\gamma - 1 > k^{1+\sigma}$ and $4\gamma - 1 > (k+1)^{1+\sigma}$. Summing up these inequalities, we obtain $2(4\gamma-1) > (k+1)^{1+\sigma} + k^{1+\sigma}$, which ensures that the last part of the numerator is positive. Hence, the fraction is positive so long as $(k+1)^{1+\sigma} > k^{1+\sigma}$ which is true for $\sigma > -1$.

Proof of Proposition 3.
The equilibrium profits obtained in the empty network can be calculated by substituting $k = 0$ into $\pi(g^k)$ :

$$\pi(g^{empty}) = \frac{\gamma(a-c)^2}{(4\gamma-1)}$$

When two firms deviate and form a link, the profits they collect can be calculated simply by substituting $k = 1$ into the equilibrium profit since the reaction functions depend only on the effort chosen by the collaborators and not by the other firms'. It becomes:

$$\pi(g^{dev}) = \frac{\gamma(4\gamma-1)(a-c)^2}{4(1-2\gamma)^2} - \beta$$

We can calculate

$$\pi(g^{dev}) - \pi(g^{empty}) = \frac{\gamma(a-c)^2(8\gamma-3)}{4(2\gamma-1)^2(4\gamma-1)} - \beta$$

We know that $\gamma > 1/2$ in an interior solution. Therefore, in an empty network, firms strictly increase their profits by forming a link when

$$\beta < \frac{\gamma(a-c)^2(8\gamma-3)}{4(2\gamma-1)^2(4\gamma-1)}$$

Finally, the stability of the complete network follows from Proposition 2, which states that profits increase in the degree of the network for $\sigma > -1$ and $\beta < \bar{\beta}$. This becomes

$$\bar{\beta} = \frac{\gamma(4\gamma-1)(a-c)^2[(n-1)^{1+\sigma} - (n-2)^{1+\sigma}][2(4\gamma-1) - (n-1)^{1+\sigma} - (n-2)^{1+\sigma}]}{[(4\gamma - (n-2)^{1+\sigma} - 1)(4\gamma - (n-1)^{1+\sigma} - 1)]^2}$$

when we replace $k = n-2$ in $\pi(g^{k+1}) - \pi(g^k)$.

Proof of Proposition 4. The equilibrium quantity can be calculated as

$$\begin{aligned} q_i(g^k) &= \frac{a - c_i(g^k)}{2} \\ &= \frac{2\gamma(a-c)}{4\gamma - k^{1+\sigma} - 1} \end{aligned}$$

Therefore, social welfare in one market is

$$\begin{aligned} W(g^k) &= \frac{q_i(g)^2}{2} + \pi_i(g) \\ &= \frac{(a-c)^2\gamma(6\gamma-1)}{(4\gamma - k^{1+\sigma} - 1)} - \beta k \end{aligned}$$

We can calculate

$$W(g^{k+1}) - W(g^k) = \frac{\gamma(6\gamma-1)(a-c)^2[(k+1)^{1+\sigma} - k^{1+\sigma}][2(4\gamma-1) - (k+1)^{1+\sigma} - k^{1+\sigma}]}{[(4\gamma - k^{1+\sigma} - 1)(4\gamma - (k+1)^{1+\sigma} - 1)]^2} - \beta$$

Similar to the proof of Proposition 2, we can observe that $W(g^{k+1}) - W(g^k) > 0$ when $\sigma > -1$ and $\beta < \breve{\beta}$.

Proof of Proposition 5. We can calculate

$$e(g^k) - e(g^{k+1}) = \frac{(a-c)[\gamma(n+1)^2 - (n - k^{1+\sigma})(n - (k+1)^{1+\sigma})] \times [(k+1)^{1+\sigma} - k^{1+\sigma}]}{[\gamma(n+1)^2 - (n - k^{\sigma+1})(k^{\sigma+1}+1)]\,[\gamma(n+1)^2 - (n - (k+1)^{\sigma+1})((k+1)^{\sigma+1}+1)]}$$

From the second order condition, the term $[\gamma(n+1)^2 - (n - k^{1+\sigma})(n - (k+1)^{1+\sigma})]$ in the numerator is positive. The denominator is positive since $n \geq (k+1)^{1+\sigma}$. The term $[(k+1)^{1+\sigma} - k^{1+\sigma}]$ determines the sign of the difference, which depends on the level of $\sigma$. We observe that when $\sigma > -1$ ($\sigma < -1$ resp.), R&D effort is declining (increasing) in the level of collaborative activity.

Proof of Proposition 6. We calculate

$$c(g^k) - c(g^{k+1}) = \frac{(a-c)\gamma(n+1)^2[n - ((k+1)^{1+\sigma} + k^{1+\sigma}) - 1] \times [(k+1)^{1+\sigma} - k^{1+\sigma}]}{[\gamma(n+1)^2 - (n - k^{1+\sigma})(k^{1+\sigma}+1)][\gamma(n+1)^2 - [n - (k+1)^{1+\sigma}][(k+1)^{1+\sigma}+1)]]}$$

The denominator is positive using the same logic as in the proof of Proposition 5. When $\sigma > -1$, the last term in the numerator is also positive. Thus, the sign of the difference depends on the term

$$[n - ((k+1)^{1+\sigma} + k^{1+\sigma}) - 1]$$

It is positive for $(k+1)^{1+\sigma} + k^{1+\sigma} < n-1$. Thus, it is non-monotonic in $k$. Costs decrease in $k$ for a low level of collaborative activity, but start to increase after a certain level of $k$.

On the other hand, when $\sigma < -1$, the term $[(k+1)^{1+\sigma} - k^{1+\sigma}]$ is negative. In addition, it is always true that $(k+1)^{1+\sigma} + k^{1+\sigma} < n - 1$ for $n > 1$. Therefore, costs always increase in collaborative activity for $\sigma < -1$.

We know that

$$c(g^k) = \frac{c\gamma(n+1)^2 - a(n - k^{1+\sigma})(k^{1+\sigma}+1)}{\gamma(n+1)^2 - (n - k^{1+\sigma})(k^{1+\sigma}+1)}$$

We can observe that the term $(n - k^{1+\sigma})(k^{1+\sigma}+1)$ equals $n$ for the complete network ($k = n-1$) when $\sigma = 0$ and for the empty network, i.e., $k = 0$. The highest cost is obtained in these networks.

Proof of Proposition 7.
We first need to show that $\pi_i^{dev} = \pi_i(g^k) - \pi_i(g^{k+1}) > 0$ for $\sigma < -1$.

$$\pi_i^{dev} = (a-c)\gamma \left[ \frac{\gamma(n+1)^2 - (n-k^{1+\sigma})^2}{[\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)]^2} - \frac{\gamma(n+1)^2 - (n-(k+1)^{1+\sigma})^2}{[\gamma(n+1)^2 - (n-(k+1)^{1+\sigma})((k+1)^{1+\sigma}+1)]^2} \right] + \beta$$

When $\sigma < -1$, we know that $k^{1+\sigma} > (k+1)^{1+\sigma}$. Thus, $(n-k^{1+\sigma})^2 < (n-(k+1)^{1+\sigma})^2$, which implies that the numerator of the first fraction is greater than the second. In order to compare the denominators we show that
$(n-k^{1+\sigma})(k^{1+\sigma}+1) - [(n-(k+1)^{1+\sigma})((k+1)^{1+\sigma}+1)] > 0$ for $\sigma < -1$.
After simple calculations, we can re-write that term as
$(n-k^{1+\sigma})(k^{1+\sigma}+1) - [(n-(k+1)^{1+\sigma})((k+1)^{1+\sigma}+1)] = \left[k^{1+\sigma} - (k+1)^{1+\sigma}\right]\left[n-1-k^{1+\sigma}-(k+1)^{1+\sigma}\right]$
We observe that both terms are positive for $\sigma < -1$. Therefore, the denominator of the first fraction is smaller than the second. We can state that for $\sigma < -1$, profits decrease in $k$ since $\pi_i(g^k) - \pi_i(g^{k+1}) > 0$.
Second, we know that when $\sigma = -1$, the difference becomes $\pi_i(g^k) - \pi_i(g^{k+1}) = \beta$. Thus, without costly link formation, profits become independent of the level of collaborative activity.
Finally, we observe that the qualitative analysis is different for different values of $\sigma$ when $\sigma > -1$. Following the proof in Goyal and Moraga-Gonzalez (2001), we know that the relationship between profits and the collaborative activity is non-monotonic for $\sigma = 0$. For $\sigma < 0$, that is, close to zero, we find the same relationship. As $\sigma$ gets closer to $-1$, we observe that profits increase in the level of collaborative activity.

Proof of Proposition 8.
We need to show that $W(g^k) - W(g^{k+1}) > 0$ for $\sigma < -1$. We have found that

$$W(g^k) = \frac{(a-c)^2 n\gamma[\gamma(n+1)^2(n+2) - 2(n-k^{1+\sigma})^2]}{2[\gamma(n+1)^2 - (n-k^{1+\sigma})(k^{1+\sigma}+1)]^2} - \beta k$$

Similar to the proof of Proposition 7, for $\sigma < -1$ the numerator is greater and the denominator is smaller in $W(g^k)$ than in $W(g^{k+1})$. Therefore $W(g^k) - W(g^{k+1}) > 0$.

We also observe in Mathematica that, when forming links is cost-free, for $n = 16$, $a = c+1$, $\gamma = 1$, and $\sigma \in [-0.25, -1)$, social welfare is increasing in the level of collaborative activity and is highest in the complete network. For $\sigma > -0.25$, there is an intermediate level of collaborative activity at which social welfare is maximized, as found in Goyal and Moraga-Gonzalez (2001) for $\sigma = 0$.

Proof of Proposition 9.
(I) Stability of the empty network:

When $k = 0$, we can calculate the profits in the empty network[8]

$$\pi_i(g^e) = \frac{(a-c)^2\gamma[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)^2 - n]^2}$$

In order to check for the stability of the empty network, we need to check whether any two firms, say $i$ and $j$, have an incentive to form an additional link, resulting in the network $g^e + g_{ij}$. Here, the externalities from indirect links do not play a role since in the empty network $k = 0$, and the deviating firms will have $k = 1$, therefore, $\sigma$ does not effect the equilibrium results of Goyal and Moraga-Gonzalez (2001), who find that the empty network is never stable. Since we will not get any other interesting result by adding linking costs (which could only make the empty network stable above a certain threshold), we simply assume that $\beta = 0$.
The profit of firms $i$ and $j$ will be

$$\pi_i(g^e + g_{ij}) = \frac{[a-c+(n-1)(e_i+e_j) - (n-2)e_m]^2}{(n+1)^2} - \gamma e_i^2$$

8 When $\sigma = 1$, $\pi_i(g^e) = \dfrac{(a-c)^2\gamma[\gamma(n+1)^2 - (n-1)^2]}{[\gamma(n+1)^2 - 2(n-1)]^2}$

The first-order condition is $(n-1)[a-c+(n-1)(e_i+e_j)-(n-2)e_m]-\gamma(n+1)^2e_i=0$. As in Goyal and Moraga-Gonzalez (2001), the R&D efforts of firms that form collaborations become strategic complements.
The profit of firms that do not have any links in the network $g^e+g_{ij}$ is

$$\pi_l(g^e+g_{ij})=\frac{[a-c+nel-2(e_i+e_j)-(n-3)e_m]^2}{(n+1)^2}-\gamma e_l^2$$

The first order condition is $n[a-c+nel-2(ei+ej)-(n-3)e_m]-\gamma(n+1)^2e_l=0$. Invoking symmetry, i.e., $e_i=e_j$ and $e_l=e_m$, we obtain the best-response functions of the firms

$$\begin{aligned} e_i &= \frac{(n-1)(a-c-(n-2)e_l)}{\gamma(n+1)^2-3n} \\ e_l &= \frac{n(a-c-4e_i)}{\gamma(n+1)^2-3n} \end{aligned}$$

For $\gamma=1$, there is a corner solution in which $e_l=0$ and $e_i=\dfrac{(n-1)(a-c)}{n(6-n)-1}$if $n=4$. The deviating profits of firm $i$ become:

$$\pi_i(g^e+g_{ij})=\frac{(a-c)^24n}{[n(6-n)-1]^2}-\beta>\pi_i(g^e)=\frac{(a-c)^2(2n+1)}{[n^2+n+1]^2}$$

We can see that the deviation profits are higher than the equilibrium profits. Therefore, the empty network is never stable as long as $\beta$ is sufficiently small.

(II) Stability of the complete network:
When $k=n-1$, we can derive the profits in the complete network

$$\pi_i(g^{n-1})=\frac{(a-c)^2\gamma[\gamma(n+1)^2-(n-(n-1)^{1+\sigma})^2]}{[\gamma(n+1)^2-(n-(n-1)^{1+\sigma})((n-1)^{1+\sigma}+1)]^2}-\beta(n-1)$$

In order to check for the stability of the complete network, we need to check whether any two firms, say $i$ and $j$, have an incentive to break an existing link. In the resulting network $g^{n-1}-g_{ij}$, firms $i$ and $j$ will have $(n-2)$ neighbors, and the remaining $(n-2)$ firms will have $(n-1)$ neighbors
The costs of firms $i$ and $j$ will be $c_i=c-e_i-(n-2)(n-1)^\sigma e_m$ and $c_j=c-e_j-(n-2)(n-1)^\sigma e_m$, while the cost of the remaining firms will be $c_l=c-e_l-(n-2)^\sigma e_i-(n-2)^\sigma e_j-(n-3)(n-1)^\sigma e_m$.
The profit function becomes,

$$\pi_i(g^{n-1}-g_{ij})=\frac{\left[a-c+(n-(n-2)^{\sigma+1})e_i-(1+(n-2)^{\sigma+1})e_j+(2(n-1)^\sigma-1)(n-2)e_m\right]^2}{(n+1)^2}-\gamma e_i^2-\beta(n-2)$$

The first order condition is $(n-(n-2)^{\sigma+1})[a-c+(n-(n-2)^{\sigma+1})e_i-(1+(n-2)^{\sigma+1})e_j+(2(n-1)^\sigma-1)(n-2)e_m]-\gamma(n+1)^2e_i=0$.
We can see that the R&D efforts of firms breaking collaboration links become strategic substitutes.

$$\pi_l(g^{n-1}-g_{ij})=\frac{\left[a-c+(n-(n-1)^{\sigma+1})e_l+(3(n-2)^\sigma-1)(e_i+e_j)+(2(n-1)^\sigma-1)(n-3)e_m\right]^2}{(n+1)^2}-\gamma e_l^2-\beta(n-1)$$

The first order condition is $(n-(n-1)^{\sigma+1})[a-c+(n-(n-1)^{\sigma+1})e_l+(3(n-2)^\sigma-1)(e_i+e_j)+(2(n-1)^\sigma-1)(n-3)e_m]-\gamma(n+1)^2e_i=0$.
Invoking symmetry,

$$\begin{aligned} e_i &= \frac{(n-(n-2)^{\sigma+1})\left[a-c+(2(n-1)^\sigma-1)(n-2)e_l\right]}{\gamma(n+1)^2-(n-(n-2)^{\sigma+1})(n-2(n-2)^{\sigma+1}-1)} \\ e_l &= \frac{(n-(n-1)^{\sigma+1})\left[a-c+(6(n-2)^\sigma-2)e_i\right]}{\gamma(n+1)^2-(n-(n-1)^{\sigma+1})(3+(n-5)(n-1)^\sigma)} \end{aligned}$$

Solving these two equations we obtain the equilibrium efforts

$$e_i = e_j = \frac{(a-c)(n-(n-2)^{\sigma+1})[\gamma(n+1)-n-(n-1)^{2\sigma+1}+(2n-1)(n-1)^{\sigma}]}{\Lambda}$$

$$e_l = e_m = \frac{(a-c)(n-(n-1)^{\sigma+1})[\gamma(n+1)-n-2(n-2)^{2\sigma+1}+(3n-2)(n-2)^{\sigma}]}{\Lambda}$$

where $\Lambda = 2(n-2)^{\sigma}n + n^2 - (n-2)^{\sigma}n^2 + 2(n-2)^{\sigma}\gamma - 8(n-2)^{2\sigma}\gamma - 2n\gamma - 5(n-2)^{\sigma}n\gamma - 3n^2\gamma - 4(n-2)^{\sigma}n^2\gamma + 6(n-2)^{2\sigma}n^2\gamma - n^3\gamma + 3(n-2)^{\sigma}n^3\gamma - 2(n-2)^{2\sigma}n^3\gamma + \gamma^2 + 3n\gamma^2 + 3n^2\gamma^2 + +n^3\gamma^2 - (n-1)^{1+2\sigma}$ $[2(n-2)^{2+2\sigma} - 3(n-2)^{1+\sigma}(n-1) - 3n + n^2 + 5\gamma + 4n\gamma - n^2\gamma] + (n-1)^{\sigma}[n + 2(n-2)^{2+2\sigma}n - 4n^2 + n^3 - (n-2)^{1+\sigma}(n-1)(3n-1) - 3\gamma + 5n\gamma + 7n^2\gamma - n^3\gamma]$

Plugging these into the profit function we get

$\pi_i(g^{n-1} - g_{ij}) = \frac{(a-c)^2\gamma[\gamma(n+1)^2 - n^2 + 2n(n-2)^{1+\sigma} - (n-2)^{2+2\sigma}][\gamma(n+1) - n + (n-1)^{\sigma}(2n-1) - (n-1)^{1+2\sigma}]^2}{\Lambda^2} - \beta(n-2)$

First, we observe that when $\sigma \to -\infty$ , $\pi_i(g^{n-1} - g_{ij}) \to \dfrac{(a-c)^2\gamma[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)-n]^2} - \beta(n-2)$ and $\pi_i(g^{n-1}) \to \dfrac{(a-c)^2\gamma[\gamma(n+1)^2 - n^2]}{[\gamma(n+1)-n]^2} - \beta(n-1)$. Therefore, when forming links is costless, i.e., $\beta = 0$, the complete network is always stable, even when the negative externalities are very high. Introducing linking costs, i.e., $\beta > 0$, the deviation profits $\pi_i(g^{n-1} - g_{ij})$ become higher than the profits under the complete network, i.e., $\pi_i(g^{n-1})$, for sufficiently low $\sigma$.

$\sigma = 0$ and $\beta = 0$ nest the setup studied by Goyal and Moraga-Gonzalez (2001), who show that the complete network is the unique stable network. We will now find the maximum $\beta$ for which the complete network is stable for the case in which there are no negative externalities ($\sigma = 0$) while there exists $\bar{\sigma} < 0$ such that if $\sigma \leq \bar{\sigma}$, then $\pi_i(g^{n-1} - g_{ij}) \geq \pi_i(g^{n-1})$. In other words we will find the the $\bar{\sigma} < 0$ for which the complete network is not pairwise stable.

In order to do so, we need to calculate the difference $\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij})$ when $\sigma = 0$ and find the maximum $\beta$ for which it is nonnegative. Taking $\gamma = 1$ for simplicity, we obtain

$$\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij}) = \frac{(a-c)^2[4n^7 + 15n^6 - 4n^5 - 20n^4 + 30n^3 + 8n^2 + 12n + 3]}{(n^2+n+1)^2(n^3+4n^2-2n+1)^2} - \beta$$

Therefore, when $\beta \leq \dfrac{(a-c)^2[4n^7 + 15n^6 - 4n^5 - 20n^4 + 30n^3 + 8n^2 + 12n + 3]}{(n^2+n+1)^2(n^3+4n^2-2n+1)^2}$, the complete network is always stable. As numerical example, if we take $a = c + 1$ and $n = 16$, the condition becomes $\beta \leq 0.000683896$. Using the same values, we observe that if $\sigma < -0.75$, then the deviation profits are higher than the profits under the complete network. In Figure 8a, we have shown the difference in profits $\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij})$ for $\beta = 0$ and $\beta = 7 \times 10^{-4}$ when $n = 16$.

We will now show that we can find some intermediate level of collaboration that is pairwise stable for the same parameter values, for the case in which there are no negative externalities, the complete network is the unique stable network.

As an example, for $\beta = 1 \times 10^{-4}$, we find that the complete network is pairwise stable when $\sigma > -1.75$ and $n = 16$. Thus, if we can find a $k < 15$ such that $\pi_i(g^k) - \pi_i(g^k - g_{ij}) \geq 0$ and $\pi_i(g^k) - \pi_i(g^k + g_{ij}) \geq 0$, it is implied that there exist networks with intermediate level of collaborations that are pairwise stable.

Using Mathematica, we were able to calculate $\pi_i(g^k) - \pi_i(g^k - g_{ij})$ and $\pi_i(g^k) - \pi_i(g^k + g_{ij})$ for any $k$. For illustrative purposes we do not include the analytical solutions, which are very long, but we present in Figure 8b an example which shows that we can find intermediate levels of connectivity that are stable.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

## 2.8 References

Becker, W. and Dietz, J. (2003). "R&D Cooperation and Innovation Activities of Firms: Evidence for the German Manufacturing Industry." *Research Policy*, 33, 2, 209-223.

D'Aspremont, C. and Jacquemin, A. (1988). "Cooperative and Noncooperative R&D in Duopoly with Spillovers." *American Economic Review*, 78, 1133-1137.

Deeds, D. L. and Hill, C. W. L. (1996). "Strategic Alliances and the Rate of New Product Development: An Empirical Study of Entrepreneurial Biotechnology Firms." *Journal of Business Venturing*, 11, 1, 41-55.

Delapierre, M. and Mytelka, L.K. (1998). "Blurring Boundaries: New Inter-firm Relationships and the Emergence of Networked, Knowledge-based Oligopolies." in M. Colombo (ed.) *The Changing Boundaries of the Firm, Explaining Evolving Inter-Firm Relations*, (London, Routledge Press), 4, 72-94.

Deroian, F. and Gannon, F. (2006). "Quality-Improving Alliances in Differentiated Oligopoly." *International Journal of Industrial Organization*, 24, 629-637.

Dodgson, Mark. (1992). "The Strategic Management of R&D Collaboration." Technology Analysis and Strategic Management, 4, 3, 227-244.

Goyal, S. and Moraga-Gonzalez, J. L. (2001). "R&D Networks." *RAND Journal of Economics*, 32, 4, 686-707.

Goyal, S. and Moraga-Gonzalez, J. L. (2003). "Firms, Networks and Markets: A Survey of Recent Research." *Revue d'Economie Industrielle*.

Goyal, S., Konovalov, A. and Moraga-Gonzalez, J. L. (2005). "Hybrid R&D." *Tinbergen Institute discussion paper* TI 2003-041/1.

Goyal, S., Konovalov, A. and Moraga-Gonzalez, J. L. (2002). "Individual Research, Joint Work and Networks of Collaboration." *mimeo*, Erasmus University and Queen Mary, University of London.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Goyal, S. and Joshi, S. (2003). "Networks of Collaboration in Oligopoly." *Games and Economic Behavior*, 43, 57-85.

Hagedoorn, J. (2002). "Inter-firm R&D Partnerships: An Overview of Major Trends and Patterns since 1960." *Research Policy,* 31, 477-492.

Hemphill, T. and Vonortas, N. S. (2003). "Strategic Research Partnerships: A Managerial Perspective." *Technology Analysis and Strategic Management*, 15, 2, 255-271.

Jackson, M. O. and Wolinksy, A. (1996). "A Strategic Model of Social and Economic Networks." *Journal of Economic Theory,* 71, 44-74.

Milgrom, P. and Roberts, J. (1992). *Economics, Organization and Management*. (Englewood Cliffs, N.J.: Prentice Hall).

Motta, M. (1992). "Cooperative R&D and Vertical Product Differentiation." *International Journal of Industrial Organization*, 10, 643-661.

Kamien, M. I., Muller, E. and Zang, I. (1992). "Research Joint Ventures and R&D Cartels." *American Economic Review,* 85, 1293-1306.

Podolny, J. M. and Page, K. L. (1998). "Network Forms of Organization." *Annual Review of Sociology,* 24, 57-76.

Veugelers, R. (1998). "Collaboration in R&D: An Assessment of Theoretical and Empirical Findings." *The Economist*, 149, 419-443.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

# Chapter 3

# Online Communication Networks and the Dynamics of Collective Action

With Christopher Kuhlman[†], Achla Marathe[†], Madhav Marathe[†], Fernando Vega-Redondo[§]

## 3.1 Introduction

The use of social networking sites (especially Facebook, Twitter and Youtube) was a distinctive feature of the recent wave of uprisings against authoritarian regimes in the Arab world, and of social actions in Western countries following the financial crisis (e.g. Occupy Wall Street). Social media was used in order to help the protesters organize and reach a critical mass of participants (Gonzalez-Bailon *et* al. 2011). Information sharing prior to, as well as during, mass demonstrations and street protests in post-election Iran, Tunisia, Egypt, similarly in Western liberal democracies, proved to be essential for the success of these protests. Citizens (potential participants and protesters) required information regarding election results, the level of discontent, the opposition or government's response, ongoing events in the streets and whether, when and where additional protests were to be held. Moreover, protesters benefited from information 'on the fly' while they were on the streets in order to know what was happening at specific locations to avoid clashes with the police (Kavanaugh et al. 2011).

Social media offers an easy, quick and inexpensive means of communication that helps the spread of information (Garrett 2006). By lowering communication and coordination costs, it facilitates collective identity, group formation, recruitment, and retention while improving group efficiency, all of which contribute to increased political participation (Bonchek 1995; 1997). Moreover, by accelerating and extending the diffusion of social movement information and of protests, digital media potentially motivates future actions elsewhere (Myers 1994).

[†]Network Dynamics and Simulation Science Laboratory, Virginia Bioinformatics Institute, Virginia Tech.
[§]Department of Economics, European University Institute


Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

In this chapter, we develop a dynamic game-theoretic model of the "on-set of revolutions" that focuses on the local spread of information through online social networks. We investigate how the flow of information through online communication channels can facilitate collective behavior. Our point of departure is that the problem of collective action is one of coordination. In collective actions, people do not act unilaterally; instead, there is an implicit coordination that takes place locally. Coordination requires that people know about each others' willingness to participate and that this information is common knowledge. We develop a model in which the communication network helps coordination by creating common knowledge within each stage, and by informing agents about earlier stages. We then study how the network structure affects the commonality of knowledge and individuals' participation decisions.

We consider a population of heterogeneous agents who differ in their willingness to participate in protests. An individual wants to participate only if joined by others.[9] The number of participants above which an individual would choose to participate defines the individual's private *threshold*. Coordination requires that people know each others' thresholds and that they know that there is a sufficient number of people who share this information, i.e., information about thresholds must be common knowledge. The communication network allows individuals to share their thresholds with their neighbors and thus helps to create common knowledge. Regarding the information-sharing mechanism, we consider *facebook*-type communication in which *friends* post their willingness to participate on each others' "*walls*", revealing the thresholds to their first and second-degree neighbors. The information on the network structure is also local, so that individuals can only observe their direct neighbors' links. Moreover, in the dynamic framework, we assume that individuals also communicate their past actions using the online social network, and thereby learn about others' past actions from the *walls* of their neighbors. The information about thresholds and past actions allows agents to make inferences about the number of revolters and to decide whether to participate or not. Finally, we analyze how the correlation of neighbors' types, i.e., network homophily, affects collective action.

[9] Typically confined to the movement, protest, and rebellion literatures is the safety in numbers argument; you are safer the more others join your actions (e.g., Kuran 1991)

 DOI: 10.2870/63947

The main features highlighted by the stylized model are (i) the implications/requirements of common knowledge, (ii) the role of network topology, (iii) the interplay between individual network characteristics and individual type, (iv) the benefits and drawbacks of individual heterogeneity, and (v) the effects of alternative mechanisms and information sharing arrangements. The contribution of this chapter to the existing literature involves three aspects of information. First of all, a key feature of social networks prevailing in the real world is that agents only have local information about the network. Consequently, we restrict the information on the topology of the network to a local level, which is determined by the information sharing mechanism considered.

Second, in game-theoretic contexts, agents' behaviors depend on whether some piece of relevant information is common knowledge or not. In our framework, this information concerns not only the network, but also the types of the individuals and their actions. In order to address this issue, we use techniques from the epistemic approach developed within game theory. We formulate how information is diffused through networks and how common knowledge is obtained in order to study the role of the network structure. In particular, we characterize the structure of the sub-networks in which all agents have common knowledge about their types and actions. We find that for a given network if there exists a set of agents who share common knowledge, then a sub-graph induced by this set must be complete bipartite. Moreover, we conduct simple simulations using complete bipartite graphs to show that similar sub-networks can result in different network dynamics in which (i) only a subset of the agents, who share common knowledge, might revolt, depending on the thresholds, (ii) it might take longer for all agents to revolt, depending on whom they are connected to, even though they share the same common knowledge.

Finally, in a dynamic extension, we examine how further information is revealed if agents observe the past actions taken by their neighbors, and how this additional information is used to make inferences about the number of revolters. This allows us to study large scale movements that unfold over time despite the small number of participants at the onset. Moreover, the analysis of type-correlation in ring networks shows that a higher degree of homophily enables people with high thresholds to coordinate and to engage in collective action, but results in a slower spread through the population.

The chapter is organized as follows. We begin with a brief overview of the relevant literature. In Section 3.3, we present the model. Section 3.4 formulates the concept of common knowledge as mentioned in the literature and adapts it to our framework. In Section 3.5, we define the equilibrium of the coordination game for the static and the dynamic frameworks. Section 3.6 presents some examples that highlight the role of correlation of neighbors' types in ring networks, which motivates the need for a theoretical analysis of homophily.

## 3.2 Related Literature

Coordination and collective action have been widely studied in many contexts.[10] One branch of this literature is derived from Mancur Olson's *The Logic of Collective Action* (1965), and especially his assertion that if the benefits of a collective good are non-excludable, then rational, self-interested individuals will have an incentive to free-ride on the contributions of others. Another branch builds on Schelling's tipping point model (Schelling 1969, 1971) in which individuals only take an action (or tip) after they observe a sufficient number of other individuals taking the same action. In this chapter, we abstract from the free-riding issue and build upon the seminal works of Granovetter (1978) and Schelling (1978), studying tipping or threshold behavior in groups. Both authors modeled the participation process as a threshold effect in which a small number of early participants can trigger a chain reaction, leading to a population-wide cascade of participation in collective behavior. We also propose that the more people participate, the more likely it is that a given individual will choose to participate in the collective action.

Similar to the spread of information and disease, most collective behaviors, including revolutions (Gould 1993), protests (Marwell and Oliver 1993) and strikes (Klandermans 1988), spread through social contacts. Hence, some conclusions drawn from research on epidemics and information networks can be generalized to the spread of collective action. On the other

[10] See reviews in Oliver (1993), Centola and Macy (2007), Siegel (2009) and Strang and Soule (1998).

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

hand, while information and disease are archetypes of simple contagions, many collective behaviors involve complex contagions that require social affirmation and reinforcement from multiple resources (Centola and Macy 2007). Complex contagion requires exposure to multiple sources (as opposed to multiple exposures) to trigger adoption as it is the case for social movements (Marwell and Oliver 1993; Opp and Gern 1993; McAdam and Paulsen 1993).

Social structures, which encompass both formal configurations such as social movement organizations or churches, and informal configurations, such as friendship and activist networks, enable individuals to organize and engage in collective action (Garrett 2006). A broad and growing interdisciplinary literature suggests that participation in collective actions depends on the structure of the social networks regarding both pattern of connections (e.g., Centola and Macy 2007; Chwe 1999; Gould 1993) and the position of individuals within the networks (e.g. Borgatti and Everett 1992; Gould 1993; Kim and Bearman 1997).

Our work is closely related to Chwe (1999, 2000), who considers social structure and individual incentives together in order to study which network structures are conducive to coordination. He presents a coordination game of incomplete information and he models social structure as a communication network through which people tell each other their willingness to participate. The objective of Chwe is posed in a normative manner, i.e., he asks what is the minimal network that allows everyone to revolt when the network itself is common knowledge. He finds structured networks that are built upon hierarchy of cliques to be the minimal network structure.[11] He also shows that "low dimensional" networks with fewer links can be better for coordination. Finally, he shows that a wide dispersion of "insurgents" can be good for coordination, but that too much dispersion can be bad.

In order to focus on the role of online social networks as communication tools, we introduce some features of these networks to the framework developed by Chwe (2000). First of all, we consider a *facebook*-type communication mechanism that allows friends to post their willingness to participate on each others' *walls*. In Chwe (2000), the communication network allows individuals to learn about the thresholds of their direct neighbors. However, in our set-

---

[11] The fact that his question is so stark and unrestricted is what explains that the networks he finds are so rigid and structured.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

up, agents can also observe the thresholds of others who are within distance-2, since neighbors have access to their respective *walls*. Motivated by online social networks, this assumption allows for common knowledge to be attainable for a larger number of individuals and in a larger variety of network structures, rather than cliques.

Second, and more importantly, Chwe (2000) assumes that the network itself is common knowledge, so that agents know about all communication that occurs between all members of the population. However, one of the key features of large social networks prevailing in the real world is that agents only have local information about the network. We try to understand the rise of "social action" in a context where information is restricted by the social network. Therefore, in our model the communication network that provides information about the thresholds of neighbors is only locally known by agents. The communication mechanism we consider allows neighbors to observe each others' *walls*, through which they obtain information about the links of their direct neighbors and about their thresholds. We show how the lack of information on the network affects common knowledge and generates different results than in Chwe (2000), even for the same network structures.

Finally, the theoretical approach adopted by Chwe (2000) is static, i.e., revolution either happens at $t = 0$ or not at all. We introduce a sophisticated behavior of agents in order to study how the interplay of knowledge and learning can bring large scale social movements. In particular, we develop a dynamic model in which the communication network informs people within each stage and about earlier stages. We contemplate a rich process of updating that produces interesting dynamics, tailored to the topology of the network. In each period, agents learn the past actions of their neighbors and, given their knowledge about thresholds in the current period, make inferences that inform their decision to participate. The dynamic approach, in which people learn progressively about the network and what others know, allows us to study how agents' behaviors evolve over time as more information is revealed. We present some examples with simple network configurations to show how the distribution of thresholds affects the speed and the spread of information.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

## 3.3 The Model

There is a finite set of people $N = \{1, 2, ..., n\}$ and each person $i \in N$ chooses an action $a_i \in \{r, s\}$, where $r$ is '*revolt*', the '*risky*' action, and $s$ is '*stay at home*', the '*safe*' action. Each person $i$ has an idiosyncratic private threshold $\theta_i \in \{1, 2, ..., n+1\}$; a person wants to revolt only if the total number of people who revolt is equal to at least his threshold. Given person $i$'s threshold $\theta_i$ and everyone's actions $a = (a_1, a_{2,} ... a_n)$, her utility is given by

$$U_i(\theta_i, a_i, a_{-i}) = \begin{cases} 0 \quad \text{if } a_i = s \\ 1 \quad \text{if } a_i = r \ \wedge \ \#\{j \in N : a_j = r\} \geq \theta_i \\ -z \text{ if } a_i = r \ \wedge \ \#\{j \in N : a_j = r\} < \theta_i \end{cases}$$

where $-z < -(n+1)^n$. In other words, a person always gets utility $0$ by staying at home. When he revolts, he gets utility $1$ if the total number of people revolting (including himself) is at least $\theta_i$. For example, $\theta_i = 3$ implies that at least two more people need to revolt in order for person $i$ to enjoy nonnegative payoffs. He gets a very large penalty $-z$, if he revolts and not enough people join him. Thus, a person will revolt as long as he is sure that there is a sufficient number of people revolting.

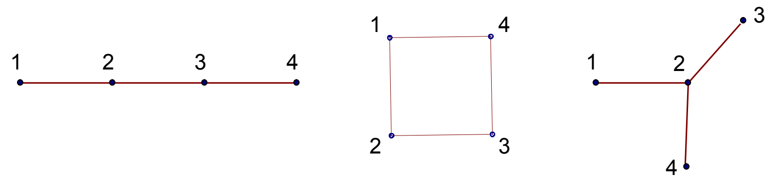
The communication network is undirected and is represented by $G(N, L)$, where $L$ denotes the set of links.[12] The communication technology we consider here is "facebook-type" communication in which people write on each others' "walls". When $ij \in L$, person $i$ and person $j$ are linked, which implies that person $i$ writes his threshold, $\theta_i$, (and action, $a_i$) on person $j$'s *wall* (and vice versa). This post is observed by $j$'s neighbors since they have access to the *wall* of $j$.[13] The network structure is not observed by everyone, but person $i$ knows about the thresholds, action plans (and past actions in the dynamic framework) of the people in his '*ball*' which is denoted by $B_i = \{j \in N_i^2\}$. Finally, let $N_i(g) = \{j \in N \setminus \{i\} : ij \in L\}$ be the set of person $i$'s neighbors and let $\eta_i(g) = \#N_i(g)$ be the cardinality of this set. Figure 1 illustrates how the communication technology allows people in different networks to learn about thresholds.

12 Chwe (2000) studies directed networks where $j \rightarrow i$ means that person $j$ talks to person $i$.

13 In Chwe, $j \rightarrow i$ means that $j$ tells his threshold to $i$ and it is only observed by $i$. The *wall* allows us to extend his framework to distance-2 neighbors while we limit the information on the network structure.

 DOI: 10.2870/63947

*Figure 1.* Line, star and square with $\theta_i = 4 \quad \forall i$



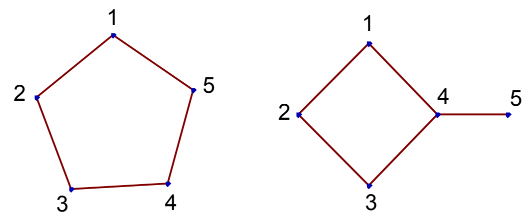

In the line, person 1 learns about the threshold of person 2 directly since they write on each others' walls. He learns about person 3 through the wall of person 2 (since person 3 writes her threshold on the wall of person 2 and it is observed by person 1). However, person 1 is not able to observe either the threshold of person 4 or the link between person 3 and person 4 because he is not in person 1's ball. Similarly, person 4 does not know about the threshold of person 1. Person 2 learns about the thresholds of agents 1 and 3 directly and she learns about the threshold of person 4 through the wall of person 3. Thus, person 2 and person 3 can observe all thresholds and links in the line network since everyone is within distance-2.

When we consider the square, everyone is within distance-2 of each other. For example, person 1 learns about agents 2 and 4 directly and he learns about person 3 through the walls of person 2 and 4. Hence, he also knows that they are linked to person 3. Due to symmetry, this is true for all agents. Thus, directly or indirectly, all agents learn about the thresholds of everyone else and everyone knows the structure of the network.

Finally, in the star network, person 2 (who is the hub) learns about the thresholds of everyone directly, while the spokes learn about each other through the wall of person 2. In addition, everyone can observe the structure of the network, as was the case in the square.

As illustrated in Figure 1, the communication network provides (limited) information about the thresholds of others. However, a person's decision to revolt depends not only on knowing others' thresholds, but also on knowing what other people know. Common knowledge among a set of people implies that: *they know each others' thresholds, action plans (and past actions) and they know that they know their thresholds and actions. Therefore they can count on each other.* In other words, everyone should know that there is sufficient discontent; everyone has to know that everyone knows, everyone has to know that everyone knows that everyone knows, and so on. (Lewis 1969; Aumann 1976). Figure 2 below highlights the importance of network structure and common knowledge in facilitating coordination.

*Figure 2.* Pentagon and kite with $\theta_i = 4 \quad \forall i$




In the pentagon, person 1 knows the thresholds of person 2 and person 5 directly and he knows the thresholds of person 3 and person 4 through the walls of person 2 and person 5, respectively. Thus, he knows the thresholds of everyone and he knows that everyone knows his threshold (by symmetry). However, he cannot observe the link between person 3 and person 4. Since he has limited information on the network structure, he does not know whether person 3 and person 4 communicate. This affects person 2's knowledge of person 4's threshold and similarly person 5's knowledge of person 3's threshold.[14] Therefore, when person 1 decides to revolt, he cannot count on the others since he does not know whether they know that there is sufficient discontent. The same argument applies for all agents and, although everyone knows everyone's thresholds, and despite the fact that there is a sufficient number of people who would get positive payoffs if they revolted (i.e., $n > \theta_i \quad \forall i$), no one revolts in the pentagon.

In the kite, as with the pentagon, we observe that person 1 knows the thresholds of everyone (everyone is within distance-2). This is also the case for person 3 and person 4. Although person 2 and person 5 do not know about each other (since they do not have a common friend), there is common knowledge of thresholds among agents 1, 2, 3, 4 and among people 1, 3, 4 and 5. Since all agents have threshold 4, the people in each group who share common knowledge about each others' thresholds know that if they jointly deviate and revolt, everyone in the group will gain. Therefore, everyone revolts.

[14] In other words, person 1 does not know whether person 3 is linked to person 4 (in which case they would know each others' thresholds). Person 1 can only observe that person 3 knows the thresholds of two other people (person 1 and person 2), and this information is not enough for person 3 to revolt.

Note that the difference between the pentagon and the kite is a structural difference, since both have the same number of people with same thresholds, and the same number of links. In Chwe (2000), if people know the thresholds of everyone, they can immediately observe the actual state of the world and revolt so long as there is sufficient discontent. If we assume the network structure is also common knowledge (as in Chwe, 2000), everyone would revolt in the pentagon. In that case, person 1 would also know that there is common knowledge among everyone, which is more than enough for all people to revolt since they have threshold 4. In our model, both information about the thresholds and information about common knowledge (determined by the network structure) matter. People do not revolt in the pentagon even when they know everyone's thresholds because, due to limited information about the network, they don't know if there is common knowledge.

Therefore, we observe that common knowledge (CK) is a key property of the network. CK is important because any joint move, deviation or adjustment by a set of agents on the basis of some information is only possible if this information is common knowledge. In the following section, we analyze formally how the communication network provides common knowledge and how the topology of the network matters in facilitating coordination.

## 3.4 Common Knowledge

In this section, we follow the Stanford Encyclopedia of Philosophy (Vanderschraaf and Sillari, 2009) for definitions and notations relating to common knowledge. Informally, an event $E$ is *mutual knowledge* among a set of agents if each agent knows that $E$. Mutual knowledge by itself implies nothing about what, if any, knowledge anyone attributes to anyone else. Suppose each student arrives for a class meeting knowing that the instructor will be late. That the instructor will be late is mutual knowledge, but each student might think that only he knows that the instructor will be late. However, if one of the students says openly "Peter told me he will be late again," then the mutually known fact is now *commonly known*. Each student now knows that the instructor will be late, that every student knows, and so on, *ad infinitum*. The agents have common knowledge in the sense articulated informally by Schelling (1960), and more

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

precisely by Lewis (1969) and Schiffer (1972). Although there are a number of ways in which the concept of common knowledge can be formalized, we adopt the set-theoretic approach.

Following C. I. Lewis (1943; 1944) and Carnap (1947), events are subsets of a set $\Omega$ of possible worlds. A distinct actual world $\omega_\alpha$ is an element of $\Omega$. An event $E \subseteq \Omega$ *obtains* (or is true) if the actual world $\omega_\alpha \in E$. The event should be consistent with the actual state.

What an agent $i$ knows about the set of possible worlds is stated formally in terms of a *knowledge operator* $\mathbf{K}_i(E)$.[15] Given an event $E \subseteq \Omega$, $\mathbf{K}_i(E)$ denotes a new event, corresponding to the set of possible worlds in which agent $i$ knows that $E$ obtains. $\mathbf{K}_i(E)$ is read as '$i$ knows (that) $E$ (is the case)'.

*Definition.* Agent $i$'s *possibility set* $P_i(\omega)$ at $\omega \in \Omega$ is defined as

$$P_i(\omega) \equiv \bigcap \{E \mid \omega \in \mathbf{K}_i(E)\}$$

The collection of sets $\mathcal{P}_i = \bigcup_{\omega \in \Omega} P_i(\omega)$ is $i$'s *private information partition*.

$P_i(\omega)$ is the intersection of all events that $i$ knows at $\omega$, $P_i(\omega)$ is the smallest event in $\Omega$ that $i$ knows at $\omega$. In other words, $P_i(\omega)$ is the most specific information that $i$ has about the possible world $\omega$. The elements of $i$'s information system represent what $i$ knows immediately at a possible world. Agents' possibility sets *partition* the state set, if axioms K1 - K5 hold. We can also write player $i$'s knowledge function as:

$$\mathbf{K}_i(E) = \{\omega \in \Omega \mid P_i(\omega) \subseteq E\}$$

We can now define mutual and common knowledge as follows:

*Definition.* Let a set $\Omega$ of possible worlds together with a set of agents $N$ be given.

1. The event that $E$ is (*first order*) *mutual knowledge* for the agents of $N$, $\mathbf{K}^1_N(E)$, is the set defined by

$$\mathbf{K}^1_N(E) \equiv \bigcap_{i \in N} \mathbf{K}_i(E)$$

2. The event that $E$ is $m^{th}$ *order mutual knowledge* among the agents of $N$, $\mathbf{K}^m_N(E)$, is defined recursively as the set

$$\mathbf{K}^m_N(E) \equiv \bigcap_{i \in N} \mathbf{K}_i(\mathbf{K}^{m-1}_N(E))$$

15 Related axioms and their properties can be found in the Appendix.

 DOI: 10.2870/63947

3. The event that $E$ is *common knowledge* among the agents of $N$, $\mathbf{K}_N^*(E)$, is defined as the set

$$\mathbf{K}_N^*(E) \equiv \bigcap_{m=1}^{\infty} \mathbf{K}_N^m(E)$$

It can be shown that:

(1) If $\omega \in \mathbf{K}_N^*(E)$ and $E \subseteq F$, then $\omega \in \mathbf{K}_N^*(F)$.

(2) $\omega \in \mathbf{K}_N^m(E)$ if and only if for all agents $i_1$,$i_2$,...,$i_m \in N$, $\omega \in \mathbf{K}_{i_1}\mathbf{K}_{i_2}..\mathbf{K}_{i_m}(E)$.

Hence, $\omega \in \mathbf{K}_N^*(E)$ if and only if (2) is the case for each $m \geq 1$. [16]

In our model, there is a finite set of agents $N = \{1, 2, ..., n\}$ and each person $i$ has a private threshold $\theta_i \in \Theta = \{1, 2, ..., n+1\}$. The communication network is undirected and it is represented by $G(N, L)$. As a benchmark, we will first formalize the case in which the network structure is known by all agents. Then, we will define common knowledge for the case in which people have limited information about the network.

### 3.4.1 Full Information on the Network Structure

We assume that the network structure is public information so that the set of possible states is defined only by the set of thresholds. The set of states is $\Omega = \Theta^n = \{1, 2, ..., n+1\}^n$, hence a state $\omega \in \Omega$ will be $n$-*tuple* such as $\omega = (2, 2, 2)$ for 3 people with $\theta_i = 2$ for all $i$. In order to demonstrate these concepts, we present the following simple example.

**Example 1**. $n = 2$, $\theta_i \in \{1, 2, 3\}$

The possible states of the world can be written as follows:

| | $\omega_1$ | $\omega_2$ | $\omega_3$ | $\omega_4$ | $\omega_5$ | $\omega_6$ | $\omega_7$ | $\omega_8$ | $\omega_9$ |
|---|---|---|---|---|---|---|---|---|---|
| Person 1 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 |
| Person 2 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |

Let us abuse the notation slightly by defining $\omega_{\theta_1\theta_2}$, i.e., replacing $\omega_1$ with $\omega_{11}$; $\omega_6$ with $\omega_{23}$, and so on.

[16] The condition that $\omega \in \mathbf{K}_{i_1}\mathbf{K}_{i_2}...\mathbf{K}_{i_m}(E)$ for all $m \geq 1$ and all $i_1, i_2, ..., i_m \in N$ is Schiffer's definition of common knowledge, which is the one most often used in the literature. Schiffer uses the formal vocabulary of *epistemic logic* and his general approach is to augment a system of sentential logic with a set of knowledge operators corresponding to a set of agents, and then to define common knowledge as a hierarchy of events in the augmented system.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

**Case 1**. There is no communication between the agents.


In this case, the private information partitions of person 1 and person 2 can be written as:

$$\mathcal{P}_1 = \{\{\omega_{11}, \omega_{12}, \omega_{13}\}, \{\omega_{21}, \omega_{22}, \omega_{23}\}, \{\omega_{31}, \omega_{32}, \omega_{33}\}\}$$

$$\mathcal{P}_2 = \{\{\omega_{11}, \omega_{21}, \omega_{31}\}, \{\omega_{12}, \omega_{22}, \omega_{32}\}, \{\omega_{13}, \omega_{23}, \omega_{33}\}\}$$

Person 1's possibility sets can be written as:

$$P_1(\omega_{11}) = P_1(\omega_{12}) = P_1(\omega_{13}) = \{\omega_{11}, \omega_{12}, \omega_{13}\}$$

$$P_1(\omega_{21}) = P_1(\omega_{22}) = P_1(\omega_{23}) = \{\omega_{21}, \omega_{22}, \omega_{23}\}$$

$$P_1(\omega_{31}) = P_1(\omega_{32}) = P_1(\omega_{33}) = \{\omega_{31}, \omega_{32}, \omega_{33}\}$$

Similarly person 2's possibility sets are:

$$P_2(\omega_{11}) = P_2(\omega_{21}) = P_2(\omega_{31}) = \{\omega_{11}, \omega_{21}, \omega_{31}\}$$

$$P_2(\omega_{12}) = P_2(\omega_{22}) = P_2(\omega_{32}) = \{\omega_{12}, \omega_{22}, \omega_{32}\}$$

$$P_2(\omega_{13}) = P_2(\omega_{23}) = P_2(\omega_{33}) = \{\omega_{13}, \omega_{23}, \omega_{33}\}$$

Now, let's say that the actual world is $\omega_\alpha = \{\omega_{22}\}$ so that both agents have threshold of 2. We know that an event $E \subseteq \Omega$ obtains (or is true) if the actual world $\omega_\alpha \in E \subseteq \Omega$. Consider the following events:

$$E_1 = \{\omega_{22}\}$$

$$E_2 = \{\text{states that both agents benefit if they revolt}\} = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$$

$$E_3 = \Omega$$

We can write players' knowledge functions, i.e., $\mathbf{K}_i(E) = \{\omega \in \Omega \mid P_i(\omega) \subseteq E\}$, as follows:

$\mathbf{K}_1(E_1) = \text{Ø}$ since $P_1(\omega_{22}) = \{\omega_{21}, \omega_{22}, \omega_{23}\} \nsubseteq E_1$

Thus, there is no state in which person 1 knows that $E_1$ occurs.

Similarly, $\mathbf{K}_2(E_1) = \text{Ø}$ since $P_2(\omega_{22}) = \{\omega_{12}, \omega_{22}, \omega_{32}\} \nsubseteq E_1$

$\mathbf{K}_1(E_2) = \text{Ø}$ since $\nexists \omega \in \Omega$ such that $P_1(\omega) \nsubseteq E_2$. Similarly, $\mathbf{K}_2(E_2) = \text{Ø}$.

$\mathbf{K}_1(E_3) = \Omega$ and $\mathbf{K}_2(E_3) = \Omega$. They both know in each state that the event $\Omega$ is the case.

Therefore, we can say that the only event that can be mutually and commonly knowledge is $E_3$.

**Case 2**. There is communication between the agents.


In this case, the private information partitions of person 1 and person 2 can be written as:

$$\mathcal{P}_1 = \{\{\omega_{11}\}, \{\omega_{12}\}, \{\omega_{13}\}, \{\omega_{21}\}, \{\omega_{22}\}, \{\omega_{23}\}, \{\omega_{31}\}, \{\omega_{32}\}, \{\omega_{33}\}\}$$

$$\mathcal{P}_2 = \{\{\omega_{11}\}, \{\omega_{21}\}, \{\omega_{31}\}, \{\omega_{12}\}, \{\omega_{22}\}, \{\omega_{32}\}, \{\omega_{13}\}, \{\omega_{23}\}, \{\omega_{33}\}\}$$

Person 1 and 2's possibility sets can be written as:

$P_1(\omega_{11}) = \{\omega_{11}\}$, $P_1(\omega_{12}) = \{\omega_{12}\}$, $P_1(\omega_{13}) = \{\omega_{13}\}$, and so on.

$P_2(\omega_{11}) = \{\omega_{11}\}$, $P_2(\omega_{12}) = \{\omega_{12}\}$, $P_2(\omega_{13}) = \{\omega_{13}\}, \ldots$

Now, for the same events $E_1, E_2, E_3$, players' knowledge functions become:

$$\mathbf{K}_1(E_1) = \mathbf{K}_2(E_1) = \{\omega_{22}\}$$

They both know in state $\omega_{22}$ that $E_1$ occurs.

$$\mathbf{K}_1(E_2) = \mathbf{K}_2(E_2) = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$$

They know in each state $\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}$ that $E_2$ occurs.

$\mathbf{K}_1(E_3) = \Omega$ and $\mathbf{K}_2(E_3) = \Omega$

They both know in each state that the event is $\Omega$.

We can now analyze mutual and common knowledge as follows:

The event $E_1$ is (*first order*) *mutual knowledge* for the agents since

$$\omega_{22} \in \mathbf{K}_N^1(E_1) \equiv \bigcap\nolimits_{i \in N} \mathbf{K}_i(E_1) = \mathbf{K}_1(E_1) \cap \mathbf{K}_2(E_1)$$

We have defined

$$\mathbf{K}_1\mathbf{K}_2(E_1) \equiv \{\omega \in \Omega : P_1(\omega) \subseteq \mathbf{K}_2(E_1)\}$$
$$\mathbf{K}_2\mathbf{K}_1(E_1) \equiv \{\omega \in \Omega : P_2(\omega) \subseteq \mathbf{K}_1(E_1)\}$$

The event $E_1$ is common knowledge among person 1 and person 2 since

$\mathbf{K}_1\mathbf{K}_2(E_1) = \{\omega_{22}\}$ and $\mathbf{K}_2\mathbf{K}_1(E_1) = \{\omega_{22}\}$. Therefore $\omega_{22} \in \mathbf{K}_N^*(E_1)$.

Similarly, we can show that the events $E_2$ and $E_3$ are mutually and commonly known by person 1 and person 2.

Also note that $\omega_{22} \in \mathbf{K}_N^*(E_1)$ and $E_1 \subseteq E_2 \subseteq E_3$, hence $\omega_{22} \in \mathbf{K}_N^*(E_2)$ and $\omega_{22} \in \mathbf{K}_N^*(E_3)$.

### 3.4.2 Local Information on the Network Structure

Let the pairwise relationship between two agents be represented by the binary variable $g_{ij} \in \{0,1\}$. When $g_{ij} = 1$, two agents are linked, while when $g_{ij} = 0$ there is no link. Now, the set of states is given by $\Omega = \Theta^n \times G$ with $G = \{0,1\}^{C_2^n}$ where $C_2^n = n(n-1)/2$ is the number of 2-combinations of $n$ nodes. Formally, a state can be written as $\omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$. Thus, a state will be a $(n + C_2^n)$-***tuple*** such as $\omega = (2, 2, 2, 0, 0, 1)$ for a network with 3 people with $\theta_i = 2$, in which $g_{12} = 0, g_{13} = 0, g_{23} = 1$.

**Example 2**. Let's take the same example with $n = 2$, $\theta_i \in \{1, 2, 3\}$

The states can be written

| | $\omega_1$ | $\omega'_1$ | $\omega_2$ | $\omega'_2$ | $\omega_3$ | $\omega'_3$ | $\omega_4$ | $\omega'_4$ | $\omega_5$ | $\omega'_5$ | $\omega_6$ | $\omega'_6$ | $\omega_7$ | $\omega'_7$ | $\omega_8$ | $\omega'_8$ | $\omega_9$ | $\omega'_9$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |
| P2 | 1 | 1 | 2 | 2 | 3 | 3 | 1 | 1 | 2 | 2 | 3 | 3 | 1 | 1 | 2 | 2 | 3 | 3 |
| $g_{12}$ | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |

Let us abuse the notation slightly by defining $\omega_{\theta_1\theta_2 g_{12}}$, i.e., by replacing $\omega_1 = (1,1,0)$ with $\omega_{11 g_{12}=0}$; $\omega'_1 = (1,1,1)$ with $\omega_{11 g_{12}=1}$; $\omega'_6$ with $\omega'_{23 g_{12}=1} = (2,3,1)$, and so on.

Private information partitions of person 1 and person 2 can be written as:

$\mathcal{P}_1 = \{\{\omega_{11g_{12}=0}, \omega_{12g_{12}=0}, \omega_{13g_{12}=0}\}, \{\omega_{21g_{12}=0}, \omega_{22g_{12}=0}, \omega_{23g_{12}=0}\}, \{\omega_{31g_{12}=0}, \omega_{32g_{12}=0}, \omega_{33g_{12}=0}\},$

$\{\omega_{11g_{12}=1}\}, \{\omega_{12g_{12}=1}\}, \{\omega_{13g_{12}=1}\}, \{\omega_{21g_{12}=1}\}, \{\omega_{22g_{12}=1}\}, \{\omega_{23g_{12}=1}\}, \{\omega_{31g_{12}=1}\}, \{\omega_{32g_{12}=1}\},$

$\{\omega_{33g_{12}=1}\}\}$

$\mathcal{P}_2 = \{\{\omega_{11g_{12}=0}, \omega_{21g_{12}=0}, \omega_{31g_{12}=0}\}, \{\omega_{12g_{12}=0}, \omega_{22g_{12}=0}, \omega_{32g_{12}=0}\}, \{\omega_{13g_{12}=0}, \omega_{23g_{12}=0}, \omega_{33g_{12}=0}\},$

$\{\omega_{11g_{12}=1}\}, \{\omega_{21g_{12}=1}\}, \{\omega_{31g_{12}=1}\}, \{\omega_{12g_{12}=1}\}, \{\omega_{22g_{12}=1}\}, \{\omega_{32g_{12}=1}\}, \{\omega_{13g_{12}=1}\}, \{\omega_{23g_{12}=1}\},$

$\{\omega_{33g_{12}=1}\}\}$

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Person 1's possibility sets can be written as:

$P_1(\omega_{11_{g_{12}=0}}) = \{\omega_{11_{g_{12}=0}}, \omega_{12_{g_{12}=0}}, \omega_{13_{g_{12}=0}}\}$; $P_1(\omega_{11_{g_{12}=1}}) = \{\omega_{11_{g_{12}=1}}\}$

$P_1(\omega_{22_{g_{12}=0}}) = \{\omega_{21_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{23_{g_{12}=0}}\}$; $P_1(\omega_{22_{g_{12}=1}}) = \{\omega_{11_{g_{22}=1}}\}$

$P_1(\omega_{33_{g_{12}=0}}) = \{\omega_{31_{g_{12}=0}}, \omega_{32_{g_{12}=0}}, \omega_{33_{g_{12}=0}}\}$; $P_1(\omega_{33_{g_{12}=1}}) = \{\omega_{33_{g_{12}=1}}\}$, and so on.

Similarly, person 2's possibility sets are:

$P_2(\omega_{11_{g_{12}=0}}) = \{\omega_{11_{g_{12}=0}}, \omega_{21_{g_{12}=0}}, \omega_{31_{g_{12}=0}}\}$; $P_2(\omega_{11_{g_{12}=1}}) = \{\omega_{11_{g_{12}=1}}\}$

$P_2(\omega_{22_{g_{12}=0}}) = \{\omega_{12_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{32_{g_{12}=0}}\}$; $P_2(\omega_{22_{g_{12}=1}}) = \{\omega_{11_{g_{22}=1}}\}$

$P_2(\omega_{33_{g_{12}=0}}) = \{\omega_{13_{g_{12}=0}}, \omega_{23_{g_{12}=0}}, \omega_{33_{g_{12}=0}}\}$; $P_2(\omega_{33_{g_{12}=1}}) = \{\omega_{33_{g_{12}=1}}\}$, and so on.

Now, consider these events:

(i) $E_1 = \{\omega_{22_{g_{12}=0}}\}$

(ii) $E_2 = \{\omega_{22_{g_{12}=1}}\}$

(iii) $E_3 =$ {states in which both agents benefit if they revolt}

$$= \{\omega_{11_{g_{12}=0}}, \omega_{11_{g_{12}=1}}, \omega_{12_{g_{12}=0}}, \omega_{12_{g_{12}=1}}, \omega_{21_{g_{12}=0}}, \omega_{21_{g_{12}=1}}, \omega_{22_{g_{12}=0}}, \omega_{22_{g_{12}=1}}\}$$

(iv) $E_4 = \{\omega_{11_{g_{12}=0}}, \omega_{12_{g_{12}=0}}, \omega_{13_{g_{12}=0}}\}$

(v) $E_5 =$ {states in which at least one agent has threshold 2 and $g_{12} = 0$}

$$= \{\omega_{21_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{23_{g_{12}=0}}, \omega_{12_{g_{12}=0}}, \omega_{32_{g_{12}=0}}\}$$

We can write the player's knowledge functions, i.e., $\mathbf{K}_i(E) = \{\omega \in \Omega \mid P_i(\omega) \subseteq E\}$, as follows:

(i) $\mathbf{K}_1(E_1) = \text{Ø}$ since $P_1(\omega_{22_{g_{12}=0}}) = \{\omega_{11_{g_{12}=0}}, \omega_{12_{g_{12}=0}}, \omega_{13_{g_{12}=0}}\} \nsubseteq E_1$

$\mathbf{K}_2(E_1) = \text{Ø}$ since $P_2(\omega_{22_{g_{12}=0}}) = \{\omega_{11_{g_{12}=0}}, \omega_{21_{g_{12}=0}}, \omega_{31_{g_{12}=0}}\} \nsubseteq E_1$

Thus, there is no state in which person 1 and person 2 know that $E_1$ obtains. In other words, the event $E_1$ is not mutually known since

$$\mathbf{K}^1_N(E_1) \equiv \bigcap_{i\in N} \mathbf{K}_i(E_1) = \mathbf{K}_1(E_1) \cap \mathbf{K}_2(E_1) = \text{Ø}.$$

(ii) $\mathbf{K}_1(E_2) = \mathbf{K}_2(E_2) = \{\omega_{22_{g_{12}=1}}\}$

$\mathbf{K}_N^1(E_2) = \mathbf{K}_1(E_2) \cap \mathbf{K}_2(E_2) = \{\omega_{22_{g_{12}=1}}\}$

Hence, they mutually know in state $\omega_{22_{g_{12}=1}}$ that $E_2 = \omega_{22_{g_{12}=1}}$ is the case.

We can also see that the event $E_2$ is common knowledge among person 1 and person 2 since

$\mathbf{K}_1\mathbf{K}_2(E_2) = \mathbf{K}_2\mathbf{K}_1(E_2) = \{\omega_{22_{g_{12}=1}}\}$. Therefore, $\omega_{22_{g_{12}=1}} \in \mathbf{K}_N^*(E_2)$.

(iii) $\mathbf{K}_1(E_3) = \mathbf{K}_2(E_3) = \{\omega_{11_{g_{12}=1}}, \omega_{12_{g_{12}=1}}, \omega_{21_{g_{12}=1}}, \omega_{22_{g_{12}=1}}\}$

$\mathbf{K}_N^1(E_3) = \{\omega_{11_{g_{12}=1}}, \omega_{12_{g_{12}=1}}, \omega_{21_{g_{12}=1}}, \omega_{22_{g_{12}=1}}\}$

They both know when $g_{12} = 1$ that $\theta_i \leq 2$ for both.

$E_3$ is common knowledge in states $\omega_{11_{g_{12}=1}}, \omega_{12_{g_{12}=1}}, \omega_{21_{g_{12}=1}}, \omega_{22_{g_{12}=1}}$.

(iv) $\mathbf{K}_1(E_4) = \{\omega_{11_{g_{12}=0}}, \omega_{12_{g_{12}=0}}, \omega_{13_{g_{12}=0}}\}$

$\mathbf{K}_2(E_4) = \emptyset$

$\mathbf{K}_N^1(E_4) = \bigcap_{i\in N} \mathbf{K}_i(E_4) = \mathbf{K}_1(E_4) \cap \mathbf{K}_2(E_4) = \emptyset$

$E_4$ is only known by person 1, thus it is not mutual knowledge.

(v) $\mathbf{K}_1(E_5) = \{\omega_{21_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{23_{g_{12}=0}}\}$

$\mathbf{K}_2(E_5) = \{\omega_{12_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{32_{g_{12}=0}}\}$

Person 1 knows in states $\omega_{21_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{23_{g_{12}=0}}$ and Person 2 knows in states $\omega_{12_{g_{12}=0}}, \omega_{22_{g_{12}=0}}, \omega_{32_{g_{12}=0}}$ that at least one of them has threshold 2 and $g_{12} = 0$.

$\mathbf{K}_N^1(E_5) = \bigcap_{i\in N} \mathbf{K}_i(E_5) = \mathbf{K}_1(E_5) \cap \mathbf{K}_2(E_5) = \{\omega_{22_{g_{12}=0}}\}$

They mutually know in state $\omega_{22_{g_{12}=0}}$ that at least one of them has threshold 2 and $g_{12} = 0$. However,

$\mathbf{K}_1\mathbf{K}_2(E_5) = \{\omega \in \Omega : P_1(\omega) \subseteq \mathbf{K}_2(E_5)\} = \emptyset$

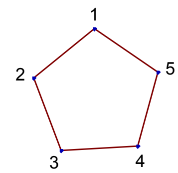
$\mathbf{K}_2\mathbf{K}_1(E_5) = \{\omega \in \Omega : P_2(\omega) \subseteq \mathbf{K}_1(E_5)\} = \emptyset$

There is no state in which one agent knows that the other knows that at least one of them has threshold 2 and $g_{12} = 0$. Therefore, the event $E_5$ is not common knowledge.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

**Example 3**. Now, we will analyze the pentagon example and formally show that the thresholds are known by all agents but that the state is not common knowledge due to limited information on the network structure.

*Figure 3.* Pentagon $n = 5$, $\theta_i \in \{1, 2, 3, 4, 5, 6\}$




We know that the set of states is $\Omega = \Theta^n \times G$ where $G = \{0,1\}^{C_2^n}$. Consider the actual state $\omega_\alpha = (4, 4, 4, 4, 4, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1)$ such that $\theta_i = 4 \ \ \forall i \in N$. The possibility sets of the agents in the actual state are given by

$$P_1(\omega_\alpha) = \{(4,4,4,4,4,1,0,0,1,1,0,0,\mathbf{1},0,1),(4,4,4,4,4,1,0,0,1,1,0,0,\mathbf{0},0,1)\}$$

$$P_2(\omega_\alpha) = \{(4,4,4,4,4,1,0,0,1,1,0,0,1,0,\mathbf{1}),(4,4,4,4,4,1,0,0,1,1,0,0,1,0,\mathbf{0})\}$$

$$P_3(\omega_\alpha) = \{(4,4,4,4,4,1,0,0,\mathbf{1},1,0,0,1,0,1),(4,4,4,4,4,1,0,0,\mathbf{0},1,0,0,1,0,1)\}$$

$$P_4(\omega_\alpha) = \{(4,4,4,4,4,\mathbf{1},0,0,1,1,0,0,1,0,1),(4,4,4,4,4,\mathbf{0},0,0,1,1,0,0,1,0,1)\}$$

$$P_5(\omega_\alpha) = \{(4,4,4,4,4,1,0,0,1,\mathbf{1},0,0,1,0,1),(4,4,4,4,4,1,0,0,1,\mathbf{0},0,0,1,0,1)\}$$

Now, consider the event $E = \{$states in which all agents have threshold 4$\} = \{\omega \in \Omega \mid \theta_i = 4 \ \forall i \in N\}$.

We can see that $\omega_\alpha \in E$ and $P_i(\omega_\alpha) \subseteq E$ for all $i \in N$.

We have defined $\mathbf{K}_i(E) = \{\omega \in \Omega \mid P_i(\omega) \subseteq E\}$. Since $P_i(\omega_\alpha) \subseteq E$, $\omega_\alpha \in K_i(E)$ for all $i \in N$.

Therefore, $\omega_\alpha \in \bigcap_{i \in N} \mathbf{K}_i(E)$. Event $E$ is *mutually known* by all agents in state $\omega_\alpha$.

In order to check whether the event $E$ is *common knowledge* in state $\omega_\alpha$, we need to show whether $\omega_\alpha \in \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E)$ for all $i, j, m \in N$.

Let us take person 1 and person 2. We have defined $\mathbf{K}_2\mathbf{K}_1(E) = \{\omega \in \Omega : P_2(\omega) \subseteq \mathbf{K}_1(E)\}$.

We need to show whether $P_2(\omega_\alpha) \subseteq \mathbf{K}_1(E)$. In words, whether person 2 knows that person 1 knows the event $E$ in the actual state.

$P_2(\omega_\alpha) = \{\omega_\alpha, \omega'_\alpha\}$, where $\omega'_\alpha = (4,4,4,4,4,1,0,0,1,1,0,0,1,0,\mathbf{0})$, i.e., the state in which all agents have threshold 4 but the link between agents 4 and 5 is deleted (line graph).

Let us define another state $\omega''_\alpha = (4,4,4,\mathbf{5},4,1,0,0,1,1,0,0,1,0,\mathbf{0})$in which $\theta_4 = 5$.

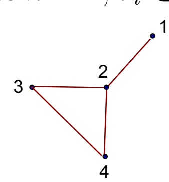
We can see that $\omega''_\alpha \in P_1(\omega'_\alpha)$. The states $\omega'_\alpha$ and $\omega''_\alpha$ are both in the same possibility set of person 1. Person 1 cannot distinguish between the states $\omega'_\alpha$ and $\omega''_\alpha$ when $g_{45} = 0$. However, $\omega''_\alpha \notin E$ since $\theta_4 = 5$. Since $P_1(\omega'_\alpha) \nsubseteq E$, then $\omega'_\alpha \notin \mathbf{K}_1(E)$. It follows that $P_2(\omega_\alpha) = \{\omega_\alpha, \omega'_\alpha\} \nsubseteq \mathbf{K}_1(E)$.

Therefore, we can state that the event *E* is *not common knowledge* in state $\omega_\alpha$.

Finally, let us show formally what happens if the network structure is fully observed by the agents. The set of states is defined by $\Omega = \Theta^n$ and the actual state is $\omega_\alpha = (4,4,4,4,4)$. Now, the possibility sets of the agents will be $P_i(\omega_\alpha) = \{(4,4,4,4,4)\} \ \ \forall i \in N$, since they can observe each others' thresholds. Taking the same event, it follows that $\omega_\alpha \in \bigcap_{i \in N} \mathbf{K}_i(E)$ and $\omega_\alpha \in \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E)$ for all $i,j,m \in N$. Therefore, we can state that the event E is *common knowledge* in state $\omega_\alpha$.

**Example 4**. Now, we analyze the kite example in which knowledge about the network is sufficient but not necessary for creating common knowledge among the agents about their thresholds.

*Figure 4.* Kite $n = 4$, $\theta_i \in \{1,2,3,4,5\}$




Consider the actual state $\omega_\alpha = (4,4,4,4,1,0,0,1,1,1)$, i.e., all agents have threshold 4 in the kite. The possibility sets of agents in the actual state can be written as

$P_1(\omega_\alpha) = \{\omega_\alpha, (4,4,4,4,1,0,0,1,1,\mathbf{0})\}$ and $P_j(\omega_\alpha) = \{\omega_\alpha\}$ for $j = 2,3,4$

Let $\omega'_\alpha = (4,4,4,4,1,0,0,1,1,\mathbf{0})$. This state is a star network with $\theta_i = 4 \ \ \forall i \in N$. Person 1 cannot distinguish between the states $\omega_\alpha$ and $\omega'_\alpha$ because he cannot observe the link between the agents 3 and 4. The others know in state $\omega_\alpha$ that state $\omega_\alpha$ occurs.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

Now, consider the event $E = \{\text{states in which all agents have threshold } 4\} = \{\omega \in \Omega \mid \theta_i = 4 \ \forall i \in N\}$

We can see that $\omega_\alpha \in E$ and $P_i(\omega_\alpha) \subseteq E$ for all $i \in N$. Hence $\omega_\alpha \in K_i(E)$ for all $i \in N$.

Therefore, $\omega_\alpha \in \bigcap_{i \in N} \mathbf{K}_i(E)$. Event $E$ is *mutually known* by all agents in state $\omega_\alpha$.

Now, we need to show that $\omega_\alpha \in \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E)$ for all $i, j, m \in N$.

For simplicity, we define another event $E' = \{\omega_\alpha, \omega'_\alpha\}$.

$\mathbf{K}_1(E') = \{\omega \in \Omega : P_1(\omega) \subseteq E'\} = \mathbf{K}_j(E') = \{\omega_\alpha, \omega'_\alpha\}$ for $j = 2, 3, 4$.

First, let us take agents 1 and 2. We have defined $\mathbf{K}_2\mathbf{K}_1(E) = \{\omega \in \Omega : P_2(\omega) \subseteq \mathbf{K}_1(E)\}$. Hence, $\mathbf{K}_2\mathbf{K}_1(E') = \{\omega_\alpha, \omega'_\alpha\}$.

Alternatively, since $P_2(\omega_\alpha) = \{\omega_\alpha\}$ and $\omega_\alpha \in \mathbf{K}_1(E)$, it follows that $P_2(\omega_\alpha) \subseteq \mathbf{K}_1(E)$ and $\omega_\alpha \in \mathbf{K}_2\mathbf{K}_1(E)$. By symmetry, we can state that $\omega_\alpha \in \mathbf{K}_j\mathbf{K}_1(E)$ for $j = 2, 3, 4$.

We can also see that $\omega_\alpha \in \mathbf{K}_m\mathbf{K}_j\mathbf{K}_1(E')$ for $j, m = 2, 3, 4$ since $\mathbf{K}_m\mathbf{K}_j\mathbf{K}_1(E') = \{\omega_\alpha, \omega'_\alpha\}$. Similarly, $\mathbf{K}_i\mathbf{K}_m\mathbf{K}_j\mathbf{K}_1(E') = \{\omega_\alpha, \omega'_\alpha\}$. By symmetry, it follows that $\omega_\alpha \in \mathbf{K}_i\mathbf{K}_j\mathbf{K}_m\mathbf{K}_1(E')$ for all $i, j, m \in N$.

We also need to show that $\omega_\alpha \in \mathbf{K}_1\mathbf{K}_j\mathbf{K}_m\mathbf{K}_i(E)$ for all $i, j, m \in N$. Taking the same event $E'$, we can easily see that $\mathbf{K}_j\mathbf{K}_m\mathbf{K}_i(E') = \{\omega_\alpha, \omega'_\alpha\}$. Since $P_1(\omega_\alpha) = \{\omega_\alpha, \omega'_\alpha\}$, it follows that $P_1(\omega_\alpha) \subseteq \mathbf{K}_j\mathbf{K}_m\mathbf{K}_i(E')$, thus $\omega_\alpha \in \mathbf{K}_1\mathbf{K}_j\mathbf{K}_m\mathbf{K}_i(E')$ for all $i, j, m \in N$.

Therefore, we can say that for all events $E$ in which $P_1(\omega_\alpha) = \{\omega_\alpha, \omega'_\alpha\} \subseteq E$, $\omega_\alpha \in \mathbf{K}_i\mathbf{K}_j..\mathbf{K}_m(E)$ for all $i, j, m \in N$. People have common knowledge about the event $\{\omega_\alpha, \omega'_\alpha\}$ in the states $\omega_\alpha$, $\omega'_\alpha$.

To sum up, example 3 shows that knowledge of the network structure is crucial to create common knowledge among the agents about the actual state. On the other hand, example 4 illustrates that agents do not need to know about all of the links in the network in order to have common knowledge about the thresholds.

We have assumed that the agents learn each others' thresholds if and only if they are within distance-2 of each other. Formally, we can state this assumption as follows:

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

**Axiom.** Consider the actual state $\hat{\omega} \in \Omega$ and the event $\hat{E} = \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_j) = (\hat{\theta}_i, \hat{\theta}_j)\}$. We assume that $\hat{\omega} \in \mathbf{K}_i(\hat{E}) \cap \mathbf{K}_j(\hat{E})$ if and only if $i \in B_j$ and $j \in B_i$ for $i, j \in N$.

This suggests that at the actual state in which $(\theta_i, \theta_j) = (\hat{\theta}_i, \hat{\theta}_j)$, person $i$ and person $j$ know the event, which is the set of states in which $(\theta_i, \theta_j) = (\hat{\theta}_i, \hat{\theta}_j)$. In other words, this event is mutual knowledge for agents $i$ and $j$ in the actual state if and only if they are within distance-2.

**Corollary.** Consider the actual state $\bar{\omega} \in \Omega$ and the event $\bar{E} = \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_j, \theta_l) = (\bar{\theta}_i, \bar{\theta}_j, \bar{\theta}_l)$ and $j, l \in N_i\}$. Then $\bar{\omega} \in \mathbf{K}_i(\bar{E}) \cap \mathbf{K}_j(\bar{E}) \cap \mathbf{K}_l(\bar{E})$.

**Proof.** Let us define some events that are consistent with the actual state $\bar{\omega}$:

$E^{ij}(\bar{\omega}) \equiv \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_j) = (\bar{\theta}_i, \bar{\theta}_j)$ and $j \in N_i\}$

$E^{il}(\bar{\omega}) \equiv \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_j, \theta_l) = (\bar{\theta}_j, \bar{\theta}_l)$ and $l \in N_i\}$

$E^{jl}(\bar{\omega}) \equiv \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_l) = (\bar{\theta}_i, \bar{\theta}_l)$ and $j, l \in N_i\}$

$E^{ijl}(\bar{\omega}) \equiv \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_j, \theta_l) = (\bar{\theta}_i, \bar{\theta}_j, \bar{\theta}_l)$ and $j, l \in N_i\}$

Since $ij \in L$, $\bar{\omega} \in \mathbf{K}_i(E^{ij}(\bar{\omega}))$ and $\bar{\omega} \in \mathbf{K}_j(E^{ij}(\bar{\omega}))$. Thus, $\bar{\omega} \in \mathbf{K}_i(E^{ij}(\bar{\omega})) \cap \mathbf{K}_j(E^{ij}(\bar{\omega}))$.

Since $il \in L$, $\bar{\omega} \in \mathbf{K}_i(E^{il}(\bar{\omega}))$ and $\bar{\omega} \in \mathbf{K}_l(E^{il}(\bar{\omega}))$. Thus, $\bar{\omega} \in \mathbf{K}_i(E^{il}(\bar{\omega})) \cap \mathbf{K}_l(E^{il}(\bar{\omega}))$.

$ij \in L$ and $il \in L$ implies $j \in B_l$. Using the Axiom, we can write $\bar{\omega} \in \mathbf{K}_j(E^{jl}(\bar{\omega})) \cap \mathbf{K}_l(E^{jl}(\bar{\omega}))$.

We need to show that $\bar{\omega} \in \mathbf{K}_i(E^{ijl}(\bar{\omega})) \cap \mathbf{K}_l(E^{ijl}(\bar{\omega})) \cap \mathbf{K}_l(E^{ijl}(\bar{\omega}))$.

If $\bar{\omega} \notin \mathbf{K}_i(E^{ijl}(\bar{\omega}))$, then $P_i(\bar{\omega}) \nsubseteq E^{ijl}$. This implies that there exists a state $\bar{\omega}' \in P_i(\bar{\omega})$ in which $\theta_j \neq \bar{\theta}_j$ or $\theta_l \neq \bar{\theta}_l$. Then either $P_i(\bar{\omega}) \nsubseteq E^{ij}(\bar{\omega})$ or $P_i(\bar{\omega}) \nsubseteq E^{il}(\bar{\omega})$. This contradicts $\bar{\omega} \in \mathbf{K}_i(E^{ij}(\bar{\omega}))$ and $\bar{\omega} \in \mathbf{K}_i(E^{il}(\bar{\omega}))$. It follows that $\bar{\omega} \in \mathbf{K}_i(E^{ijl}(\bar{\omega}))$.

Similarly, we can write $\bar{\omega} \in \mathbf{K}_j(E^{ijl}(\bar{\omega}))$ and $\bar{\omega} \in \mathbf{K}_l(E^{ijl}(\bar{\omega}))$. Therefore, $\bar{\omega} \in \mathbf{K}_i(E^{ijl}(\bar{\omega})) \cap \mathbf{K}_l(E^{ijl}(\bar{\omega})) \cap \mathbf{K}_l(E^{ijl}(\bar{\omega}))$. This completes the proof.

Also note that $E^{ijl}(\bar{\omega}) \subseteq E^{ij}(\bar{\omega}) \cap E^{il}(\bar{\omega}) \cap E^{jl}(\bar{\omega})$. We can state that if $\bar{\omega} \in \mathbf{K}_i(E^{ijl}(\bar{\omega}))$, then $\bar{\omega} \in \mathbf{K}_i(E^{ij}(\bar{\omega}))$, $\bar{\omega} \in \mathbf{K}_i(E^{jl}(\bar{\omega}))$ and $\bar{\omega} \in \mathbf{K}_i(E^{il}(\bar{\omega}))$ for all $i, j, l \in N$. In words, an agent knows an event in the actual state if he knows a subset of that event (i.e., more specific information).

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

**Lemma.** Given the actual state $\bar{\omega} \in \Omega$ and the event $\tilde{E}^{ijl}(\bar{\omega}) = \{\omega \in \Omega : \omega = [(\theta_i)_{i\in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_j, \theta_l) = (\bar{\theta}_i, \bar{\theta}_j, \bar{\theta}_l)$ and $j, l \in N_i$,$\}$, $\bar{\omega} \in \mathbf{K}^*_M(\tilde{E}^{ijl})$ where $M = \{i, j, l\}$.

**Proof.** Since $ij, il \in L$, using the Corollary we can write $\bar{\omega} \in \mathbf{K}_i(\tilde{E}^{ijl}) \cap \mathbf{K}_j(\tilde{E}^{ijl}) \cap \mathbf{K}_l(\tilde{E}^{ijl})$.

First, we will show that $\mathbf{K}_i(\tilde{E}^{ijl}) = \mathbf{K}_j(\tilde{E}^{ijl}) = \mathbf{K}_l(\tilde{E}^{ijl}) = \tilde{E}^{ijl}$ for $j, l \in N_i$.

We know that the knowledge operator satisfies $\mathbf{K}_i(E) \subseteq E$ (see Appendix). So, we can write $\mathbf{K}_i(\tilde{E}^{ijl}) \subseteq \tilde{E}^{ijl}$. Thus, we need to show that $\tilde{E}^{ijl} \subseteq \mathbf{K}_i(\tilde{E}^{ijl})$. This implies that for all $\omega \in \tilde{E}^{ijl}$, it must be that $\omega \in \mathbf{K}_i(\tilde{E}^{ijl})$. If $\tilde{\omega} \in \tilde{E}^{ijl}$, then $\tilde{\omega} = [(\theta_i)_{i\in N}, (g_{ij})_{i<j}]$ with $(\theta_i, \theta_j, \theta_l) = (\bar{\theta}_i, \bar{\theta}_j, \bar{\theta}_l)$ by definition. Since $ij, il \in L$, the Corollary suggests that $\tilde{\omega} \in \mathbf{K}_i(\tilde{E}^{ijl})$. This is also true for $j$ and $l$. Thus, $\mathbf{K}_i(\tilde{E}^{ijl}) = \mathbf{K}_j(\tilde{E}^{ijl}) = \mathbf{K}_l(\tilde{E}^{ijl}) = \tilde{E}^{ijl}$. We can write $\mathbf{K}_i\mathbf{K}_j\mathbf{K}_l(\tilde{E}^{ijl}) = \mathbf{K}_i\mathbf{K}_j(\tilde{E}^{ijl}) = \mathbf{K}_i(\tilde{E}^{ijl})$. Since $\bar{\omega} \in \mathbf{K}_i(\tilde{E}^{ijl})$, it follows that $\bar{\omega} \in \mathbf{K}_i\mathbf{K}_j\mathbf{K}_l(\tilde{E}^{ijl})$. The same argument applies to $j$ and $l$, thus, $\bar{\omega} \in \mathbf{K}^*_M(\tilde{E}^{ijl})$.

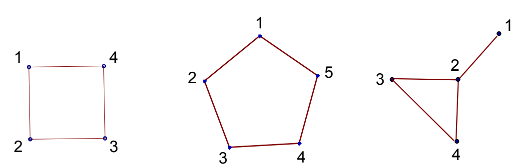
**Proposition 1.** Given the actual state $\hat{\omega} \in \Omega$ and the event $\hat{E} = \{\omega \in \Omega : \omega = [(\theta_i)_{i\in N}, (g_{ij})_{i<j}]$ with $(\theta_i)_{i\in M\subseteq N} = (\hat{\theta}_i)_{i\in M\subseteq N}\}$, $\hat{\omega} \in \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(\hat{E})$ for all $i, j, m \in M \subseteq N$ if and only if

(1) $i \in B_j \ \ \forall i, j \in M \subseteq N$

(2) $\forall jl \in L$, either (a) $ij \in L \vee il \in L$ or (b) $\exists k : ik, jk, lk \in L \ \ \forall i, j, k, l \in M \subseteq N$

The proposition specifies the necessary and sufficient conditions for a subset of agents to have common knowledge about the thresholds of everyone in the subset. Before the proof, let us present Figure 5, which illustrates what the conditions suggest.

*Figure 5.* Square, pentagon and kite




Condition (1) implies that when $i \in B_j$ (and $j \in B_i$) $\forall i, j \in M$, people either know each others' thresholds directly, or indirectly through the wall of a common friend. In all three networks above, the maximum distance between any two nodes is 2, and thus, everyone knows the thresholds of everyone else. This condition also ensures that the graph is not empty.

As we have discussed before, in the pentagon, everyone knows the thresholds of everyone else. However, in order for an agent to know if the others know about each other, the agent would need information about the network structure. For example, person 1 has to communicate with either person 3 or person 4 directly to observe the link between them. Therefore, the second condition ensures that each agent knows what a second agent knows about a third party. Thus, the event is common knowledge in the actual state among all agents.

Condition (2) suggests that for every link $jl \in L$ with $j, l \in M$, (2a) every other node $i \in M$ should be connected to one of the nodes of that link ($ij \in L$ or $il \in L$) or (2b) there must be another node $k$ that all three are connected to ($ik \in L$ , $jk \in L$ and $lk \in L$). In addition to knowing about the others' thresholds, this condition ensures that an agent also knows that the others know about each others' thresholds.

In the square, we observe that for every link (say the link between person 1 and person 2), the remaining people (person 3 and person 4) are linked to one of those people, so that they know that person 1 and person 2 communicate, and thus know about each other. This applies to every link; therefore the thresholds are common knowledge among everyone in the square. However, in the pentagon, although person 1 knows the thresholds of everyone, he cannot observe the communication between person 3 and person 4. Therefore, he does not know whether person 3 also knows about the thresholds of everyone (person 4 and person 5). Moreover, person 1 does not know whether person 2 knows about person 4, since he cannot observe that person 3 is their common neighbor. Condition (2a) suggests that person 1 should be connected to either person 3 or person 4, so that he can observe that they communicate and know about each other, which in turn allows person 1 to know that person 2 knows about person 4 and that person 5 knows about person 3. Similarly, person 2 should connect either to person 4 or to person 5, and so on.

Finally, we observe that condition (2a) is sufficient but not necessary to have common knowledge among a set of agents. We can have structures in which (2a) is not satisfied but the thresholds are still common knowledge. In the kite, we observe that person 1 is not linked to person 3 or person 4 directly (violating condition 2a), and thus, he cannot observe the link between them. However, person 1 knows that they know about each other through person 2, to whom all of them are linked. Condition (2b) captures these cases, in which agents do not need to know about all links but they need to know that the others' know about each other so that the thresholds are common knowledge among everyone.[17]

**Proof of the Proposition:**

First, let us summarize the definitions and the assumptions we have made so far:

$N_i(g) \equiv \{j \in N : ij \in L\}$ and $B_i(g) \equiv \{j \in N : j \in N_i^2\}$

$E^{i...m}(\hat{\omega}) \equiv \{\omega \in \Omega : \omega = [(\theta_i)_{i \in N}, (g_{ij})_{i<j}]$ with $(\theta_i)_{i,..,m} = (\hat{\theta}_i)_{i,..,m} \;\; m \leq n\}$

Mutual knowledge: $\mathcal{MK}(M) \equiv \hat{\omega} \in \bigcap_{i \in M} \mathbf{K}_i(E^{i...m}(\hat{\omega}))$

Common knowledge: $\mathcal{CK}(M) \equiv \hat{\omega} \in \mathbf{K}_M^*(E^{i...m}(\hat{\omega})) = \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m}(\hat{\omega}))$ for all $i, j, m \in M$

(Axiom) $i \in B_j \leftrightarrow \hat{\omega} \in \mathbf{K}_i(E^{ij}(\hat{\omega})) \cap \mathbf{K}_j(E^{ij}(\hat{\omega}))$

(Corollary) $j, l \in N_i \rightarrow \hat{\omega} \in \mathbf{K}_i(E^{ijl}(\hat{\omega})) \cap \mathbf{K}_j(E^{ijl}(\hat{\omega})) \cap \mathbf{K}_l(E^{ijl}(\hat{\omega}))$

(Lemma) $j, l \in N_i \rightarrow \hat{\omega} \in \mathbf{K}_M^*(E^{ijl}(\hat{\omega}))$ where $M = \{i, j, l\}$

The conditions are:

(1) $i \in B_j \;\; \forall i, j \in M$

(2a) $\forall jl \in L, ij \in L \vee il \in L \;\; \forall i, j, l \in M$

(2b) $\exists k : ik, jk, lk \in L \;\; \forall i, j, l \in M$

We need to show that $\mathcal{CK}(M) \leftrightarrow (1) \wedge (2a \vee 2b)$.

*If* part:

$$\text{NTS: } (1) \wedge (2a \vee 2b) \rightarrow \mathcal{CK}(M)$$
$$\equiv (1 \wedge 2a) \vee (1 \wedge 2b) \rightarrow \mathcal{CK}(M)$$

I) $(1 \wedge 2a) \rightarrow \mathcal{CK}(M)$ :

Before starting the proof, let us show that $(2a) \rightarrow (1)$ for nonempty networks:[18]

[17] If we write condition (2b) such that $k$ is interpreted as one of the two nodes ($j$ or $l$), then we could make this claim with only one condition.

[18] We do not even need condition (1) if we reformulate condition (2) such that we make sure there is at least one link in $M$.

Condition $(2a)$ implies that: If $jl \in L$, then it must be that either $ij \in L$ or $ik \in L$.

First, $jl \in L$ implies that $j \in B_l \wedge l \in B_j$.

Second, $ij \in L \to i \in B_j \wedge j \in B_i$, and since $ij \in L \wedge jk \in L$, it follows that $i \in B_k$

Finally, $ik \in L \to i \in B_k \wedge k \in B_i$, and since $ik \in L \wedge jk \in L$, $i \in B_j$.

Therefore, in both cases we get $i \in B_j \wedge j \in B_i \wedge j \in B_k \wedge k \in B_j \wedge i \in B_k \wedge k \in B_i$, which is equivalent to condition $(1)$: $i \in B_j \ \ \forall i, j \in M$.

Moreover, we also know from the pentagon example that $(1) \nrightarrow \mathcal{CK}(M)$. Since $(2a)$ implies $(1)$ for nonempty networks, condition $(1)$ only ensures that $M$ is not an empty graph. Therefore, we need to show that $(2a) \to \mathcal{CK}(M)$ for nonempty networks, i.e., that $(2a)$ is sufficient for $\mathcal{CK}(M)$ :

NTS: $\forall jl \in L,\ (ij \in L) \vee (il \in L) \to \hat{\omega} \in \mathbf{K}_i \mathbf{K}_j ... \mathbf{K}_m (E^{i...m}(\hat{\omega}))$ for all $i, j, l, m \in M$

For any $jl \in L$,

either A) $ij \in L \to i, l \in N_j \to \hat{\omega} \in \mathbf{K}^*_M (E^{ijl}(\hat{\omega}))\ \forall i, j, l \in M$ by the Lemma.

or B) $il \in L \to i, j \in N_l \to \hat{\omega} \in \mathbf{K}^*_M (E^{ijl}(\hat{\omega}))\ \forall i, j, l \in M$ by the Lemma.

This completes the first part of the proof.

Note that the kite example satisfies $\mathcal{CK}(M)$ but not $(2a)$. Together with the proof above, this implies that $(2a)$ is sufficient but not necessary for $\mathcal{CK}(M)$, i.e., $(2a) \to \mathcal{CK}(M)$ and $\mathcal{CK}(M) \nrightarrow (2a)$.

II) $(1 \wedge 2b) \to \mathcal{CK}(M)$ :

Following the argument above, we now need only to show that $(2b) \to \mathcal{CK}(M)$ for nonempty graphs. We need to show that:

$$\exists k : ik, jk, lk \in L \to \hat{\omega} \in \mathbf{K}_i \mathbf{K}_j ... \mathbf{K}_m (E^{i...m}(\hat{\omega}))\ \ \forall i, j, l, m \in M$$

If $\exists k : ik, jk, lk \in L$, then $i, j, l \in N_k$.

A) $i, j, l \in N_k \to \hat{\omega} \in \mathbf{K}_i (E^{ijl}(\hat{\omega})) \cap \mathbf{K}_j (E^{ijl}(\hat{\omega})) \cap \mathbf{K}_j (E^{ilj}(\hat{\omega}))$ by the Corollary.

B) Following the Lemma, we have $i, j, l \in N_k \to \hat{\omega} \in \mathbf{K}_M (E^{ijl}(\hat{\omega}))\ \ \forall i, j, l \in M$.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

(I) and (II) complete the proof of $(1) \wedge (2a \vee 2b) \rightarrow \mathcal{CK}(M)$.

<u>*Only if* part:</u>

$$\mathcal{CK}(M) \rightarrow (1) \wedge (2a \vee 2b)$$
$$\equiv (\mathcal{CK}(M) \rightarrow (1)) \wedge (\mathcal{CK}(M) \rightarrow (2a \vee 2b))$$

I) $\mathcal{CK}(M) \rightarrow (1)$ :

$$\begin{aligned}
&\equiv \quad \neg(1) \rightarrow \neg\mathcal{CK}(M) \\
&\equiv \quad \neg(i \in B_j \ \forall i, j \in M) \rightarrow \hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m}(\hat{\omega})) \ \ \forall i, j, m \in M \\
&\qquad \rightarrow \hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j(E^{ij}(\hat{\omega})) \ \ \forall i, j \in M
\end{aligned}$$

In order to show this we can write the following:

$$\begin{aligned}
\neg(i \ \in \ B_j \ \forall i, j \in M) &\rightarrow \exists i \in M : i \notin B_j \\
&\leftrightarrow \hat{\omega} \notin \mathbf{K}_i(E^{ij}(\hat{\omega})) \cap \mathbf{K}_j(E^{ij}(\hat{\omega})) \text{ by the Axiom.} \\
&\rightarrow \hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j(E^{ij}(\hat{\omega})) \ \ \forall i, j \in M \\
&\rightarrow \hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_l(E^{ijl}(\hat{\omega})) \ \ \forall i, j, l \in M
\end{aligned}$$

This completes the first part of the proof.

An alternative proof is as follows:

$$\begin{aligned}
\mathcal{CK}(M) \quad &\equiv \quad \hat{\omega} \in \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m}(\hat{\omega})) \text{ for all } i, j, m \in M \\
&\rightarrow \quad \hat{\omega} \in \mathbf{K}_i\mathbf{K}_j\mathbf{K}_l(E^{ijl}(\hat{\omega})) \text{ for any } i, j, l \in M. \\
&\rightarrow \quad \hat{\omega} \in \mathbf{K}_i(E^{ijl}(\hat{\omega})) \cap \mathbf{K}_j(E^{ijl}(\hat{\omega})) \cap \mathbf{K}_l(E^{ijl}(\hat{\omega}))
\end{aligned}$$

Since $E^{ijl}(\hat{\omega}) \subseteq E^{ij}(\hat{\omega})$, it follows that $\hat{\omega} \in E^{ij}(\hat{\omega})$ and $\hat{\omega} \in \mathbf{K}_i(E^{ij}(\hat{\omega})) \cap \mathbf{K}_j(E^{ij}(\hat{\omega}))$.

Following the Axiom, we have $i \in B_j \wedge j \in B_i$.

Similarly, $E^{ijl}(\hat{\omega}) \subseteq E^{il}(\hat{\omega})$ and $E^{ijl}(\hat{\omega}) \subseteq E^{jl}(\hat{\omega})$.

Thus, we have $\hat{\omega} \in \mathbf{K}_i(E^{il}(\hat{\omega})) \cap \mathbf{K}_l(E^{il}(\hat{\omega}))$ and $\hat{\omega} \in \mathbf{K}_j(E^{jl}(\hat{\omega})) \cap \mathbf{K}_l(E^{jl}(\hat{\omega}))$.

These imply that $i \in B_l$ and $j \in B_l$. Therefore, we have shown that $i \in B_j \ \ \forall i, j \in M$.

We also know from the pentagon example that $(1) \nrightarrow \mathcal{CK}(M)$, i.e., everyone is within distance-2 but there is no common knowledge among all agents. Therefore, the relationship between two sets that satisfy Condition $(1)$ and $\mathcal{CK}(M)$ can be written as $\mathcal{CK}(M) \subset (1)$.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

II) $\mathcal{CK}(M) \rightarrow (2a \vee 2b)$ :

$$\begin{aligned} &\equiv \neg(2a \vee 2b) \rightarrow \neg\mathcal{CK}(M) \\ &\equiv (\neg 2a) \wedge (\neg 2b) \rightarrow \neg\mathcal{CK}(M) \\ &\equiv \neg(\forall jl \in L,\ (ij \in L) \vee (il \in L) \vee (\exists k : ik, jk, lk \in L\, \forall i, j, l \in M)) \\ &\quad \rightarrow \hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j ... \mathbf{K}_m(E^{i...m}(\hat{\omega}))\ \ \forall i, j, m \in M \end{aligned}$$

In order to show this we can write the following:

$$\begin{aligned} &\neg(\forall jl \in L, (ij \in L) \vee (il \in L) \vee (\exists k : ik, jk, lk \in L\, \forall i, j, l \in M)) \\ &\equiv \exists i, j, l \in M \text{ with } jl \in L : (ij \notin L) \wedge (il \notin L) \wedge (\nexists k : ik, jk, lk \in L\, \forall i, j, l \in M) \end{aligned}$$

First note that, if $M$ is an empty subgraph so that $N_i = \{\}\ \ \forall i \in M$, then $B_i = \{\}\ \ \forall i \in M$. By the Axiom, $\hat{\omega} \notin \mathbf{K}_i(E^{ij}(\hat{\omega})) \cap \mathbf{K}_j(E^{ij}(\hat{\omega}))$ for any $i, j \in M$. Since $E^{ij..m}(\hat{\omega}) \subseteq E^{ij..m}(\hat{\omega})$, $\hat{\omega} \notin \bigcap_{i \in M} \mathbf{K}_i(E^{ij..m}(\hat{\omega}))$. It follows that $\hat{\omega} \notin \mathbf{K}^*_M(E^{ij..m}(\hat{\omega}))$ for $i, j.., m \in M$.

For nonempty subgraphs, note that $\exists j, l \in M$ with $jl \in L$.

If $\exists i \in M : (ij \notin L) \wedge (il \notin L) \wedge (\nexists k : ik, jk, lk \in L\, \forall i, j, l \in M)$, then either ($i \in B_j \backslash N_j$ and $i \in B_l \backslash N_l$ ) or ( $i \notin B_j$ and/or $i \notin B_l$ ).

The latter (i.e., $\exists i \in M : i \notin B_j$, which is equivalent to $\neg(1)$ ) implies that $\hat{\omega} \notin \mathbf{K}_i(E^{ij}(\hat{\omega})) \cap \mathbf{K}_j(E^{ij}(\hat{\omega}))$ for any $i, j \in M$ and/or $\hat{\omega} \notin \mathbf{K}_i(E^{il}(\hat{\omega})) \cap \mathbf{K}_l(E^{il}(\hat{\omega}))$ for any $i, l \in M$. As we have shown in the first part of the proof $\neg(1) \rightarrow \neg\mathcal{CK}(M)$.

On the other hand, if $i \in B_j \backslash N_j$ and $i \in B_l \backslash N_l$, then by assumption $\hat{\omega} \in \mathbf{K}_i(E^{ij}(\hat{\omega})) \cap \mathbf{K}_j(E^{ij}(\hat{\omega}))$ and $\hat{\omega} \in \mathbf{K}_i(E^{il}(\hat{\omega})) \cap \mathbf{K}_l(E^{il}(\hat{\omega}))$ for $i, j, l \in M$. We also know that $jl \in L \rightarrow j \in B_l$, thus, $\hat{\omega} \in \mathbf{K}_j(E^{jl}(\hat{\omega})) \cap \mathbf{K}_l(E^{jl}(\hat{\omega}))$.

Moreover, if $i \in B_j \backslash N_j$, then $\exists m \neq l : im, jm \in L$. It must be that $lm \notin L$, otherwise it violates condition (2b), i.e., $\nexists k : ik, jk, lk \in L\, \forall i, j, l \in M$ . Similarly, since $i \in B_l / N_l$, it must be that $\exists m' : im', lm' \in L$ with $jm' \notin L$. Therefore, as in the pentagon example $i, j, l, m, m'$ are all within distance-2, thus, their thresholds are *mutually known*, i.e., $\hat{\omega} \in \mathbf{K}_i(E^{ijlmm'}(\hat{\omega})) \cap \mathbf{K}_j(E^{ijlmm'}(\hat{\omega})) \cap \mathbf{K}_l(E^{ijlmm'}(\hat{\omega})) \cap \mathbf{K}_m(E^{ijlmm'}(\hat{\omega})) \cap \mathbf{K}_{m'}(E^{ijlmm'}(\hat{\omega}))$.

We will focus on the pentagon to show that the event $E^{ijlmm'}(\hat{\omega})$ is not *common knowledge*, i.e., $\hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j\mathbf{K}_l\mathbf{K}_m\mathbf{K}_{m'}(E^{ijlmm'}(\hat{\omega}))$.

Let us take agents $j$ and $l$ with $jl \in L$. We have defined $\mathbf{K}_l\mathbf{K}_j(E) \equiv \{\omega \in \Omega : P_l(\omega) \subseteq \mathbf{K}_j(E)\}$.

We need to show whether $P_l(\hat{\omega}) \subseteq \mathbf{K}_j(E^{ijlmm'}(\hat{\omega}))$. If true, this statement implies that agent $l$ knows that $j$ knows the event $E^{ijlmm'}(\hat{\omega})$ in the actual state $\hat{\omega}$.

We can write $P_l(\hat{\omega}) = \{\hat{\omega}, \hat{\omega}'\}$ where $\hat{\omega}' = [(\theta_i)_{i,..,m'} = (\hat{\theta}_i)_{i,..,m'}, (g_{ij})_{i<j}, im \notin L]$, i.e., a state in which all thresholds and links are consistent with the state except that $im \notin L$.

Let us define another state $\hat{\omega}'' = [(\theta_j)_{j,..,m'} = (\hat{\theta}_j)_{j,..,m'}, \theta_i = \hat{\theta}_i', (g_{ij})_{i<j}, im \notin L]$, in which $\theta_i = \hat{\theta}_i' \neq \hat{\theta}_i$ and $im \notin L$.

We can see that $\hat{\omega}'' \in P_j(\hat{\omega}')$ since $i \notin B_j$ when $im \notin L$. The states $\hat{\omega}'$ and $\hat{\omega}''$ are in the same possibility set of person $j$, meaning that he cannot distinguish between the states $\hat{\omega}'$ and $\hat{\omega}''$ when $im \notin L$. However, $\hat{\omega}'' \notin E^{ijlmm'}(\hat{\omega})$ since $\theta_i = \hat{\theta}_i' \neq \hat{\theta}_i$. Since $P_j(\hat{\omega}') \nsubseteq E^{ijlmm'}(\hat{\omega})$, $\hat{\omega}' \notin \mathbf{K}_j(E^{ijlmm'}(\hat{\omega}))$. It follows that $P_l(\hat{\omega}) = \{\hat{\omega}, \hat{\omega}'\} \nsubseteq \mathbf{K}_j(E^{ijlmm'}(\hat{\omega}))$, thus, $\hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j(E^{ijlmm'}(\hat{\omega}))$. Therefore, $\hat{\omega} \notin \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m}(\hat{\omega}))\ \forall i, j, l, m \in M$. This completes the proof.

**Definition.** A *complete bipartite graph* is a bipartite graph (i.e., a set of graph vertices decomposed into two disjoint sets such that no two graph vertices within the same set are adjacent) in which every pair of graph vertices in the two sets are adjacent.

**Proposition 2.** Given a graph $G(N, L)$, if there exists a set of people $M \subseteq N$ such that $\hat{\omega} \in \mathbf{K}_M^*(E^{i...m}(\hat{\omega})) = \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m}(\hat{\omega}))$ for all $i, j, m \in M$, then a sub-graph induced by this set $M$ must be a complete bipartite graph.

**Proof.** Following the definition of complete bipartite graphs, let us consider two sets $M_1, M_2 \subseteq M \subseteq N$ such that every pair of graph vertices in the two sets are adjacent. Formally, $ij \in L$ $\forall i \in M_1$ and $\forall j \in M_2$. This satisfies Condition (1) of Proposition 1, i.e., $i \in Bj\ \forall i, j \in M$. Now, let us take $i_1 \in M_1$. Since $j_{1,}j_{2,..}j_{m_2} \in N_{i_1}$, using the Lemma we have $\hat{\omega} \in \mathbf{K}^*_{i_1 \cup M_2}(E^{i_1 j_1 j_2 ... m_2}(\hat{\omega}))$. Similarly, if we take $i_2 \in M_1$, then we have $\hat{\omega} \in \mathbf{K}^*_{i_2 \cup M_2}(E^{i_2 j_1 j_2 ... m_2}(\hat{\omega}))$. We can repeat this procedure for all $i \in M_1$. Since $i_1, i_2, .., i_{m_1} \in N_{j_1}$, it also follows from the Lemma that $\hat{\omega} \in \mathbf{K}^*_{j_1 \cup M_1}(E^{i_1 .. i_{m_1} . j_1}(\hat{\omega}))$. If $M_1$or $M_2$ is a singleton (as in the star graph), then

Condition (2b) of Proposition 1, i.e., $\exists k : ik, jk, lk \in L \ \ \forall i, j, l \in M$, is also satisfied; and this completes the proof.

In general, let us take any $m \in M_1$. Since $ij \in L \ \ \forall i \in M_1$ and $\forall j \in M_2$, we have $\forall ij \in L$, $mj \in L$. This is equivalent to Condition (2a) of Proposition 2. Therefore, $\hat{\omega} \in \mathbf{K}^*_{M_1 \cup M_2}(E^{i...m}(\hat{\omega}))$.

*Figure 6.* Complete Bipartite Graphs

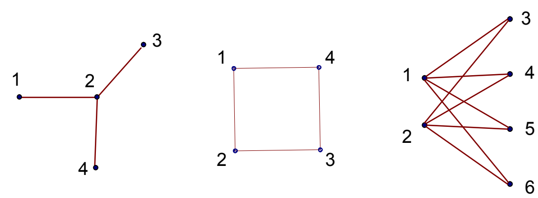



Figure 6 illustrates three examples of complete bipartite graphs, including the star and the square, in which the thresholds are commonly known among all agents. As discussed before, the star is the extreme case in which common knowledge is obtained only though the hub's wall.

## 3.5 Deviations and the Equilibrium

### 3.5.1 Static Framework

**Deviations:** A subset of agents $M \subseteq N$ revolts when

(1) $\hat{\omega} \in \mathbf{K}^*_M(E^{i...m}) = \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m})$ for all $i, j, m \in M$

(2) $\theta_i \leq \#M \ \ \forall i \in M$

Therefore, the subset $M$ deviates when the number of people in $M$ who share the common knowledge of the thresholds is sufficiently high. In other words, they revolt when it is common knowledge (everyone knows that everyone knows) that if all $i \in M$ deviate, then everyone gains. Coming back to the pentagon example, person 1 knows that person 3 and person 4 will gain by deviating (since he knows the thresholds of everyone) but he does not know whether they know that they will gain by deviating, i.e., it is not common knowledge.

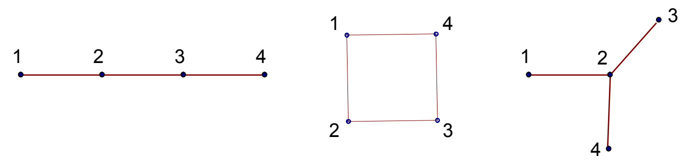
**Equilibrium:** We say that an action profile $a^* = (a_1^*, a_2^*, ..., a_n^*)$ is an equilibrium if and only if $\nexists M \subset N$ :

(1) $\hat{\omega} \in \mathbf{K}_M^*(E^{i...m}) = \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m})$ for all $i, j, m \in M$

(2) $\forall i \in M \;\; U_i(r, a_{-M}^*) \geq U_i(a^*)$ with strict inequality for some $j \in M$.

Having defined deviations and the equilibrium, we can now re-analyze some examples and observe the interaction between the network structure, common knowledge and the thresholds of people.

*Figure 7.* Line, star and square with $\theta_i = 4 \;\; \forall i$

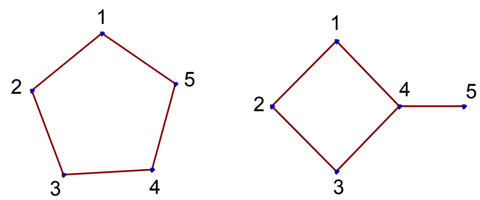



In the line, the actual state of the world can be written as $\omega_\alpha = (4, 4, 4, 4, 1, 0, 0, 1, 0, 1)$. As discussed before, person 2 and person 3 know that the state is $\omega_\alpha$. Person 1 knows that person 2 has threshold 4 and he knows that person 3 has threshold 4 through the wall of person 2, but he does not know either the threshold of person 4 or the link between person 3 and 4. Hence, all person 1 knows is that the thresholds are in the set $\{4441, 4442, 4443, 4444, 4445\}$; thus, person 1 does not revolt. Similarly, person 4 does not revolt. Although person 2 and person 3 know that the state is $\omega_\alpha$, in this example no one revolts, i.e., $a^* = (s, s, s, s)$.

In the square, the actual state of the world is $\omega_\alpha = (4, 4, 4, 4, 1, 0, 1, 1, 0, 1)$ and the types are common knowledge among all agents ($\forall jl \in L$, either $ij \in L$ or $il \in L \;\; \forall i, j, k \in M$). Thus, the state $\omega_\alpha$ is commonly known. We also observe that $\theta_i \leq \#M \;\; \forall i \in M$. Therefore, $a^* = (r, r, r, r)$ in the square. Note that, even though person 2's knowledge about everyone's thresholds is the same in both the line and the square, she only revolts in the latter because she knows that person 1 and person 4 communicate.

When we consider the star, the actual state can be written as $\omega_\alpha = (4, 4, 4, 4, 1, 0, 0, 1, 1, 0)$ and, as we have discussed before, the types become common knowledge through the wall of the hub ($\exists k : ik, jk, lk \in L \;\; \forall i, j, k, l \in M$). Since $\theta_i \leq \#M \;\; \forall i \in M$, everyone revolts in the star i.e., $a^* = (r, r, r, r)$.

*Figure 8.* Pentagon and kite with $\theta_i = 4 \quad \forall i$



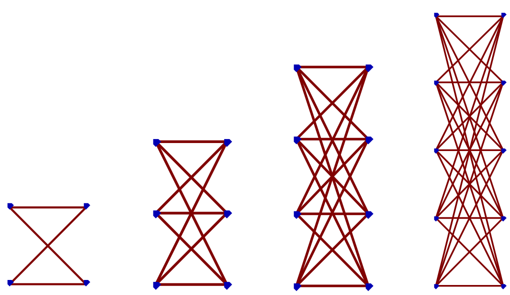

We have already studied the pentagon in which $\omega_\alpha = (4,4,4,4,4,1,0,0,1,1,0,0,1,0,1)$. The thresholds are mutually known by all agents (everyone is within distance-2), however, due to the limited information on the network structure, the actual state is not common knowledge among the agents. Thus, they do not revolt, i.e., $a^* = (s,s,s,s,s)$.

Note that we can find smaller subsets of agents (stars) $M_1,..,M_5 \subseteq N$ in the pentagon such as $M_1 = \{1,2,3\}$, $M_2 = \{2,3,4\},..,M_5 = \{5,1,2\}$. The thresholds of the agents in these subsets are commonly known among those agents since all of them share a common wall. Therefore, if the thresholds are $\theta_i = 3 \quad \forall i$, then $\theta_i \leq \#M$ is satisfied for each subset. In this case, $a^* = (r,r,r,r,r)$.

However, even if the thresholds are $\theta_i = 4 < \#N$ in the pentagon, $a^* = (s,s,s,s,s)$ because only in subsets with 3 agents do they commonly know each others' thresholds. $\theta_i \leq \#M$ is not satisfied.

In the kite, we have $M_1 = \{1,2,3,4\}$ and $M_2 = \{1,3,4,5\}$ such that the thresholds are common knowledge in each subset and $\theta_i \leq \#M \quad \forall i \in M$. Agents in the same subset know that if they jointly deviate and revolt, everyone in that set will gain. Therefore, everyone revolts in the kite, and $a^* = (r,r,r,r,r)$.

*Figure 9.* Regular Complete Bipartite Graphs



Finally, following Proposition 2, we know that in the complete bipartite graphs, there is common knowledge of thresholds among all of the people. Everyone knows everyone's thresholds and everyone knows that everyone knows everyone's thresholds. Therefore, $M$ revolts so long as $\theta_i \leq \#M$.[19]

The regular complete bipartite graphs in Figure 9 illustrate that the number of people for whom the thresholds are common knowledge is not limited. As long as the structure remains the same, infinitely many agents can have common knowledge about the state of the world. In this case, whenever there exists a subset $M \subseteq N$ for which $\theta_i \leq \#M \ \ \forall i \in M$, $a_i = r$ for all $i \in M$.

### 3.5.2 Dynamic Framework

In this section, we introduce dynamics to the model described above. We argue that in each time period, people obtain local information about the past actions of others in addition to information about their thresholds. The number of revolters in the previous periods is not publicly known, but only locally observed. We aim to study the role of online social networks as communication tools; thus, we argue that people only trust in the information they receive from their local environment and not from the mass media.

Focusing on a *facebook*-type mechanism, we assume that people post their actions on their direct neighbors' "*walls*" and that past actions are observed by the neighbors at distance-2 through the wall of a common neighbor. We argue that people make inferences about the number of revolters in the previous periods by using knowledge about their neighbors' thresholds and by learning about their neighbors' past actions. In other words, learning that someone with threshold $\bar{\theta}$ revolted at $t-1$ reveals that the number of revolters must be at least $\bar{\theta}$, since everyone knows that people only revolt when they know that there is sufficient number of revolters. Therefore, we introduce more sophisticated behavior in the dynamic framework and we need to assume that it is commonly known that everyone is sophisticated enough to make these inferences.

[19] We will see in the following section that the equilibrium is the same in both the static and the dynamic frameworks, since there can be no more information revealed as a result of past actions because all agents already share all information.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

The set of states is given by $\Omega(t) = \Theta^n \times A^n(t-1) \times G$ with $G = \{0,1\}^{C_2^n}$ and $A(t-1) = \{r, s\}$. Formally, a state can be written as $\omega = [(\theta_i)_{i\in N}, (a_i)_{i\in N}, (g_{ij})_{i<j}]$. Thus, a state will be a $(2n + C_2^n)$-*tuple,* for example, $\omega(t) = (2,2,2,s,r,r,0,0,1)$ for a network with 3 people with $\theta_i = 2$ where $g_{12} = 0, g_{13} = 0, g_{23} = 1$ and $a_1(t-1) = s$, $a_2(t-1) = r$, $a_3(t-1) = r$.

**The timing of the dynamic game**

$t < 0$. Nature determines people's preferences (thresholds to revolt), which are private information.

$t = 0$. People communicate their thresholds and action plans with their direct neighbors (which are also observed by second-degree neighbors). They then choose their own actions, $r$ or $s$, (decide whether to revolt or not).

$t > 0$. The remaining people, who have chosen $s$ in the previous periods, observe the thresholds and past actions of the people in their '*ball*' and decide whether to revolt or not.[20]

**The Law of Motion** ($a_i(t)$):

$a_i(t<0) = s \quad \forall i \in N$

$a_{i\in M}(t=0) = r \Longleftrightarrow \exists M \subseteq N :$

(1) $\hat{\omega} \in \mathbf{K}_M^*(E^{i...m}) = \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m})$ for all $i, j, m \in M$

(2) $\theta_i \leq \#M \;\; \forall i \in M$

$a_{i\in M}(t>0) = r \Longleftrightarrow \exists M \subseteq N :$

(1) $\hat{\omega} \in \mathbf{K}_M^*(E^{i...m}) = \mathbf{K}_i\mathbf{K}_j...\mathbf{K}_m(E^{i...m})$ for all $i, j, m \in M$

(2) $\exists K \subseteq N$ with $a_k(t-1) = r$ and $\hat{\omega} \in \mathbf{K}_M^*(E^{ki...m})$ for all $i, j, m \in M$ and $k \in K$:

(2.a) $\theta_i - \max\{\max\{\theta_k\}_{k\in K}, \#K\} \leq \#M \;\; \forall i \in M$

(2.b) If $\theta_k = 1$ for some $k \in K$, then $\theta_i - \sum_{k\in K} \theta_k \leq \#M \;\; \forall i \in M$.

[20] In this set framework, we only allow for moving from $s$ to $r$. One can think about introducing the possibility of changing the action from $r$ to $s$.

 DOI: 10.2870/63947

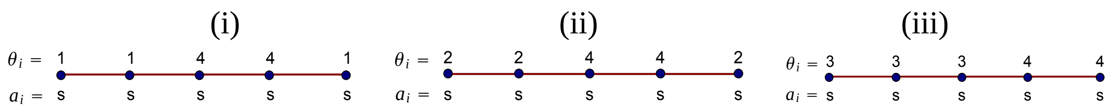
### 3.5.3 Examples

We now present the following examples to explain the timing of the game and the law of motion formulated above.

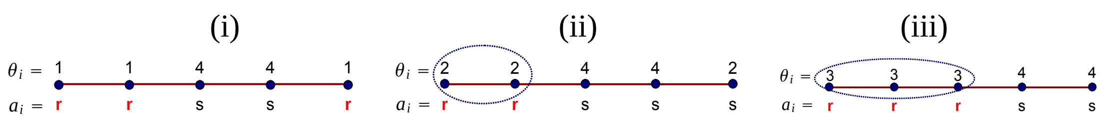
**Line**

First, we analyze the line graph with three different distributions of thresholds. The figures illustrate when and how people decide to revolt in periods $t = 0$ and $t = 1$ :

$\mathbf{t} < \mathbf{0}$. Nature determines people's private thresholds. We assume that $a_i(t < 0) = s \quad \forall i \in N$.




$\mathbf{t} = \mathbf{0}$. Agents communicate their thresholds with their neighbors and choose $r$ or $s$.

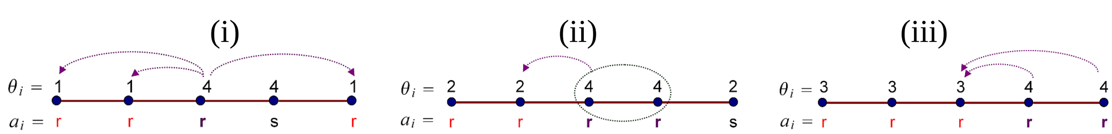



(i) The agents with $\theta_i = 1$, unilaterally revolt.

(ii) Communication between the agents with $\theta_i = 2$ allows them to coordinate and jointly revolt.

(iii) The agents with $\theta_i = 3$ can coordinate and revolt since their thresholds are commonly known among the group.

$\mathbf{t} = \mathbf{1}$. Agents communicate their past actions. People who have chosen $s$ in the previous periods decide whether to choose $s$ or $r$.




(i) The agent with $\theta_{i \in M} = 4$ can observe the subset $K$ of agents with $\theta_k = 1$ and $a_k(t = 0) = r$. Since $\theta_i - \sum_{k \in K} \theta_k \leq \#M$, he unilaterally revolts.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

(ii) The two agents for whom $\theta_{i\in M} = 4$ observe $\theta_k = 2$ with $a_k(t = 0) = r$. The threshold of $k$ is mutually and commonly known by the agents with $\theta_{i\in M} = 4$. Since $\theta_i - \max\{\theta_k\}_{k\in K} \leq \#M$ $\forall i \in M$, they coordinate and jointly revolt. Note that, the agent on the left can observe two agents on his left with $\theta_k = 2$ and $a_k(t = 0) = r$. Thus, he knows the number of revolters is at least two. However, this is not enough for him to revolt unilaterally since his threshold is $4$. He needs to coordinate with the other agent on his right. Since both of them commonly know the agent with $\theta_k = 2$ and $a_k(t = 0) = r$, they coordinate and revolt.

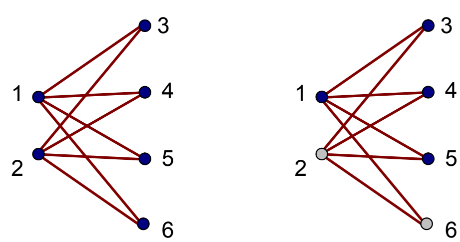
(iii) The remaining agents with $\theta_{i\in M} = 4$ can observe the agent with $\theta_k = 3$ with $a_k(t = 0) = r$. Although they do not observe all three agents with $\theta_k = 3$ and $a_k(t = 0) = r$, the information about the past action and the threshold of only one of the revolters will allow them to infer that there are at least three revolters. In this case, they do not need to coordinate since this information is enough for both of them to revolt unilaterally.

### Complete Bipartite Graphs

We now present some examples involving complete bipartite graphs in which we know that the thresholds of all agents are common knowledge. The examples are different but they build upon each other. Moreover, we conduct simple simulations to show that similar sub-networks can result in different network dynamics

**Case 1**: $n = 6$ and $\theta_i = 6$ for $i = 1, 2, .., 6$

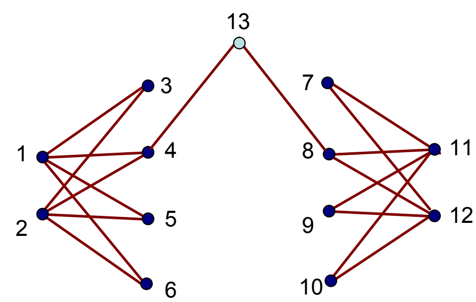
**Case 2**: $n = 6$ and $\theta_i = 4$ for $i = 1, 3, 4, 5$ and $\theta_i = 7$ for $i = 2, 6$




Case 1 Case 2

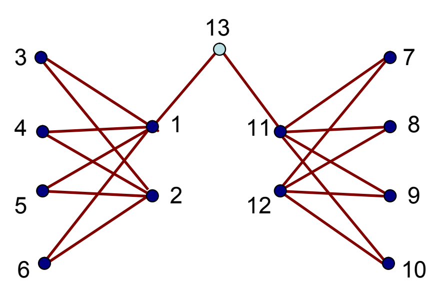
Case 1 and Case 2 illustrate that either everyone or a subset of the individuals who share common knowledge can coordinate and revolt at $t = 0$, depending on the thresholds. In Case 1, everyone revolts at $t = 0$, while in Case 2 agents with threshold $7$ do not revolt. Since there is no more information to be revealed in later stages, agents 2 and 6 never revolt in Case 2.

**Case 3:** $n = 13$ and $\theta_i = 6$ for $i = 1, 2, ..12$ and $\theta_i = 5$ for $i = 13$




In this case, the agents in each complete bipartite network choose $r$ at $t = 0$ since the number of people who share common knowledge is equal to their thresholds. Although individual 13 is connected to both sets, he does not share the same common knowledge. He observes two distinct sets with 4 people in each (including himself) that share common knowledge. In particular, he observes two star networks (with 4 people in each) in which agents 4 and 8 are the hubs. He only knows that the others who have threshold $6$ communicate with 3 other people, and this information is not enough for him to know whether they will revolt or not. Hence, he does not revolt at $t = 0$. At $t = 1$, it is enough for him to observe that a person with threshold $6$ has revolted to infer that there are at least $6$ participants. Therefore, he revolts at $t = 1$. In this case, he actually observes 6 people who have revolted. If individual 13 had threshold $7$ instead of $5$, he could still revolt at $t = 1$.

**Case 4**: $n = 13$ and $\theta_i = 6$ for $i = 1, 2, ..12$ and $\theta_i = 5$ for $i = 13$

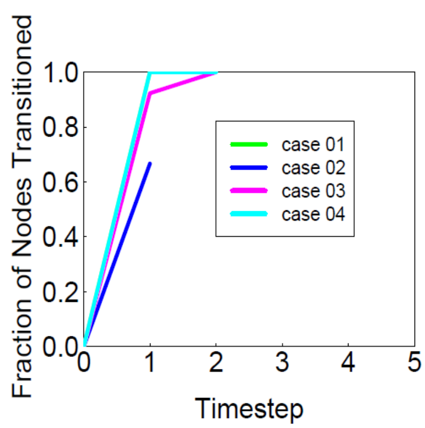



Case 4 differs from Case 3 depending on who person 13 is connected to. In contrast to the previous case, here person 13 also revolts at $t = 0$ because the number of people who share


DOI: 10.2870/63947

common knowledge in both sets that person 13 can observe is 6. Now, person 13 knows that the others can observe a sufficient number of people to make them revolt. Therefore, he also revolts. Since person 13 is connected to agents with higher degrees than in the previous case, he gets more information, which allows him to revolt earlier.

We have conducted some simple simulations to implement the cases considered above. The following figure illustrates how many individuals choose to revolts at each stage.[21] We can observe that even similar sub-networks can result in different network dynamics: (i) only a subset of the agents, who share common knowledge, might revolt depending on the thresholds, (ii) it might take longer for all agents to revolt depending on who they are connected to, even though they share the same common knowledge.




So far, we have presented simple examples to illustrate the mechanism of the dynamic model and the novelty of the sophisticated behavior we have introduced. Although we have not characterized any equilibrium, our next step will be to obtain more generalized theoretical results by studying particular networks and to investigate thoroughly the role of the interaction between the network structure and the distribution of thresholds.

[21] Note that the first stage of our game ($t = 0$) corresponds to $t = 1$ in the figure.


DOI: 10.2870/63947

## 3.6 Homophily

In this section, we show how the likelihood of being connected to people of the same type, known as *homophily*, affects the dynamics of collective action. This section aims to illustrate the trade-off between the number of initial revolters and the speed of the spread of revolution. We argue that in order for *activists* (people with low thresholds) to revolt in the beginning, the degree of homophily needs to be sufficiently high that they can coordinate. A further increase in the probability of being linked to the same type enables higher threshold agents to coordinate and revolt at the initial stages. However, as homophily gets very high, the spread of information to *conservatives* (people with the highest threshold in the population) becomes more difficult thus, it takes longer for everyone to revolt. In this section, we analyze some simple examples to illustrate this trade-off in order to motivate future theoretical analysis of homophily.

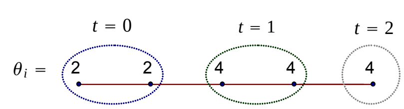
### 3.6.1 Examples

**Line**

Consider a line graph with $n = 5$, $\theta_i = 2$ for $i = 1, 2$ and $\theta_i = 4$ for $i = 3, 4, 5$.

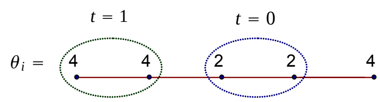
(a) 


There is no revolt in the first example. People with threshold $\theta_i = 2$ (*activists*) cannot coordinate since they cannot communicate with each other and $a = (s, s, s, s, s)$ for $t \geq 0$.

(b) 

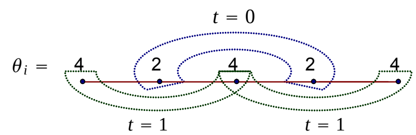

When the activists are direct neighbors, they can coordinate and revolt at $t = 0$. The *conservatives* who can observe their action coordinate and revolt at $t = 1$. Finally, at $t = 2$, the last person observes the action, $r$, of his first and second-degree neighbors and, since he knows that their threshold is 4, he infers that there must be at least 2 more revolters. This information allows him to revolt at $t = 2$.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947




(c)

In this case, the last conservative only observes the action of the activists. Thus, he knows that there are 2 people who chose to revolt. He cannot observe the others however, and so the information he has is not enough for him to revolt. The equilibrium is $a^* = (r, r, r, r, s)$.



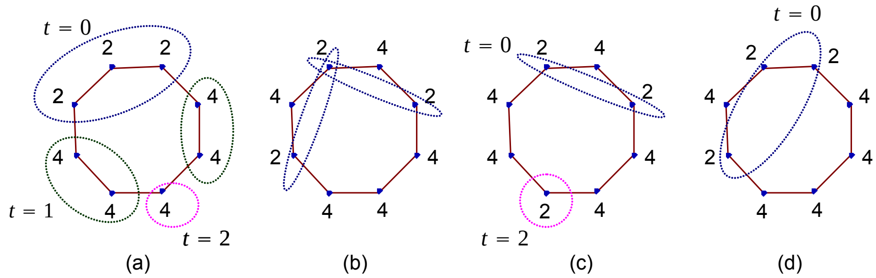

(d)

Finally, in the last figure, the activists are within distance-2, they are still able to coordinate and revolt at $t = 0$. The conservatives can observe the action of the activists and they are also close enough to coordinate and revolt at $t = 1$. Note that the common knowledge and the coordination is not among the three conservatives. Person 1 and person 5 do not share common knowledge of thresholds. If the thresholds of the conservatives were 5, they would not be able to coordinate and jointly revolt even though there is a sufficient number of people to make all of them revolt at $t = 1$.

In the simple examples above, we first observe that the *activists* should be close enough to coordinate so that there is revolt at $t = 0$. Second, there should be a sufficient number of *conservatives* who are close enough to coordinate in order to allow for the spread of collective action. Finally, when the activists are dispersed but close enough to coordinate, revolt is faster as we saw in the last example. However, one might argue that the degree of the players also plays a role in the line graph. Therefore, in the following example we study ring networks in which all agents have 2 neighbors.

**Ring**

Now let us consider a ring with $n = 8$, $\theta_i = 2$ for $i = 1, 2, 3$ and $\theta_i = 4$ for $i = 4, .., 8$




(a) The *activists* share common knowledge and jointly revolt at $t = 0$. The *conservatives* that are close enough to the activists (within distance-2) observe the past actions of the activists and infer that there are at least 2 revolters. They thus coordinate, and revolt at $t = 1$. One of the agents with threshold $4$ does not observe the past action of the activists. He revolts at $t = 2$ after he observes the past actions of the other conservatives.

(b) The activists are not direct neighbors, but they are close enough to coordinate in pairs. They therefore revolt at $t = 0$. Since the activists are dispersed, their actions can be observed by the rest of the people. The conservatives are close enough to coordinate; thus, all of them revolt at $t = 1$.

These two figures confirm our observations from the previous line examples. First, the activists need to be close enough to coordinate for revolt to start. Second, when the activists are dispersed and can coordinate; the spread of the collective action is faster.

(c) Here, two activists can coordinate and revolt at $t = 0$. The conservatives are able to observe and coordinate in pairs to revolt at $t = 1$. Note that, in this case, the remaining activist revolts at $t = 2$ after he observes the actions of the conservatives. Here we obtain a *late activist* who revolts even after the conservatives due to limited information.

(d) In this example, the activists are close enough to communicate. They can coordinate in pairs and all revolt at $t = 0$. Here, the conservative who is linked to the activists has enough information to allow him to revolt at $t = 1$ without the need to communicate and coordinate with other conservatives.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947

These examples motivate us to develop a theoretical framework to study the role of network homophily in collective action. For example, one can parameterize the analysis by defining homophily as the fraction of an agent's neighbors that are of the same type. Formally,

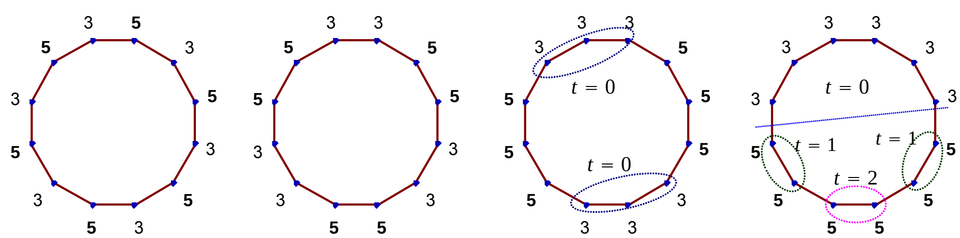
$$\rho_i \equiv \frac{\#\{j \in N_i : \theta_j = \theta_i\}}{\#N_i} \text{ and } \rho(G) \equiv \frac{\sum_{i \in N} p_i}{\#N}$$

Let us assume that there are two types of thresholds: $\theta_i \in \{L, H\}$, where $L$ is low and $H$ denotes the high threshold. Moreover, let $N^L$ and $N^H$ define the number of low and high types in the population, respectively. We can define the fraction of types as

$$n^L \equiv \frac{N^L}{N^L + N^H} \text{ and } n^H \equiv \frac{N^H}{N^L + N^H}$$

Here we will assume that $n^L = n^H = \frac{1}{2}$ for simplicity. We will also assume that the degree of homophily for both types ($\rho^L$ and $p^H$, respectively) is symmetric.

We can now analyze the trade-offs in the previous examples, in a systematic way using the following example. Consider a ring with $\theta_i \in \{3, 5\}$ and $N^L = N^H = 6$.




(a) (b) (c) (d)

These networks differ with respect to the degree of homophily. In the first ring, we observe that all individuals are adjacent to people of different types. Formally, $\rho_i = 0 \quad \forall i \in N$ and $p(a) = 0$. In the second graph, everyone has one neighbor who is of the same type, i.e., $\rho_i = 1/2 \quad \forall i \in N$, hence $p(b) = 1/2$. We observe that the ring is composed of homogeneous sections (set of nodes that contain groups of the same type) of size 2. In the next ring, we observe that the homogeneous sections are of size 3. Here, the agents in the middle of the sections have $\rho_i = 1$ since both of their neighbors are of the same type. There are 4 agents with $\rho_i = 1$, while the others have $\rho_i = 1/2$. Therefore, we can write $1/2 < p(c) < 1$. Finally, the last ring has 2 sections with size 6 each and the network homophily is the highest, i.e., $p(c) < p(d) < 1$.


DOI: 10.2870/63947

In this example, at least 3 activists must share common knowledge of thresholds in order to be able to coordinate and start the protest since their threshold is $3$. In (a) and (b) the collective action never starts since the *activists* are not close enough to coordinate. As the degree of homophily increases, the size of the homogeneous sections increases, allowing for coordination. Therefore in (c) we observe that the activists choose $r$ at $t = 0$ and the other agents revolt at $t = 1$ after they observe the action of the activists. A further increase of $\rho(G)$ results in larger sections of both types. This does not affect the coordination of activists since they are always able to coordinate above a certain level of network homophily. However, as the section size of the conservatives increases, it takes longer for the information to reach everyone. This results in a slower spread of collective action to the population.

## 3.7 Conclusion

We aim to improve our understanding of how collective action spreads in large and complex networks in which agents use online social networks as communication tools. This chapter is part of a research project that uses an individual based modeling environment that accommodates a number of diffusion models for spreading the infections/ideologies on large social networks. The project involves three parts: a theoretical-analytical part, a theoretical-numerical part, and an empirical part. The first part of the study is partly covered in this chapter. We develop a dynamic game-theoretic model of the "on-set of revolutions" that focuses on the local spread of information in order to study how network structure, knowledge and information-sharing interact in facilitating coordination.

In this chapter, we provide a detailed description of our model and a comparison with some seminal works from the literature (e.g. Chwe, 2000). Our work differs from existing models in three aspects, all of which relate to how information is modeled. First, we relax the assumption that agents have full information about the structure of the network. In our model, agents have local information about the network structure, as mediated by the network itself. Second, we formulate the diffusion of information and common knowledge by relying on the epistemic framework developed within game theory in order to study the role of the network structure. In particular, we find that for a given network, if there exists a set of agents who share common knowledge, then a sub-graph induced by this set must be complete bipartite.

Finally, we introduce dynamics by assuming that people communicate both their thresholds and their past actions. While the static approach involves characterizing the strategic equilibria of the game, based on the information that agents have about the situation, the dynamic approach involves considering how agents' behaviors evolve over time as further information is revealed. The observation of other agents' past actions allows agents to make inferences based on the knowledge of their neighbors' types. This behavior might explain how large-scale movements unfold over time, in spite of a small number of early revolters.

For both the static and dynamic approaches, an important consideration is how agents' types are correlated in the social network. This pertains, in particular, to the degree to which homophily is displayed. Thus, for example, whether low-threshold agents (*activists*) tend to be connected to activists (high homophily) or, to high-threshold individuals (*conservatives)* is crucial to understand the range and speed of social action. So far, we have analyzed simple examples using ring networks. We observe that a higher degree of homophily enables people with high thresholds to coordinate and revolt, at the expense of a slower spread of collective action in the population. This motivates us to study the role of homophily theoretically. We will undertake the study of a theoretical model for simple setups that are either regular (say, lattices) or that exhibit a recurrent structure (e.g. linked stars). This will provide basic insights for the implications of the network structure on common knowledge and homophily.

In addition to the theoretical-analytical part described above, the project will also have computational and empirical levels. The second part still involves a theoretical analysis of the model but one based on mean-field and numerical methods. Social networks do not display the simple structure considered in the first part. Specifically, they display substantial heterogeneity and complex structures. This part concerns the inclusion into the model of more complex and realistic features that are novel, interesting and important but that cannot be solved by analytical methods.[22] Due the additional complexity, we will rely on numerical or mean-field

[22] Those additional features may include the addition of a second dimension like reported type (in contrast with real type) as in Kuran (1989); weights on the links, nodes; stochastic processes, etc.

techniques to explore the model through large-scale and systematic simulations. This part aims at integrating features that may enrich or modify the perspective and insights obtained in the previous analysis. It can therefore be viewed as a robustness test of the more stylized theoretical analysis undertaken in the first part of the project.

For instance, this part will involve generating large complex networks (Poisson, scale-free, etc.) with stochastic patterns of homophily associated with them. Given any family of networks under consideration, we could structure the analysis by considering the effects of two parameters: average degree $z$ and the network homophily $\rho$. It would be interesting to understand how the range and speed of social action depends on these parameters. In line with other studies in the literature (e.g. Watts, 2002) we might expect that the effect of an increase in $z$ would be non-monotonic - increasing at the beginning and decreasing thereafter. It would also be very interesting to understand how this dependence on $z$ interacts with $\rho$, and how all of this depends on the family of networks under consideration (say, Poisson versus scale-free). As mentioned, a study of this problem may be undertaken numerically. The study should also be approachable by mean-field methods analogous to those used by Newman *et* al (2001), Watts (2002), and many others in the epidemic literature.

Finally, the third part of the project is empirical and involves testing our novel modeling approach by using real-life data on social networks. We consider two possible approaches: (i) Simulating the process in artificially created ("synthetic") networks that are constructed to mimic some of the network characteristics (degree distribution, clustering, internode correlations) that are known to arise in specific social networks. This would allow us to predict how information spreads in a network. In general this should depend on the nature of the information involved and the specific mechanism at work (e.g. whether the diffusion unfolds through virtual networks, word-of-mouth communication, etc.) (ii) Focusing on existing social networks on which one can gather large-scale specific data (e.g. Facebook-based networks). In this case, for example, one may have specific applications in mind (e.g. the spread of insurgency in a popular revolt, as in the so-called "Arab Spring" revolts) or the spread in the use of a consumer product (say, the ipad, some internet game, or the purchase of some music).

The idea is to test our model with large-scale data of social networks and some contemporaneous "revolutionary movement" unfolding on it in order to test our model, and to compare its implications with those of other models which conceive the process as one of contagion (e.g. Gonzalez-Bailon *et* al. 2011). In these papers, the only relevant consideration is how many "infected neighbors" an agent has (in absolute or relative terms). However, the role played by the network in the diffusion of mutual and common knowledge is neglected. In this sense, our model seems more suitable than epidemic models for describing social processes that are based on Facebook-type platforms, where there is a bidirectional exchange of information and knowledge.

## 3.8 Appendix

The knowledge operator $\mathbf{K}_i$ satisfies certain axioms, including:

K1: $\mathbf{K}_i(E) \subseteq E$

K2: $\Omega \subseteq \mathbf{K}_i(\Omega)$

K3: $\mathbf{K}_i(\cap_k E_k) = \cap_k \mathbf{K}_i(E_k)$

K4: $\mathbf{K}_i(E) \subseteq \mathbf{K}_i(\mathbf{K}_i(E))$

K5: $-\mathbf{K}_i(E) \subseteq \mathbf{K}_i(-\mathbf{K}_i(E))$

In words, K1 says that if $i$ knows $E$, then $E$ must be the case. K2 says that $i$ knows that some possible world in $\Omega$ occurs no matter which possible world $\omega$ it is. K3 says that $i$ knows a conjunction if and only if $i$ knows each conjunct. K4 is a *reflection axiom*, which says that if $i$ knows $E$, then $i$ knows that she knows $E$. Finally, K5 says that if the agent does *not* know an event, then she knows that she does not know. We can also write that:

K1 and K2: $\mathbf{K}_i(\Omega) = \Omega$

K3: If $E \subseteq F$, then $\mathbf{K}_i(E) \subseteq \mathbf{K}_i(F)$

K1 and K4: $\mathbf{K}_i(E) = \mathbf{K}_i(\mathbf{K}_i(E))$

*Definitions.* Let a set $\Omega$ of possible worlds together with a set of agents $N$ be given.

1. The event that $E$ is (*first order*) *mutual knowledge* for the agents of $N$, $\mathbf{K}_N^1(E)$, is the set defined by

$$\mathbf{K}_N^1(E) \equiv \bigcap_{i \in N} \mathbf{K}_i(E)$$

2. The event that $E$ is $m^{th}$ *order mutual knowledge* among the agents of $N$, $\mathbf{K}_N^m(E)$, is defined recursively as the set

$$\mathbf{K}_N^m(E) \equiv \bigcap_{i \in N} \mathbf{K}_i(\mathbf{K}_N^{m-1}(E))$$

3. The event that $E$ is *common knowledge* among the agents of $N$, $\mathbf{K}_N^*(E)$, is defined as the set

$$\mathbf{K}_N^*(E) \equiv \bigcap_{m=1}^{\infty} \mathbf{K}_N^m(E)$$

It can be shown that:

(1) If $\omega \in \mathbf{K}_N^*(E)$ and $E \subseteq F$, then $\omega \in \mathbf{K}_N^*(F)$.

(2) $\omega \in \mathbf{K}_N^m(E)$ if and only if for all agents $i_1$,$i_2$,...,$i_m \in N$, $\omega \in \mathbf{K}_{i_1}\mathbf{K}_{i_2}..\mathbf{K}_{i_m}(E)$.

Hence, $\omega \in \mathbf{K}_N^*(E)$ if and only if (2) is the case for each $m \geq 1$.

The condition that $\omega \in \mathbf{K}_{i_1}\mathbf{K}_{i_2}...\mathbf{K}_{i_m}(E)$ for all $m \geq 1$ and for all $i_1, i_2, ..., i_m \in N$ is Schiffer's definition of common knowledge, which is the one that is most often used in the literature. Schiffer uses the formal vocabulary of *epistemic logic* and his general approach is to augment a system of sentential logic with a set of knowledge operators corresponding to a set of agents, and to define common knowledge as a hierarchy of events in an augmented system.

## 3.9 References


Aumann, R. (1976). "Agreeing to Disagree." *Annals of Statistics*, 4, 1236-9.

Bonchek, M. S. (1995). "Grassroots in Cyberspace: Recruiting Members on the Internet or Do Computer Networks Facilitate Collective Action? A Transaction Cost Approach." Presented at the 53rd Annual Meeting of the Midwest Political Science Association, Chicago, IL.

Bonchek, M. S. (1997). "From Broadcast to Netcast: the Internet and the Flow of Political Information." Ph.D. Dissertation, Department of Political Science, Harvard University.

Borgatti, .S. P., and Everett, M. G. (1992). "Notions of Position in Social Network Analysis." *Sociological Methodology*, 22, 1-35.

Carnap, R. (1947). *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago, University of Chicago Press.

Centola, D., and Macy, M. "Complex Contagions and the Weakness of Long Ties." *American Journal of Sociology*, 113, 702-34.

Chwe, M. (2000). "Communication and Coordination in Social Networks." *Review of Economic Studies*, 67, 1-16.

Chwe, M. (1999). "Structure and Strategy in Collective Action." *American Journal of Sociology*, 105, 128-156.

Garrett, R. K. (2006). “Protest in an Information Society: A Review of Literature on Social Movements and New ICTs.” *Information, Communication, and Society*, 9 (2), 202-224.

Gonzalez-Bailon, S. *et* al. (2011). “The Dynamics of Protest Recruitment through an Online Network.” *Scientific Reports*, 1,197.

Gould, R. V. (1993). “Collective Action and Network Structure.” *American Sociological Review*, 58 (2), 182-96.

Granovetter, M. (1978). “Thresholds Models of Collective Behavior.” *American Journal of Sociology*, 83, 1420-43.

Kim, H., and Bearman, P. S. (1997). “The Structure and Dynamics of Movement Participation.” *American Sociological Review*, 62, 70-93.

Klandermans, B. (1988). “Union Action and the Free-Rider Dilemma.” *Research in Social Movements, Conflict and Change*, 10, 77–92.

Kuran, T. (1989). “Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution.” *Public Choice*, 61, 41-74.

Kuran, T. (1991). “Now out of Never: The Element of Surprise in the East European Revolution of 1989.” *World Politics*, 44 (1), 7-48.

Lewis, C. I. (1943). “The Modes of Meaning.” *Philosophy and Phenomenological Research*, 4, 236-250.

Lewis, D. (1969). *Convention: A Philosophical Study*. Harvard University Press, Cambridge, Massachusetts.

Marwell, G., and Oliver, P. E. (1993). *Critical Mass in Collective Action*. Cambridge, Cambridge University Press.

McAdam, D., and Paulsen, R. (1993).“Specifying the Relationship between Social Ties and Activism.” *American Journal of Sociology*, 993, 640-667.

Myers, D. J. (1994). “Communication Technology and Social Movements: Contributions of Computer Networks to Activism.” *Social Science Computer Review*, 12, 251-60.

Newman, M. E. J., Strogatz, S. H., and Watts, D. J. (2001). “Random Graphs with Arbitrary Degree Distributions and Their Applications.” *Phys. Rev.*, E 64.

Oliver, P. (1993). “Formal Models of Collective Actions.” *Annual Review of Sociology*, 19, 271–300.

Olson, M. (1965). *The Logic of Collective Action.* Cambridge: Harvard Univ. Press.

Opp, K., and Gern, C. (1993). “Dissident Groups, Personal Networks and Spontaneous Cooperation: the East German Revolution of 1989.” *American Sociological Review,* 58, 659-680.

Schelling, T. (1960). *The Strategy of Conflict*. Harvard University Press, Cambridge, Massachusetts.

Schelling, T. (1969). “Models of Segregation.” *American Economic Review*, 59, 488-93.

Schelling, T. (1971). “Dynamic Models of Segregation.” *Journal of Mathematical Sociology*, 1, 143-86.

Schelling, T. (1978). *Micro-Motives and Macro-Behavior.* New York: W. W. Norton.

Schiffer, S. (1972). *Meaning*. Oxford University Press, Oxford.

Siegel, D. A. (2009). “Social Networks and Collective Action.” *American Journal of Political Science*, 53 (1), 122-138.

Strang, D., and Soule, S. A. (1998). “Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills.” *Annual Review of Sociology*, 24, 265-90.

Vanderschraaf, P., and Sillari, G. (2009) “Common Knowledge.” *The Stanford Encyclopedia of Philosophy*, Spring 2009 Edition, Edward N. Zalta (ed.),
URL = <http://plato.stanford.edu/archives/spr2009/entries/common-knowledge/>.

Watts, D. (2002). “A Simple Model of Global Cascades on Random Networks.” *PNAS,* 99 (9), 5766-5771.

Korkmaz, Gizem (2012), Network Structure Matters: Applications to R&D collaboration, collusion, and online communication networks
European University Institute DOI: 10.2870/63947