DECENTRALIZED INTERACTIONS IN NETWORKS

A Dissertation in
Business Administration & Operations Research

by

Yang Zhang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2013
The dissertation of Yang Zhang was reviewed and approved* by the following:

Gary E. Bolton  
O. P. Jindal Chair Professor of Management Economics, University of Texas at Dallas  
Dissertation co-Advisor  
Co-Chair of Committee

Susan Xu  
Professor of Supply Chain & Information Systems, Penn State  
Dissertation co-Advisor  
Co-Chair of Committee

Elena Katok  
Ashbel Smith Professor of Operations Management, University of Texas at Dallas

Anthony Kwasnica  
Associate Professor of Business Economics, Penn State

Kalyan Chatterjee  
University Distinguished Professor of Economics, Penn State

John E. Tyworth  
John J. Coyle Endowed Professor of Supply Chain Management  
Head of the Department of Supply Chain & Information Systems

*Signatures are on file in the Graduate School
My dissertation concerns decentralized decision making of networked agents. By means of laboratory experiments, Chapter 1 and 2 respectively examine how the network structure affects coordination and free riding, referring to the case where a player’s incentive to take a certain action increases / decreases if more players connected to her takes the same action. Typical applications include technology adoptions under network effects (coordination), and the supply of informational goods in peer-to-peer networks (free riding).

In Chapter 1, we find that both global and local networks impact coordination when agents have incomplete information about the network structure. Simple topological measures (e.g. network density and one’s number of connections) appear sufficient for predicting the level of successful coordination.

In Chapter 2, we find a strong effect of one’s local network on free riding, which is robust regardless of whether agents have full or localized view of the network. The patterns of free riding under both information settings can be explained by a unified behavioral model involving one’s local network connectivity, and one’s tendency of reacting to past neighbor actions and actions of oneself.

In Chapter 3 we study how a network is formed by decentralized decisions of agents. Through a simple analytical model, we explain the emergence of core-periphery landscape in knowledge citation networks as a dynamic process, driven by innovation and citation of individual forward-looking authors who contribute knowledge in their own interests. We show that the core-periphery structure forms on the equilibrium path with an initial state that contains sufficiently large amount of knowledge.
TABLE OF CONTENTS

List of Figures .................................................................................................................. v
List of Tables ..................................................................................................................... vi
Acknowledgements .......................................................................................................... vii

Chapter 1 Network Effects on Decentralized Operations: A Laboratory Investigation .......... 1
  1. Introduction .................................................................................................................. 1
  2. Model, Experimental Design & Hypotheses ............................................................... 44
  3. Results ......................................................................................................................... 11
  4. Conclusions and Discussion ....................................................................................... 22
  References ...................................................................................................................... 24

Chapter 2 Network, Information, and Free Riding ............................................................ 28
  1. Introduction .................................................................................................................. 28
  2. Literature Background .............................................................................................. 29
  3. The Background Model .............................................................................................. 31
  4. Experimental Design and Hypotheses ...................................................................... 36
  5. Results ......................................................................................................................... 41
  6. Conclusions and Discussion ....................................................................................... 51
  References ...................................................................................................................... 53

Chapter 3 Innovation, Citation, and the Emergence of Knowledge Core ......................... 57
  1. Introduction & Background Literature ...................................................................... 57
  2. The Baseline Model ................................................................................................... 59
  3. Equilibrium Analysis ................................................................................................. 61
  4. Conclusions and Discussion ....................................................................................... 65
  References ...................................................................................................................... 66

Appendices ....................................................................................................................... 68
  Appendix A. Background Theory for Chapter 1 ............................................................ 68
  Appendix B. Experimental Implementation for Chapter 1 ............................................. 71
  Appendix C. The Heterogeneity of Coordination in Treatment LrcLden, Chapter 1 ...... 74
  Appendix D. Experimental Instruction for Chapter 1 .................................................... 77
  Appendix E. The Experimental Software Interface for Chapter 1 ............................. 79
  Appendix F. The Derivation of Equilibria for Chapter 2 .............................................. 83
  Appendix G. Experimental Instruction for Chapter 2 .................................................... 87
  Appendix H. The Experimental Software Interface for Chapter 2 ............................. 92
  Appendix I. Proofs of Chapter 3 .................................................................................... 97
LIST OF FIGURES

Figure 1-1. Network structure for a single cohort ................................................................. 9
Figure 1-2. Treatment-level coordination .............................................................................. 12
Figure 1-3. Pooling equilibria in high density treatments .................................................... 13
Figure 1-4. Separating equilibria in low density treatments .................................................. 14
Figure 1-5. Global network effect ......................................................................................... 16
Figure 1-6. Economic return effect ....................................................................................... 18
Figure 1-7. Strategic risk ...................................................................................................... 20
Figure 2-1. Network structures used in experiments .............................................................. 37
Figure 2-2. Treatment-level network effects ......................................................................... 43
Figure 2-3. Local network effect under incomplete information ........................................... 45
Figure 2-4. Global network effect under incomplete information ......................................... 46
Figure 2-5. Information effect: aggregate level ..................................................................... 46
Figure 2-6. Information effect: position-wise ....................................................................... 47
Figure 2-7. Individual behavior: myopia ............................................................................... 49
Figure 2-8. Individual behavior: inertia ................................................................................ 49
Figure 3-1. Exemplary citation networks .............................................................................. 57
Figure 3-2. Example of core-periphery citation networks that our model could generate....... 65
Figure C-1. Coordination trends for each network cohort in Treatment $LrLden$ ..................... 75
Figure E-1. The snapshots for the experimental software interface ....................................... 83
Figure F-1. Network structures used in the experiment ......................................................... 85
Figure H-1. Quiz .................................................................................................................. 93
Figure H-2. Decision making ............................................................................................... 95
Figure H-3. Feedbacks .......................................................................................................... 97
LIST OF TABLES

Table 1-1. Subject’s information structure about the network they play in............................5
Table 1-2. Summary of experimental design........................................................................8
Table 1-3. Equilibrium structure...........................................................................................9
Table 1-4. Notations for our dataset ....................................................................................12
Table 1-5. Local network effect in low density networks....................................................14
Table 1-6. Equilibrium selection: equilibria as null hypotheses to test.........................15
Table 1-7. Global network effect: rank sum tests ...............................................................16
Table 1-8. Cohort-averaged coordination rates for case (iv) of Table 1-7......................17
Table 1-9. Economic return effect in low density networks..............................................19
Table 1-10. Economic return effect in low density networks (continued).......................19
Table 2-1. The base game ..................................................................................................36
Table 2-2. Summary of experimental design......................................................................37
Table 2-3. The neighbor degree distributions for $G_h$ and $G_l$ .........................................38
Table 2-4. Equilibrium structure..........................................................................................40
Table 2-5. Notations for our dataset ....................................................................................42
Table 2-6. Sign rank tests on equilibria as null hypotheses .............................................44
Table 2-7. Sign rank tests on equilibria as null hypotheses (cont.).....................................44
Table 2-8. Information effect: independent-sample Mann Whitney tests.........................48
Table 2-9. Individual behavior: logistic regression .............................................................50
Table C-1. Logit regression for $LrcLden$ data....................................................................75
Table C-2. Wald test for cohort heterogeneity in $LrcLden$ dataset....................................76
Table F-1. The neighbor degree distributions for $G_h$ and $G_l$ ...........................................84
Table F-2. Bayesian Nash equilibria for incomplete information treatments...................85
Table F-3. Symmetric Nash equilibria for complete information treatments.....................87
First and foremost, I express my gratitude to my advisors Professor Gary Bolton and Professor Susan Xu for their continuous guidance and encouragement throughout my doctoral study. In particular I also thank Prof Elena Katok, for her advices in many aspects of my research, coursework, and personal affairs. I thank Prof Chatterjee and Prof Kwasnica for the thoughtful advice on my dissertation. I appreciate the Faculty of Supply Chain & Information Systems of Penn State for supporting me and my career development (especially Prof Bansal, Prof Moritz, and Prof Tyworth). Funding from National Science Foundation under Doctoral Dissertation Grant SES-1123580 is acknowledged and appreciated. This dissertation is dedicated to my mother, Jing Zhang.
Chapter 1

Network Effects on Decentralized Operations: A Laboratory Investigation

1. Introduction and Literature Background

Operations management research into networks frequently assumes that network activities can be optimized by a central authority. While the assumption is appropriate for many applications (e.g. distribution networks and network flow problems: Tempelmeier 2011, Ahuja et al. 1993, etc.), interactions on the network are decentralized in other cases, where actors respond to their own incentives and their decisions generate externalities to the networked society. Applications include technology adoption and diffusion (e.g. Tucker 2008, Ryan and Tucker 2007, Dutta et al. 2007), consumer herding (Debo and Veeraraghavan 2009), collaboration and information sharing across firms (Baum et al. 2010, Cowan et al. 2007, Buhman et al. 2005). In these examples individual actors coordinate their decisions with those of their partners in the network. This paper contributes to the research on decentralized operations in networks that involve a coordination game.1 We conduct an experimental investigation on the decision making of agents coordinating their actions in the network.

For a firm that provides technologies under positive network externalities (e.g. Radio-frequency identification (RFID), or communication products like the iPhone), the importance of user network size in generating profit has been long recognized: Since the classical models of Katz and Shapiro (1992, 1986, 1985), it is assumed that everyone’s adoption raises the value of the technology / product to everyone else, so that larger network size makes the product more valuable to every customer. Firms therefore often engage in enthusiastic competitions in expanding the scale of their user networks. However in this paper, we shall suggest to managers the role of user network structure in determining the outcome of coordinative activities such as technology adoption. To see the intuition, notice that an individual’s benefit from using a product sometimes depends on the adoption decisions from a subset (not all) of users in the network.

---

1 Coordination game is generally regarded as a class of supermodular game, with player actions being strategic complementary. See Cachon and Netessine (2004) for discussion of supermodular games in operations management.
whom she replies on that product to interact with. As a result, what matters to decision making is the network structure, which determines “who interacts with whom”. The concept of network in this paper encompasses that of industrial collaboration networks (when the game is played by firms / organizations), and that of social networks (when the game is played between people holding certain relationships such as friends, kinship, or colleagues). The broad research question that guides our work is: How does the structure of network influence the coordination behavior of agents within the network? In the networks we study, players have incomplete information on the network structure beyond their direct connections. In theory, full-coordination and no-coordination are always pooling equilibria, and there are also separating equilibria depending on the ego-connectedness of players.

Our work is motivated in part by the extant literature on revenue management and product / technology adoption with network effects. Classical models (Katz and Shapiro 1992, 1986, 1985), Riggins et al. 1994) assume a complete network where every user coordinates her product adoption with everyone else. Based on the complete network framework, Lee and Mendelson (2008) investigates the optimal production and pricing strategy of the selling firm under competition. Lee et al. (2006) challenge the complete network setup, and argues the existence of “local network bias” that prevents the shift of technology. Sundararajan (2007) develops a model where each customer only coordinates with those directly networked with her. Galeotti et al. (2010) models the scenario in a more general economic setting. We examine how the agent coordination is affected by the underlying network structure from a behavioral viewpoint. The behavior in our experiment exhibits a clear pattern of selection among the theoretical equilibria. Specifically, we find that network density, individual connectivity, and economic return modulate the amount of strategic risk in coordinating and can hence account for the observed pattern of equilibrium selection. In this spirit, our work helps the manager better understand the demand structure shaped across the user network and plan her revenue management strategy accordingly.

This paper also contributes to the literature on decentralized operations in networks. Lovejoy and Sinha (2010) study the efficiency of communication topologies in incubating innovation. Leider et al. (2009) analyze the influence of social networks on trust and reciprocity. Through multi-agent simulations, Hanaki et al. (2007) investigate how individual cooperative behavior

---

2 For example, suppose a customer decides whether to buy the new generation of iPhone that carries a particular video-chat unit only working with the same type of phone. For the unit to function, the customer has to coordinate his purchase with the people whom she intends to video-chat (e.g. her family, friends, colleagues), but obviously she does not need to coordinate this decision with everyone else.
coevolves with interaction architecture. Baum et al. (2010) and Cowan et al. (2007) investigate the emergence of R&D networks from firm collaboration and knowledge generation. The majority of works in this field lacks comparative statics that forecasts the specific style that equilibria change with network layouts, which are established and tested in our work as experimental hypothesis.

There are also numerous field studies in specific industries that involve network effects: Tucker (2008) on the adoptions of video-messaging units, Ryan and Tucker (2007) dealing with videoconferencing technology diffusion with benefit-cost heterogeneity, Conley and Udry (2007) examining the diffusion of agricultural innovations in Ghana, Ackerberg and Gowrisankaran (2006) analyzing the adoptions of ACH e-payment system, Rysman (2004) addressing the competition in yellow-page market, and Saloner and Shepard (1995) studying the ATM installations of banks. Our approach complements the field empirics in that we provide laboratory observations that strictly control the network effect aside from other noisy factors which can be hard to tackle in the field data.

Also related to our work is the economics literature on coordination. Van Huyck et al (1990) finds that group size plays a critical role in coordination: Larger groups of people find it more difficult to coordinate while small groups find it easier (This result inspires a stream of later works including Weber, 2006). Meanwhile, there is a burgeoning research area in experimental economics that addresses the impact of group structure on coordination. Most of works there, however, only consider bilateral games played between each pair of connected players: Berninghaus et al. (2002) examine the equilibrium selection and evolution of bilateral games as affected by certain interaction structures, based on the models of Ellison (1993) and Berninghaus and Schwalbe (1996). Cassar (2007) extends Berninghaus et al. (2002) by introducing random networks and small-world networks as interaction architectures. An earlier work, Keser et al. (1998) compare bilateral interactions upon a complete network and that on a circle. In contrast to these papers, the multilateral setup of our game entails both one’s local connectivity and global network density as determinants of equilibria, for both of which we find behavioral evidence (Section 3). The nature of networked coordination in our paper is similar to the “consensus” in

---

3 With bilateral games, the player collects payoff from separate games played with each of her connected partners (referred as neighbors) in the network. In our game however, the player reaps profit from one unified game played multilaterally with all her neighbors. While in both cases (bilateral and multilateral games) the player has to take into account all players’ actions simultaneously, our setup allows us to explicitly incorporate network configurations into the equilibrium prediction (see Proposition 1-1 and 2, Section 2).

4 Cassar (2007) examines both coordination games and cooperation games on networks.
Judd et al. (2010), while in our experiment players have incomplete information of the network configuration beyond their own direct connections, which is a common feature in practice and allows us to decompose network effects from one’s local network and global network. Corbae and Duffy (2008) investigate coordination on endogenously formed networks with subjects self-determining whom to coordinate with. We focus on coordination behavior on exogenous networks.

The rest of paper is organized as follows. Section 2 introduces the analytical model and experimental design. In Section 3 we analyze the data. Section 4 concludes and discusses limitation.

2. Model, Experimental Design & Hypotheses

We first lay out the game that we investigate in the experiment. Then we describe the theory behind the game, and then introduce the experimental design and hypotheses.

2.1 The Network Game

We connect a group of subjects into a network. Directly connected subjects are referred to as neighbors. Each subject will play with her neighbors a simultaneous-move coordination game that resembles the typical problem of strategic product adoption: If a player chooses to coordinate, she adopts the new product (cellphone) and receives a benefit of \( R \) from each of her neighbors who also coordinates (purchases the same type of phone). If the player chooses to defect (chooses not to coordinate), she remains with the status quo earning a secure profit of \( C \), independent of her neighbor actions. This payment structure reflects strategic complementarities and positive externalities typical in networks where coordination is the issue. We refer to the ratio \( R/C \) as the economic return. Instructions given to subjects are reprinted in Appendix.

Prior to choosing whether to coordinate, subjects are given (incomplete) information about the structure of the network. They know that all members of the network (including themselves) have either 2 or 3 neighbors; we refer to the number of neighbors a subject has as her degree. Subjects know the distribution that neighbor degree follows. This information setting captures a

---

5 Suppose, for example, a player has precise information of how many friends she herself has, yet has only rough estimates on the number of friends that each of her neighbors has.

6 Throughout the paper, the connections are undirected, i.e. we do not distinguish the direction of connections.
common feature to many social / industrial networks: A player has precise information of her own connectivity, yet has only rough image on the connectivity of her neighbors. We say that a network has higher density (or is denser) if for players of every degree type, the neighbor degree distribution is (stochastically) higher. See (A-2) in Appendix A for a formal definition of network density. The information structure of the game is critical to our purposes and for the convenience of the reader is summarized in Table 1-1.

<table>
<thead>
<tr>
<th>Information</th>
<th>Told to Subjects?</th>
<th>Nature of Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>neighbor degree distribution*</td>
<td>Yes</td>
<td>Public</td>
</tr>
<tr>
<td>economic return*</td>
<td>Yes</td>
<td>Public</td>
</tr>
<tr>
<td>own degree**</td>
<td>Yes</td>
<td>Private</td>
</tr>
<tr>
<td>global network structure</td>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

* informed through written instructions.
** informed via computer interface during the experiment.

2.2 Supporting Theory

The theory is inspired by the general models of Galeotti et al. (2010) and Sundararajan (2007). Our version allows explicit computation of equilibria to facilitate laboratory implementation, and sets forth an equilibrium selection principle. For a full treatment of the background theory see Appendix A.

Denote by \( N_i \) the set of neighbors of player \( i \), and \( k_i \) the degree of player \( i \). \( \bar{k} (\bar{k}) \) is the minimal (maximal) degree value in the network. Let \( x_i \) be player \( i \)'s (binary) decision, \( x_i = 1 \) if the player coordinates, and \( x_i = 0 \) if the player defects. Player \( i \)'s objective is to maximize her expected payoff, by deciding whether to coordinate or to defect:

\[
\max_{x_i = 0, 1} \left\{ E \left( R \sum_{j \in N_i} x_j \right) \right\}.
\]

(1-1)

The game is of incomplete information, with player’s degree as her private type, and neighbor degree distribution as distribution of partner types. In many practical networks, the neighbor degrees are positively correlated (Barabasi and Albert, 1999). In a few of other network models the degrees of one’s neighbors are (at least asymptotically) independent (c.f. Erdos and Renyi, 1960). We focus on the cases where neighbor degrees are either independent or positively correlated (termed as “positive neighbor affiliation” and discussed in Appendix A), which are cases realistically appealing.
We write a pure strategy for player $i$ as a mapping from $i$’s degree to her action: $\sigma_i: k_i \rightarrow x_i$. Throughout the paper the focus is given to symmetric strategies, so we omit the player index and write $\sigma_i \equiv \sigma$. A mixed strategy for a player can be written as a mapping from the player’s degree to a probability distribution over her binary actions: $\sigma: k \rightarrow (p_k, 1 - p_k)$, where $p_k$ is the probability that the degree-$k$ player assigns to coordination (hence $1 - p_k$ the probability of defection). We call an increasing (decreasing) strategy when the coordination probability it assigns increases (decreases) in degree. The solution concept to the game is Bayesian Nash equilibrium, which we refer to as simply equilibrium. We refer to a (degree-)pooling equilibrium as an equilibrium in which all players take the same action regardless of degree. A (degree-)separating equilibrium is a (possibly) mixed strategy equilibrium in which a player’s action differs by her degree. As implied by the nature of coordination problems, there exist multiple equilibria. Proposition 1-1 below characterizes the set of equilibria of our game.

**PROPOSITION 1-1. Equilibrium structure.** a. degree-pooling equilibria. There exists a pooling equilibrium where all players defect. If $C < Rk$ another pooling equilibrium exists with everyone coordinating. b. degree-separating equilibria. There exists separating equilibrium that is increasing in degree. Moreover, such equilibrium involves a threshold $t$, such that a player coordinates if having more than $t$ neighbors, defects if having less than $t$ neighbors, and (only if $t$ is integer) coordinates with probability $p_t \in (0, 1)$ if having exactly $t$ neighbors. When neighbor degrees are independent, the threshold $t$ is solved from the following equation.

$$t = \frac{C}{R \left( 1 - \sum_{k=1}^t G(k) + p_t G(t) \right)}$$

(1-2)

where $G(k)$ is the probability that a single neighbor’s degree equals to $k$.

While a formal proof is left to Appendix A, the intuition behind Proposition 1-1 is simple: If $C < Rk$, coordination is optimal even for the player of the lowest degree type, when all neighbors coordinate. Thus there exists a pooling equilibrium with everyone coordinating. To see the existence of separating equilibrium, suppose a player presumes her neighbors are using increasing strategy. If a player has one more neighbor, then the player’s prior on neighbor degrees will either remain the same, if neighbor degrees are independent, or shifted upward (i.e. The player believes...)

---

7 We restrict our attention to symmetric strategies (and equilibria) because of the *ex ante* symmetry of players.

8 See, among other references, Gibbons (1992) for knowledge on Bayesian Nash equilibrium.
that their neighbors have *ex ante* higher degrees due to the increase of her own degree), if neighbor degrees are positively correlated. In both cases, the expected total neighbor coordination level will increase because 1) the coordination from incumbent neighbors weakly rises due to the *ex ante* higher neighbor degrees and the presumed increasing strategy and 2) the newly added neighbor has a nonnegative probability to coordinate. Since the sum of neighbors’ actions *ex ante* increases, the player in question should raise her probability of coordination, and thus reinforces the presumed increasing strategy and constitutes an equilibrium. In a binary action context the equilibrium breaks down to a threshold structure: While one’s degree increases, her probability of coordination increases from 0 (defection), to some positive value ∈ (0,1) (indifference), then to 1 (coordination). The indifference case would correspond to the degree threshold in equilibrium (i.e. t in Proposition 1-1).

Proposition 1-1 shows that the network game has multiple equilibria. The primary tension behind this multiplicity, as posed by Harsanyi and Selten (1988), pits the payoff gain from joint coordination (‘I gain from coordinating if others coordinate’) against the strategic risk inherent in choosing to coordinate (‘I lose if I coordinate but others do not’). If the payoff gain is pursued by players, we should observe the full coordination outcome; if the concern of strategic risk prevails in decision making, we should observe limited coordination. We shall continue to discuss equilibrium selection in Section 3 based on observations from data and a formal definition of strategic risk.

Proposition 1-1 indicates that the optimality of coordination depends on a player’s local network (degree) and the value of the economic return. As a novelty of our experiment, the equilibrium level of coordination depends on the underlying global network structure: Denser global network allows for more space for payoff gain (against strategic risk). To proceed, define that a (Bayesian Nash) equilibrium payoff-dominates another if the former generates higher *ex ante* social welfare than the latter does. This is consistent with the spirit of Pareto-ranking of Bayesian games (Van Zandt and Vives 2007).

**Proposition 1-2. Comparative statics of equilibria.** If the network density is increased, there exists an equilibrium that weakly payoff-dominates (≥) the current one, with lowered equilibrium threshold and every degree type coordinating with weakly higher (≥) probability.

The proof is found in the appendix and we here outline the intuition. Fixing one’s degree, if the network becomes denser, then a player’s neighbor will *ex ante* have more neighbors. This will lead to a rise of total coordination level in the neighborhood of the player, presuming
coordination of individual neighbor does not decline in the denser network. Then the player in question will be incentivized to increase her coordination probability regardless of her degree, which reinforces the presumption and constitutes an equilibrium. Since this new equilibrium has every degree type coordinating with increasing probability (i.e. equilibrium threshold is lowered), the expected social welfare must rise and thus the new equilibrium payoff-dominates the present one. In the same spirit, we refer to the all-coordination equilibrium as payoff-dominant equilibrium.

2.3 Experimental Design and Hypotheses

The theory in Section 2.2 implies that the equilibrium of the game is affected by (global and local) network characteristics and economic return. Therefore, our experiment manipulates the network density and the economic return associated with the game. Each factor takes on two levels, high or low, in a fully crossed 2×2 factorial design, for a total of four treatments. The design is reported in Table 1-2. In each treatment, 40 subjects are connected in separate networks. In the experiment we use $G_h$ and $G_l$ (shown in Figure 1-1) as network structures respectively supporting high and low density and both exhibiting positive correlation of neighbor degrees. Applying the formal network density definition (A-2) in Appendix A it is straightforward to verify that $G_h$ is indeed denser than $G_l$ (Claim B-1, Appendix B). The proportion of degree-three players, denoted by $q$, is 0.5 in the high density network and 0.25 in the low density network.

<table>
<thead>
<tr>
<th>2×2 factorial design</th>
<th>Network density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low: $G_l$</td>
<td>High: $G_h$</td>
</tr>
<tr>
<td>Economic return</td>
<td></td>
</tr>
<tr>
<td>Low: R/C=80/130</td>
<td>LrcLden</td>
</tr>
<tr>
<td>High: R/C=80/100</td>
<td>HrcLden</td>
</tr>
<tr>
<td></td>
<td>LrcHden</td>
</tr>
<tr>
<td></td>
<td>HrcHden</td>
</tr>
</tbody>
</table>

To obtain statistically independent observations, we divided subjects into separate, anonymous cohorts of 8 subjects each. Each (network) cohort is structured as either $G_h$ or $G_l$. Subjects within each cohort are randomly assigned to the network positions in an equally likely manner. Each treatment has 5 cohorts (5 independent observations), and the resulting 20 network cohorts are labeled as N1, N2…N20 respectively. Each treatment was run in a series of sessions, ranging from 8 to 24 subjects per session.
Recall that an important point is that subjects do not know the structure of the global network they are playing in: They do not know the size or the structure (Figure 1-1) of the cohorts they are organized into. They only know the neighbor degree distribution, which is induced by the random assignment over network positions. For the particular network structures in Figure 1-1, the equilibria are summarized in Table 1-3. Appendix B illustrates the important issues of laboratory implementation: generating neighbor degree distribution from the random assignment of players over network positions, deriving equilibria in Table 1-3, and effectively instructing the neighbor degree distributions to subjects.

![Network Structure](image)

Figure 1-1. Network structure for a single cohort

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Separating Equilibria*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_2$</td>
</tr>
<tr>
<td>$HrcHden$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$HrcLden$</td>
<td>0.5</td>
</tr>
<tr>
<td>$LrcHden$</td>
<td>0.625</td>
</tr>
<tr>
<td>$LrcLden$</td>
<td>0.75</td>
</tr>
</tbody>
</table>

* Pooling equilibria on both coordination and defection exist in each treatment.
Our experiment is aimed at explaining to the manager how network structure affects the coordination of agents being networked. We shall approach this goal in two regards of managerial significance: On the one hand, we examine whether the agents at given network positions behave as the equilibrium predicts. On the other hand, we investigate how the agent behavior changes with the network configuration. In doing so, we combine the theoretical results in Section 2.2 and our experimental setup, and derive the following hypotheses.

**HYPOTHESIS 1-1.** *Equilibrium play.* In each treatment, the actual pattern of coordination is consistent with one of the theoretical equilibria prescribed in Table 1-3.

**HYPOTHESIS 1-2.** *Network effects.*

**a. local network effect.** In a given network (with fixed density), players with more neighbors coordinate with weakly higher (≥) probability.

**b. global network effect.** Players of each degree type coordinate with higher probability in denser networks.

Hypothesis 1-1 poses the basic question whether the actual play of the game can be explained as (or anyway close to) its equilibrium. Hypothesis 1-2 is built upon the intuition of Proposition 1-1 and -2 but differs from the theory in two ways. First, as a behavioral conjecture, it is constructed such that we can test network effects on general coordination situations where an equilibrium is not necessarily in place. Second, when play is at equilibrium, Hypothesis 1-2 serves as an equilibrium selection principle for our game. To see so, notice that global network effect selects equilibria across games (networks): It hypothesizes that higher coordination level for each degree type will be achieved in denser networks, which is a possible (though not necessary) direction of change demonstrated in Proposition 1-2. Corollary 1-1 below serves an example of applying Hypothesis 1-2 to reduce the equilibrium set of our experimental games.

**COROLLARY 1-1.** *Equilibrium selection.* Given the equilibrium set in Table 1-3 and the equilibrium selection principle in Hypothesis 1-2, if separating equilibria emerge in low density networks, then all-coordination pooling equilibria will be played in high density networks.

To follow Corollary 1-1, notice that the separating equilibrium for HrcLden (LrcLden) is $p_2 = 0.5, p_3 = 1$ ($p_2 = 0.75, p_3 = 1$) (see Table 1-3). Since the proposed equilibrium selection

---

*To see that, notice that equilibrium is a prerequisite to the results of Proposition 1-1 and 2, but is no longer required in the statement of Hypothesis 2. In that sense, Hypotheses 1 and 2 are aimed at independent purposes (respectively, equilibrium and network effects) and will be tested independently.*
scheme requires equilibria in denser networks to assign higher coordination probability for each
degree type, the only legitimate equilibria in high density networks are degree-pooling with
coordination probability 1 (refer to Table 1-3). In that spirit, implausible combinations of
equilibria over treatments are ruled out by the proposed equilibrium selection principle (c.f.
Hypothesis 1-2).

In each treatment subjects play the above-described coordination game in a repetition of 20
rounds. Repeating the play allows us to observe more experienced subject behavior. Meanwhile,
it seems unlikely that subjects update their prior on networks in later rounds of the game, given
the limited information they possess. As we shall see, the assumption of non-updated network
priors fits well with the data. The random rotation of subjects over network positions (mentioned
earlier) at each round keeps neighborhoods constantly change and every round is hence treated as
an independent one-shot game. The rotation of neighbors is restrained within each cohort, so as to
maintain the statistical independence across cohorts.

The data was collected from 160 subjects from March through June 2010, at the Laboratory
for Economics Management and Auctions (LEMA) at Penn State University. Subjects are mainly
undergraduate students at Penn State, recruited from an online information system. Cash is the
only motivation offered for subject participation. The software is programmed in zTree
(Fischbacher, 2007). Instruction and software screenshots are found in the Appendix.

3. Results

In testing Hypotheses 1-1 and 1-2, we disentangle the two sources of network effects in the
experimental design: 1) We compare the game plays of different degree types in the same
network (i.e. under the same network density), so as to tease out the effect of local network while
controlling that of global network. 2) We analyze the data from the same degree type but in
different networks, so as to fix the local network while examining the global network effect. In all
cases regarding the network effects, the level of economic return is controlled analogously.

Define coordination rate as the percentage of subjects who coordinate. The notations in
Table 1-4 below apply to our dataset. Labels of treatments (e.g. LrcLden) and cohorts (e.g. N16)
are used as their corresponding indicator variables.
Table 1-4. Notations for our dataset

<table>
<thead>
<tr>
<th>Notation*</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_i$</td>
<td>degree-$i$ ($i \in {2,3}$)</td>
</tr>
<tr>
<td>$X_{di}$</td>
<td>degree type $i$ in treatment $X$ ($i \in {HrcHden, LrcHden, HrcLden, LrcLden}$)</td>
</tr>
<tr>
<td>$action$</td>
<td>= 1 if coordination, = 0 if defection</td>
</tr>
<tr>
<td>$period$</td>
<td>round of the game, taking values of 1,2… 20.</td>
</tr>
<tr>
<td>$X_Y$</td>
<td>the interaction between two variables $X$ and $Y$ (e.g. $period_{LrcLden}$).</td>
</tr>
</tbody>
</table>

* Some of the notations are used in Appendix.

![Figure 1-2. Treatment-level Coordination](image)

Before proceeding to the main body of data analysis, Figure 1-2 above outlines the overall landscape of our dataset. As shown in Figure 1-2, coordination is unambiguously successful when the network has high density: Coordination rates in both high density treatments are higher than 90% in every period, and higher than 95% after period 3. In all practical terms, payoff-dominant equilibria are played in these networks. In contrast, the low density networks achieve a substantially lower rate of coordination, ranging roughly between 50% and 70% each period.

To understand further coordination behavior, we decompose the dataset in various ways, and test for equilibria and the local network effect (Observation 1-1), the global network effect (Observation 1-2), and the economic return effect (Observation 1-3). For the observed effects, we provide a multi-fold explanation based on the concept of strategic risk established by Van Huyck.
et al. (1990), and that of potential gain based on Brandts and Cooper (2006) and Goeree and Holt (2005).

**Observation 1-1. Equilibria & Local network effect.** In high density networks, players with different degree types behave similarly towards coordination, consistent with the pooling equilibrium. In low density networks, high-degree players play significantly higher level of coordination than do low-degree players. At treatment level this behavioral pattern supports separating equilibrium. Overall, the hypotheses on equilibrium play and local network effect (Hypotheses 1-1 and 1-2a) are supported.

To illustrate, Figure 1-3 and -4 plot coordination rates by player degree types, at fixed levels of network density and economic return (i.e. in fixed treatments). As implied by Figure 1-3, the average degree-3 player coordination rates do not appear significantly different than those for degree-2 players when network density is high. The average difference in coordination rate between two degree types is 4% (4.5%) for treatment HrcHden (LrcHden), against a base coordination rate of 95.5% for low-degree players in both treatments. Indeed, the pooling equilibria is favored in high density networks, although an exact statement is weakly rejected in a paired-sample signed rank test with cohort units, yielding p-values of 0.0556 (0.0897) for HrcHden (LrcHden). As such, the hypothesis on equilibrium play (Hypothesis 1-1) is supported in high density networks.

![Figure 1-3. Pooling equilibria in high density treatments](image)

---

10 The test is drawn on paired samples, since each subject generates data for different degree types across rounds of game in a given cohort (due to random positioning after each round --- see Section 2.3).
Strong patterns of degree-separating plays are present in networks with low density. Subjects with more neighbors are more likely to coordinate (Figure 1-4). This is consistent with the conjectured local network effect (Hypothesis 1-2a), and is confirmed by the paired-sample signed-rank tests in Table 1-5 (z-value<0, p-value<0.05).

Table 1-5. Local network effect in low density networks

<table>
<thead>
<tr>
<th>Ho: coordination rate</th>
<th>z-value</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_2 - d_3 = 0$**</td>
<td>-2.023</td>
<td>0.0431</td>
</tr>
<tr>
<td><strong>HrcLden</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LrcLden</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Two-tailed
** The data unit is coordination rate for each degree type for each cohort.

Now we test for specific equilibria. One-sample signed rank tests (Table 1-6) with cohort as unit confirm that the behavioral pattern is consistent with the aggregate levels of coordination anticipated by separating equilibria in Hypothesis 1-1: All p-values are larger than 0.15. The observed equilibria are $p_2 = 0.5, p_3 = 1$ for HrcLden and $p_2 = 0.75, p_3 = 1$ for LrcLden (see Table 1-6 for test details). We should however acknowledge that, as suggested by Shachat (2002), Walker and Wooders (2001), and Erev and Roth (1998), these test results only support equilibria

---

11 Notice that the one-sample signed-rank tests imply that the low degree players in LrcLden play larger coordination rate (0.75) than those in HrcLden (0.5), which may seem conflicting with the information we read from Figure 1-4. This happens because there are a few cohorts in LrcLden in which coordination is played at rather low levels. When a nonparametric test is applied with cohort as unit of analysis, the impact of the coordination poverty of these cohorts has been largely reduced. Thus, the one-sample signed-rank tests suggest a seemingly higher level of coordination in LrcLden than Figure 1-4 does. For more details on the heterogeneity of behavior within LrcLden data see Appendix C.
at aggregate level (i.e. aggregated over players with the same degree type) and are not proof that mixed equilibria are *exactly* executed on the individual level.

Table 1-6. Equilibrium selection: equilibria as null hypotheses to test

<table>
<thead>
<tr>
<th>Ho: coordination rate***</th>
<th>HrcLden</th>
<th>z-value</th>
<th>p-value**</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>d2</td>
<td>0.135</td>
<td>0.8927</td>
</tr>
<tr>
<td>0.97 (1)*</td>
<td>d3</td>
<td>-1.219</td>
<td>0.2228</td>
</tr>
<tr>
<td>Ho: coordination rate LrcLden</td>
<td>z-value</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>d2</td>
<td>-1.214</td>
<td>0.2249</td>
</tr>
<tr>
<td>1</td>
<td>d3</td>
<td>-1.406</td>
<td>0.1599</td>
</tr>
</tbody>
</table>

* In practical terms we use 0.97 instead of 1. The test of coordination rate=100% gives p-value of 0.0422, but perfect coordination here is too strict a null hypothesis to survive.

** Two-tailed.

*** The data unit is coordination rate for each degree type for each cohort.

To summarize the observation, the actual coordination behavior in our experiment can be approximated by theoretical equilibria (on aggregate level). Therefore, Hypothesis 1-1 is supported. Specifically, the observed pattern of behavior falls into the prediction of Corollary 1-1 (i.e. pooling equilibrium in high density networks, while separating equilibria in low density networks), and evidences the local network effect and equilibrium selection (Hypothesis 1-2).

**Observation 1-2. Global network effect.** *Individuals of each degree type coordinate at higher level in high density networks than in low density networks, consistent with the global network effect hypothesis (Hypothesis 1-2b).*
Table 1-7. Global network effect: rank sum tests

<table>
<thead>
<tr>
<th>Rank-sum difference in coordination rates**</th>
<th>z-value</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>case (i): $HrcHden_{d2} - HrcLden_{d2}$</td>
<td>2.402</td>
<td>0.0163</td>
</tr>
<tr>
<td>case (ii): $HrcHden_{d3} - HrcLden_{d3}$</td>
<td>2.660</td>
<td>0.0078</td>
</tr>
<tr>
<td>case (iii): $LrcHden_{d2} - LrcLden_{d2}$</td>
<td>2.102</td>
<td>0.0356</td>
</tr>
<tr>
<td>case (iv): $LrcHden_{d3} - LrcLden_{d3}$</td>
<td>1.491</td>
<td>0.1360</td>
</tr>
</tbody>
</table>

* Two-tailed
** The unit of analysis is coordination rate for each degree type in each cohort.

In all cases present in Figure 1-5, high network density leads to considerably better coordination than does low density. This finding is endorsed by the rank sum tests (Table 1-7) on the coordination difference driven by network density across treatments. In Table 1-7, the signs of test statistics suggest that, other things equal, coordination levels are higher in denser networks. For the case in which the test is not significant (case (iv), $p$-value=0.1360), we break down the corresponding data by cohorts. As shown in Table 1-8, while high degree subjects in all cohorts of treatment $LrcHden$ coordinate perfectly in every round, there are two cohorts (N16, N20) out
of five in \( LrcLden \) in which degree-three subjects play far less than 100% coordination \((0.775, 0.45)\). In this case, high density network obviously dominates the low density one in cohort-wise coordination rates. Altogether, we conclude the global network effect: For any given degree type, the level of player coordination increases with network density --- Hypothesis 1-2b is supported.

Table 1-8. Cohort-averaged coordination rates for case (iv) of Table 1-7

<table>
<thead>
<tr>
<th>cohorts of ( LrcHden )</th>
<th>coordination rate</th>
<th>cohorts of ( LrcLden )</th>
<th>coordination rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>N11</td>
<td>1</td>
<td>N16</td>
<td>0.775</td>
</tr>
<tr>
<td>N12</td>
<td>1</td>
<td>N17</td>
<td>1</td>
</tr>
<tr>
<td>N13</td>
<td>1</td>
<td>N18</td>
<td>1</td>
</tr>
<tr>
<td>N14</td>
<td>1</td>
<td>N19</td>
<td>1</td>
</tr>
<tr>
<td>N15</td>
<td>1</td>
<td>N20</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Integrating Observations 1 and 2, we conclude that theoretical equilibria are acceptable measures of aggregate player coordination. Furthermore, the observed equilibrium style indicates that both local and global network structures have effects on equilibrium selection.

**Observation 1-3 Economic return.** Economic return weakly affects the pattern that coordination evolves in low density networks.
Figure 1-6 separates the data by levels of economic return, while fixing network density and degree type. At time-aggregated level, economic return seems not to influence coordination: Every two cases in Figure 1-6 differing only in economic return hold approximately the same coordination rate averaged over rounds. Nevertheless, coordination changes in distinct directions over time with different economic return levels in the two treatments with low network density (Figure 1-6): When economic return is high, coordination appears to be stable throughout the experiment. Nevertheless, although starting off higher, coordination continuously deteriorates in the treatment with lower economic return. Using nonparametric rank-sum tests, Table 1-9 compares the coordination rates in the low density networks averaged over Round 1-10 of the game and those over Round 11-20. The test outcomes in Table 1-9 agree with the visual message from Figure 1-6 (i.e. $z$-value is negative for rounds 1-10 then positive for rounds 11-20), albeit the detected trends are not strong enough (all $p$-values $>$0.15). Similar tests indicate that, for players with both degree types, lower economic return results in (though statistically weakly) more decrement of coordination level from the 1st to the 2nd half of the game in the low density treatments (see Table 1-10).
Table 1-9. Economic return effect in low density networks

<table>
<thead>
<tr>
<th>Range over which the data is averaged</th>
<th>Rank-sum difference in coordination rates**</th>
<th>z-value</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rounds 1-10</strong></td>
<td>$H_{rcLden_{d2}} - L_{rcLden_{d2}}$</td>
<td>-0.424</td>
<td>0.6714</td>
</tr>
<tr>
<td></td>
<td>$H_{rcLden_{d3}} - L_{rcLden_{d3}}$</td>
<td>-1.410</td>
<td>0.1585</td>
</tr>
<tr>
<td><strong>Rounds 11-20</strong></td>
<td>$H_{rcLden_{d2}} - L_{rcLden_{d2}}$</td>
<td>0.313</td>
<td>0.7540</td>
</tr>
<tr>
<td></td>
<td>$H_{rcLden_{d3}} - L_{rcLden_{d3}}$</td>
<td>0.111</td>
<td>0.9113</td>
</tr>
</tbody>
</table>

* Two-tailed  
** The unit of analysis is coordination rate for each degree type in each cohort, averaged over respective time periods.

Table 1-10. Economic return effect in low density networks (continued)

<table>
<thead>
<tr>
<th>Rank-sum difference in the decrement of coordination rates**</th>
<th>z-value</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{rcLden_{d2}} - L_{rcLden_{d2}}$</td>
<td>-0.731</td>
<td>0.4647</td>
</tr>
<tr>
<td>$H_{rcLden_{d3}} - L_{rcLden_{d3}}$</td>
<td>-1.730</td>
<td>0.0837</td>
</tr>
</tbody>
</table>

* Two-tailed  
** The unit of analysis is the difference between coordination rate averaged over rounds 1-10 and that averaged over rounds 11-20, for each degree type in each cohort.

A more elaborate study on the behavioral trend in the low density treatment with low economic return ($L_{rcLden}$) requires understanding on the heterogeneity at cohort level. This will be explored in the Appendix C.

To summarize the results, we find that subjects’ play of coordination is close to equilibria of the game at aggregate level (Hypothesis 1-1). Agents play payoff-dominant equilibria in high density networks, performing virtually full coordination. Meanwhile, separating equilibria emerge in low density networks, where players with more connections coordinate at higher level. The observed pattern of behavior suggests the effectiveness of global and local network structures as equilibrium selection device (Hypothesis 1-2) on the networked coordination games we study.
Why do we observe such coordination pattern in our game? To find an answer, we define strategic risk $p_i$ as the probability perceived by a generic player that her degree-$i$ neighbor will defect (Note that the player has only stochastic information on neighbor degrees). Strategic risk is an important concept in the theories of equilibrium selection (Harsanyi and Selten 1988, Harsanyi 1995), and is central to understanding the experimental coordination results of Van Huyck et al. (1990). Coordination is robust if its optimality survives high variation on strategic risk -- $(\rho_2, \rho_3)$ plane. Represented by the shaded area in Figure 1-7, the range in which coordination is ex ante more profitable than defection is always larger for high degree than that for low degree in every treatment. Therefore, coordination is more robust for players with more neighbors. In other words, one’s local network determines the robustness of her coordination.

Figure 1-7. Ranges of strategic risk in which coordination is optimal by treatment and degree types

On the global network level, high network density raises the proportion of high-degree players in the system, the ones for whom coordination is most robust. Furthermore, as a secondary effect, coordination becomes optimal for the low-degree players due to perceived strategic risk because, in a dense network, they are more likely to have a neighbor of high degree.

---

Later in this section we provide an alternative explanation of local network effect from the perspective of potential gain of coordination -- a concept driving the laboratory coordination results in Brandts and Cooper (2006) and Goeree and Holt (2005).
This enforces the pooling equilibrium on coordination (Observations 1-1 and 1-2). In contrast to the uniform trend in high density networks, the two low-density treatments are somewhat dissimilar from one another. In treatment HrcLden, coordination is less optimal for the low degree players than that in the high density networks because 1) their low-degree reduces the coordination robustness so that the benefit of coordinating hangs crucially on their neighbor actions, and 2) their neighbors are more likely to be low-degree also. As a result, low degree players turn to defection more often than they do in the high density networks. On the other hand, coordination from high degree players remains robust, attributed to the high economic return of the treatment. This pattern of separating is stabilized as equilibrium. When economic return is also low (LrcLden), even coordination decisions from high degree players become sensitive to neighbor actions, and are deterred by the fact that their neighbors are more probably low degree. Hence in the LrcLden treatment, the action of coordinating is fragile for both types of players. Consequently, behavior exhibits instability and the overall coordination level declines (Observation 1-3).

While the network effects in our experiment are apparent and intricate, the reader may be surprised by the extent that economic factor impacts coordination in our treatments, which appears less remarkable than those demonstrated in the experiments of Brandts and Cooper (2006) and Goeree and Holt (2005), etc. Next we resolve this confusion. On the local network level, a low degree player faces a game that is equivalent to a high degree player having one of her neighbors defecting for sure. Therefore, the high degree type should be more incentivized to coordinate than does the low degree, given the former sees more potential gain from coordinating. Global network, on the other hand, determines the distribution of players for whom coordination generates high or low potential gain. A denser grand network thus has more players who are more incentivized to coordination, which meanwhile strengthens the beliefs of the rest on neighbor coordination. Thus denser network leads to better coordination for each degree type and overall. An analogous argument explains the coordination situation in low density network. In this manner, our experiment also confirms (in an incomplete information context) the classical

---

13 To see, label by 1…k the neighbors of a degree-k player. Then a degree-k player faces a problem \[ \text{max } \left( E \left( \sum_{j=1}^{k} x_j \right), C \right) \] (equation (1-1)). Suppose the degree of the player changes to k+1 and there is one player defecting with probability 1. Then the degree-(k+1) player faces \[ \text{max } \left( R E \left( \sum_{j=1}^{k} x_j + x_{k+1} \right), C \right), \] which is identical to the problem of degree-k player given \( x_{k+1} = 0 \).
laboratory results of Brandts and Cooper (2006) and Goeree and Holt (2005) on the incentive effects on coordination, but using networks as instruments.

4. Conclusions and Discussion

Our laboratory experiments shed light on how the network that determines “who interacts with whom” affects the strategic decision making of firms or individuals being networked. The game we investigate exhibits a coordination nature that captures a wide range of decentralized networked operations, including product adoption, technology diffusion, consumer herding, among others. We find that agent (firm/individual) behavior is well approximated by theoretical equilibria derived from a model with incomplete network information. Specifically, payoff-dominant pooling equilibria are reached in high density networks where coordination is nearly perfect, whereas in sparser networks, levels of coordination are partial and on the aggregate level are consistent with separating equilibria. The pattern of coordination behavior we observe reflects two effects. a) local network effect: In a given global network, agents with more local connections are more likely to coordinate; and b) global network effect: Holding one’s number of local connections fixed, an agent is more likely to coordinate when the global network becomes denser. In our game, local and global network characteristics, together with economic factors, modulate the robustness of coordination, shaping the pattern of equilibrium selection we observe.

Our findings have several implications for Operations Management research on networks. For managers of network-effect products, our results indicate that the size of user network should not be the only concern for inducing higher level of network coordination. In some case, it is more profitable to focus on a user network with a desired structure rather than that of large size. For instance, our experimental result (Observation 1-1, Section 3) suggests that separating equilibrium emerges in the treatment HrcLden (LrcLden), with the coordination rate of degree-2 players being 0.5 (0.75) and that of degree-3s being 1. Recall that the network used in low density treatments ($G_1$) has a size of 8 players, 2 of whom are degree-three and the rest degree-two. Therefore, the total expected number of coordinating players is $2p_3 + 6p_2 = 2 + 6 \times .5 = 5$ per network in HrcLden and 6.5 for that in LrcLden. Now suppose we have a level of network density adequately high to induce all-coordination pooling equilibrium, such as the density of $G_h$ (Section 2.3). For this level of network density, we only need a network size of 5 (7) to achieve (surpass) the same level of total coordination generated by the size-8 networks in HrcLden

---

14 Recall that $p_k$ is the probability that the degree-$k$ player assigns to coordination (Section 2.2).
Therefore, it may be better off for the management to invest in increasing the density of their user network instead of expanding the network size, depending on the relative costs of doing so.

The effects of degree and network density on coordination observed in our study can be perceived as corresponding counterparts to the effects of ego network and large scale network studied in the context of inter-firm collaboration. In that literature, Schilling and Phelps (2007) has a brief summary of the usual constructs of ego and large scale networks: The ego network usually refers to one’s immediate neighborhood (Ahuja 2000, Baum et al. 2000), and is measured by individual centrality (Smith-Doerr et al. 1999) or number of alliances (Shan et al. 1994). The large-scale network, on the other hand, refers to the configuration of overall alliance network in which the firm is located, and can be measured by clustering or reach (Schilling and Phelps, 2007). In our study, we also identify the effects of local and global networks on agents’ decision making, but in a game-theoretical context.

Our findings also have implications for a firm’s revenue management policies. The two types of network effects we discovered pose different considerations for a firm to price its product over the network: On the one hand, our finding that coordination level increases with network density suggests the firm provide discount on product price to users in sparse network to help reduce the difficulty of coordinated product adoption with low network density. On the other hand, we observe coordination increases with the degree of individual user. As a result, the firm may consider price discrimination with regard to network positions, for which the specific form would depend on the problem setup: For example, the optimal pricing scheme may assign lower price to individuals with less number of connections, in order to ease their adoption. Or as a possible alternative, the company may also issue discounts to high degree customers, in order to incentivize their purchases and generate the desired externality. See Candogan et al. (2012) for a related model on price discrimination of network goods that assumes complete network information on consumer side.

\[ \text{(LrcLden)} \]\(^{15}\) We focus on the cases where network density is independent of network size (see the definition of network density as \((A-2)\) in Appendix A). This excludes the non-interesting case of complete network, where neighbor degrees are network size-1 with probability 1, and other cases where the corresponding degree distribution involves support on any degree value violating the upper bound set by the network size. The independence between network density and network size enables our calculations of the expected number of coordinating players in equilibrium.
The theory for our networked coordination game is independent of network size (see Section 2.2). That said, it is certainly an interesting question whether it provides accurate forecasts on actual behavior when the network expands or shrinks in scale. In that line, our experiment can be re-examined under various network sizes. Also, in practice, individuals in social / economic networks might also hold reasonable knowledge outside one's direct relations. Thus another extension of our experiment is to enrich the network information that players have. Third, while the data obtained in our game hues close to equilibrium, there is nevertheless a substantial amount of heterogeneity in play in low density networks (see Appendix C for analyses). This phenomenon is consistent with studies of mixed strategies in other settings (e.g. Shachat, 2002, Walker and Wooders, 2001), where individuals also exhibit significant diversity in strategic play. A more nuanced study of the behavioral dynamics that lead to the equilibria, including heterogeneity in low density networks, requires an assessment of subject learning (ex. Erev and Roth 1998), for which a formal investigation awaits.

References


25


Tempelmeier, Horst. *Inventory Management in Supply Networks*. 2011.
Chapter 2

Network, Information, and Free Riding

1. Introduction

Many economic activities involve decentralized provision of public goods (Olson 1971, Bergstrom, et al. 1986). Classic studies find that group size plays a critical role in cooperating for public goods contribution (c.f. Ledyard 1997, Isaac and Walker 1988, Isaac, et al. 1994, Chamberlin 1974, Hamburger, et al. 1975). There has been much less experimental work, however, on the effect of group structure on free riding / public goods provision. We conduct experimental research on public goods cooperation games in which each player engages with a connected subset of the player population. In other words, the externality of public goods is localized: One can only benefit from the goods contributed by those directly connected to oneself in the society. The game configuration determining “who benefits from whose contribution” defines a network. We ask the following question: How does the network structure, and knowledge about network structure, impact public goods supply in the network?

The spillover of knowledge in R&D networks, and associated literature, serve to illustrate the important notions related to this research. As elaborated in Eeckhout and Jovanovic (2002), Cassiman and Veugelers (2002), a firm’s innovations (e.g. experimenting with new technology, creating new knowledge) may spread to its business partners in the industry, thus creating incentive for the latter to free ride on the outcome of innovation (instead of conducting the same costly innovations by themselves). If we connect firms with business partnership, we have obtained a network. In addressing how the structure of this network affects the incentive for innovation, Bramoullé and Kranton (2007) and Bramoullé, et al. (2011) propose game-theoretical models viewing innovational outcomes as public goods. Galeotti et al (2010) models the problem by assuming that players have incomplete information on network structure. The work pursued here is related to these models.

Theories on network games have received much prosperity in recent years. Many of existing models pursue full information on complex network structures, and thus render highly complicated results (c.f. Goyal 2009). Assuming incomplete information on network structures, Galeotti et al. (2010) and Sundararajan (2007) establish neat comparative statics between
equilibrium and the network properties. Recent laboratory studies (Rosenkranz and Weitzel 2011 and Pla et al. 2009), interestingly, observe that subjects under complete network information behave as if their information is incomplete. In our experiment, we find that a monotonic network effect predicted by the theory under incomplete network information applies to both cases with complete and incomplete network information settings. The observed patterns of free riding under both information settings can fit into a unified behavioral model that involves one’s network connectivity, and one’s tendency of reacting to past actions of neighbors and oneself. Therefore, our experimental findings suggest that theories might not factor in all details of a network, but rather focus on a set of simple measures (density, degree) that yield simple yet most powerful predictions for behavior.

In Section 2, we review the relevant literature. Section 3 introduces the analytical model. The experimental design and empirical hypotheses are established in Section 4. In Section 5 we analyze the data. Section 6 concludes and discusses the limitations.

2. Literature Background


Directly related to our paper is a handful of laboratory studies on cooperation for public goods in networks. Rosenkranz and Weitzel (2011) analyze an experiment with complete network information based on the model of Bramoullé and Kranton (2007). The experiment disentangles the sources of influence on cooperation from global network and local network. The results of

---

16 In this case the network is formed by players’ decisions.
Rosenkranz and Weitzel (2011) suggest that subjects receiving complete network information behave as if they only have partial knowledge of the entire network (although there is no actual experiment with incomplete information reported in Rosenkranz and Weitzel (2011)). Our work, furthermore, directly derive hypotheses from the theory of incomplete information network game (Galeotti et al. 2010), and compare behaviors under both complete and incomplete information to the theoretical hypotheses. Pla et al. (2009) reports a preliminary experiment aiming to relate behavior under limited network information to subjects’ bounded rationality. Unlike ours, their experiment is based on a strategic complementary game, and is conducted between human and automated computers. Assuming full information of network layouts, Cassar (2007) examines both cooperation (in form of prisoner’s dilemma) and coordination games upon various interaction topologies including random networks and small-world networks. While we investigate cooperation for public goods in networks, there is a literature on experimental coordination in networks, e.g. Berninghaus et al (2002), Keser et al. (1998), and Corbae and Duffy (2008)\textsuperscript{17}. See Kosfeld (2007) for a review of experiments on network games.

Our paper is practically motivated by the management science literature on decentralized operations in networks. Baum et al. (2010) and Cowan et al. (2007) investigate the emergence of R&D networks from firm collaboration and knowledge generation. Jaffe and Trajtenberg (2002) and Jaffe et al. (1993) identify that knowledge spillovers by patents are largely restricted within the boundary of countries. The localized nature of knowledge diffusion in their papers suggests the potential impact of networking between information creating entities on knowledge diffusion. In the field of information systems, the problem of free riding in peer-to-peer networks has attracted numerous studies, e.g. Li et al. (2012), Asvanund et al. (2004), Krishnan et al. (2004), Krishnan et al. (2003) and Golle et al. (2001). By means of multi-agent simulations, Hanaki et al. (2007) investigate how individual cooperative behavior coevolves with interaction architecture. Due to the complexity of networking issues studied, the majority of works in this field lacks comparative statics that forecasts the specific style that equilibria change with the network configuration. Such comparative statics are however established and tested in our work as experimental hypothesis.

\textsuperscript{17} Corbae and Duffy (2008) investigate coordination in endogenously formed networks with subjects self-determining whom to coordinate with.
3. The Background Model

We first describe the network and information arrangements, and then introduce the game under different information settings. The theory of incomplete information game is due to Galeotti et al (2010), and we follow their notations.

3.1. Network and Information

The network is constituted by a set of players, $N$, and the connections between players. Connections are binary, so that players are either connected or not. Connected players are neighbors. Let $N_i$ be the set of player $i$’s neighbors or player $i$’s neighborhood, $k_i$ be the number of neighbors that player $i$ has, or the degree of player $i$. In this paper we consider degree as the measure of one’s local network. $\kappa$ is the set of feasible degree values. For convenience, we may ignore the player index when talking about properties of generic player, e.g. $k$ can be the degree of a generic player.

We consider two scenarios regarding player knowledge about the network, complete information and incomplete information. In the complete information setting, players have knowledge of the entire network structure. The rest of this section handles the incomplete information setting, which applies to large social networks where individuals only know about their own connectivity but are uncertain about that of others.

In the incomplete information setting, each player is privately informed of her own degree $k$ (her type), outside her neighborhood she has knowledge of the neighbor degree distribution $G(\cdot | k)$, a joint probability distribution of her neighbor degrees conditional on her own degree. Specifically, let $G(k|k)$ be the probability that the player’s neighbor degrees equal to the $(k$-dimensional) vector $k$, when the player has degree $k$. The underlying network formation procedure that generates the neighbor degree distribution is anonymous and will not be required for deriving equilibrium. We refer to the set of distributions $G := G(\cdot | \cdot)$ as network density, which is a measure of the global network. Define $f_m$ as a non-decreasing mapping $f_m: \kappa^m \rightarrow \mathbb{R}$, for $m \leq k$. Let $f := \{f_m\}_{m \leq k}$. Converting any subset of neighbor degree vector to a scalar, $f$ can be thought as a family of scoring mechanisms: Any collective uplift of (any subset of) neighbor degree values will lead to higher scores. Let $k_m$ be a $m$-dimension subset of neighbor degree distribution.

---

18 Throughout the paper, the connections are undirected, i.e. we do not distinguish the direction of connections.
19 We are aware of other measures on one’s local network (and the global network – introduced later) in sociology (Mizruchi and Marquis 2006). Studying how those measures affect free riding would be certainly interesting but out of the scope of this paper.
degree $k$, obtained by some prescribed selection rule. Let $E_{\mathcal{G}(\cdot|k)}[f_m] = \Sigma_{k \in \mathcal{C}} G(k|k)f_m(k(m))$, which can be understood as an ex ante score of neighborhood of a given degree type $k$, mechanism $f_m$, and the network density $G$. Based on that, we introduce a concept that describes how neighbor degrees are correlated: A network exhibits positive neighbor affiliation (Galeotti et al. 2010), if
\[ E_{\mathcal{G}(\cdot|k)}[f_k] \geq E_{\mathcal{G}(\cdot|k')}[f_k], \] for all $k' > k$, any non-decreasing $f_k$ and any prescribed rule of selecting $k$ out of $k'$ neighbors. The neighbor affiliation is negative when the inequality above is reversed.

Positive (Negative) neighbor affiliation means that higher own degree will shift up (down) the distribution of neighbor degrees (in terms of the one-dimensional score induced by the $f$ operator). The network density $G'$ is said to be higher than another network density $G$, if for any non-decreasing $f_k$
\[ E_{\mathcal{G'}(\cdot|k)}[f_k] > E_{\mathcal{G}(\cdot|k)}[f_k], \forall k. \] That is, if one network allows, for every degree type, higher neighbor degree distribution than does the other, then the former network is denser (Galeotti et al. 2010).

3.2. Payoff

Let $x_i$ be player $i$’s (binary) decision, $x_i = 1$ or 0. We refer to the action 1 as contribute or cooperate, and action 0 as free ride, and the player who contributes (free rides) with probability 1 as contributor (free rider) respectively. Player $i$’s payoff, denoted by $v_i(x_i; x_{N_i})$, is determined by her own decision $x_i$ and actions of her neighbors, $x_{N_i}$. Neighbor decisions are strategic substitutable:
\[ v_i(1; x) - v_i(0; x) \leq v_i(1; x') - v_i(0; x'), \] for any vector of contribution levels $x \succeq x'$. The payoff setting captures the nature of public goods: When neighbor’s contribution levels are higher, the agent in question has less incentive to contribute. For example, it applies to the issue of firm innovation discussed in Section 1: The firm’s incentive to innovate is weakened in a context of high neighbor innovation, because the

---

20 Namely, the selection rule associated with $k(m)$ should select $m$ values out of the total $k$ values present in the vector $k$. While the selection rule can be arbitrary, the idea of establishing a notion of neighbor degree correlation is to make it insensitive to the underlying selection rules (see the definition of positive neighbor affiliation in (2-1)).
firm can free ride the innovational outcomes of its partners. Throughout we assume the game satisfies the following properties.

**Assumption 2-1.** $v_i(x_i; (x, 0)) = v_i(x_i; x) \forall i \in N$. \hspace{1cm} (2-4)

**Assumption 2-2.** $v_i(1; 1) - v_i(0; 1) < 0$, $v_i(1; 0) - v_i(0; 0) > 0 \forall i \in N$ \hspace{1cm} (2-5)

Assumption 2-1 indicates that having a neighbor who is a free rider is payoff-equivalent to not having that neighbor. As discussed in (Galeotti, et al. 2010), Assumption 2-1 is satisfied in many economic models including those where one’s profit depends on the sum of neighbor actions. Assumption 2-2 sets boundary conditions for the game: For any player if no neighbor free rides (contributes), then the player in question should optimally choose to free ride (contribute).

Assumption 2-2 and (2-3) define a public goods game of threshold structure, similar to that of Cadsby and Maynes (1999):

**Remark 2-1.** **Threshold public goods game.** There exists a threshold $X_{N_i}$ in neighbor contribution vector such that player $i$ has the following best response: If $X_{N_i} < X_{N_i}$, then $x_i^* = 1$; if $X_{N_i} > X_{N_i}$, then $x_i^* = 0$. Equality may appear on either side or appears with the case where the player is indifferent between two actions.

Without loss of generality, we can label a player by her degree type. Therefore, we sometimes write payoff function for player $i$ as $v_{k_i}(x_i; x_{N_i})$.

### 3.3. Equilibrium

#### 3.3.1. The complete information game

In the complete information setting, we concentrate on symmetric Nash equilibria in which players at symmetric network positions contribute with the same probability. We label network positions by English letters: $a, b, \ldots$ Let $p_i$ be the probability of contribution of position $i, \forall i = a, b, \ldots$.

Under complete information, the complexity of network structure is fully taken into account of player strategy. Any (local) change of network structure will therefore broadcast externalities across the whole system due to the interdependency of agent’s network positions. Thus it is generally hard to obtain systematic predictions on how equilibria change with network characteristics. Under complete network information, Bramoulle et al. (2011) and Bramoulle and Kranton (2007) have established network-embedded equilibrium properties for continuous-action
strategic substitute games with linear best reply. Their theories, however, cannot apply here because our game features binary action and allows for mixed equilibrium. As we shall see next, tractable relationships exist between equilibrium structure and network configuration, when the game takes place under incomplete information.

3.3.2. The incomplete information game

In the incomplete information setting, the focus is given to symmetric strategy \( \sigma \) mapping from player’s degree to her probability distribution over actions: \( \sigma : k \rightarrow (p_k, 1 - p_k) \), where \( p_k \) is the probability that the degree-\( k \) player assigns to contribution (hence \( 1 - p_k \) the probability of free riding).\(^{21}\) A strategy is non-increasing if \( p_k \) weakly decreases (\( \leq \)) with \( k \). Let

\[
U(x_i, \sigma, k_i) = \int_{x_{N_i}} v_{k_i}(x_i; x_{N_i}) d\phi(x_{N_i}, \sigma, k_i),
\]

where \( \phi(\cdot, \sigma, k_i) \) is the probability distribution over neighbor actions induced by the neighbor degree distribution \( G(\cdot | k_i) \) and the strategy \( \sigma \). The payoff satisfies degree substitution (Galeotti et al. 2010) if for non-increasing \( \sigma \),

\[
U(1, \sigma, k_i) - U(0, \sigma, k_i) \leq U(1, \sigma, k'_i) - U(0, \sigma, k'_i), \forall k_i > k'_i.
\]

The degree substitution means that a higher degree type will have less incentive to contribute under non-increasing strategy. The solution concept to the incomplete information game is Bayesian Nash equilibrium (Gibbons 1992, Osborne and Rubinstein 1998, etc.). Whenever without confusion, we do not literally distinguish solution concepts to different games and simply refer them as equilibrium. We refer to a degree-pooling equilibrium as an equilibrium in which all players take the same (probabilistic) action regardless of degree. A degree-separating equilibrium is an equilibrium in which a player’s (probabilistic) action differs by her degree.

**Proposition 2-1. Equilibrium & Local network effect under incomplete information.**

There does not exist degree-pooling equilibrium in pure strategy. There exists a degree-separating equilibrium where a player contributes if having less than \( t \) neighbors, free rides if having more than \( t \) neighbors, and (if \( t \) is integer) contributes with probability \( p_t \in (0,1) \) if having exactly \( t \) neighbors. If the network displays negative neighbor affiliation, then the degree-separating equilibrium is unique.

\(^{21}\) Given our notations, \( p_1, p_2, \ldots \) are the cooperation probabilities of players of degree-1, -2, \ldots under incomplete information, while \( p_a, p_b, \ldots \) represent the cooperation probabilities of players at position-\( a, -b, \ldots \) under complete information.
While the formal proof can be adapted from that of Proposition 4 in Galeotti et al. (2010), we provide the intuition here. The nonexistence of degree pooling equilibrium is a consequence of Assumption 2-2. To see the intuition for degree separating equilibria, suppose that a player presumes her neighbors are using non-increasing strategy. If the player has one more neighbor, then the player’s prior on neighbor degrees will shift downward, given that neighbor degrees are negatively correlated. The expected total neighbor contribution level will increase because 1) the contribution from incumbent neighbors weakly rises under the presumed non-increasing strategy and 2) the newly added neighbor has a nonnegative probability to contribute (If the new neighbor free rides for sure then it is equivalent to not having her as neighbor due to Assumption 2-1, in which case 2) is simply ignored). Thus, the player in question should decrease her probability of contribution (regardless of degree type), and thus reinforces the presumed non-increasing strategy and potentially constitutes an equilibrium. In a binary action context the equilibrium breaks down to a threshold structure: While one’s degree increases, her probability of contribution decreases from 1 (contributor), to some positive value ∈ (0,1) (indifference), then to 0 (free rider). The indifference case would correspond to the equilibrium threshold (i.e. t in Proposition 2-1).

Proposition 2-1 identifies the effect of one’s local network on public goods contribution. To see, notice that the proposition investigates how the equilibrium contribution tendency changes with degree in a given global network (with fixed network density). The next proposition will provide directional insights into the effect of global network.

**Proposition 2-2. Global network effect under incomplete information.** For degree-separating equilibria: If the network density is increased, then for any non-increasing equilibrium, there exists an equilibrium with weakly higher (≥) equilibrium threshold so that every degree type contributes with weakly higher probability.

We provide the intuition for Proposition 2-2 from a contradiction. Start with the hypothesis that the equilibrium threshold is lowered in the denser network. Notice for every degree type, if the network becomes denser, then a player’s neighbor will *ex ante* have more neighbors. Then the player’s neighbors should free ride more (given more neighbors they have and the hypothesized lowered equilibrium threshold in the denser network). Then the degree threshold (above which the player should free ride) of the player in question rises (regardless of player’s degree type). That leads to contradiction with the hypothesis. Therefore, the threshold of the new equilibrium in
the denser network should be weakly higher.\textsuperscript{22} In a binary action setting, this implies every degree type contributes with weakly higher probability. A formal proof is due to that of Proposition 5 of Galeotti et al (2010).

4. Experimental Design and Hypotheses

In our experiment neighbors play a public goods game as shown in Table 2-1, referred as the \textit{base game}. It is easy to verify that the game satisfies Assumptions 2-1 and 2-2, and exhibits strategic substitutability (2-3) and the threshold public goods property (Remark 2-1).

Table 2-1. The base game

<table>
<thead>
<tr>
<th>Your Profit</th>
<th>Number of Your Neighbors who Contribute</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td><strong>Your Choice</strong></td>
<td><strong>Contribute</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Free ride</strong></td>
</tr>
</tbody>
</table>

Our experiment manipulates the \textit{network structure} and the \textit{network information} associated with the game, so that each treatment \textit{per se} accommodates a distinct set of equilibria. Each factor takes on two levels in a fully crossed 2×2 factorial design (high and low network density, complete and incomplete network information), yielding in total 4 \textit{treatments} as recorded in Table 2-2 below. Throughout the experiment we fix the game played in all treatments (as in Table 2-1). The alias for each treatment is a combination of the underlined fragments of the denotation of treatment variable levels (ex. $Lden_I$ assembles $Lden$ representing low network density, and $I$ abbreviated for incomplete network information).

\textsuperscript{22} It is worth noted that the alternative reasoning, starting with the assumption of increasing equilibrium threshold in denser network, \textit{cannot} make a contradiction. To see, it cannot be predicted whether a player’s neighbors should be contributing more or less, given they have more neighbors (in the denser network) while implementing higher equilibrium threshold. Thus the direction of change in the degree threshold used by the player in question is not clear. That is why the proof is made by assuming lowered equilibrium threshold in denser network and then creating the contradiction.
Table 2-2. Summary of experimental design

<table>
<thead>
<tr>
<th>factorial design</th>
<th>Network Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low density: $G_l$</td>
</tr>
<tr>
<td>Network Information</td>
<td>Lden_I</td>
</tr>
<tr>
<td></td>
<td>High density: $G_h$</td>
</tr>
<tr>
<td></td>
<td>Hden_I</td>
</tr>
</tbody>
</table>

|                  | Complete network information |
|                  | Lden_C                    |
|                  | Hden_C                    |

Figure 2-1. Network structures used in experiments

In all treatments, subjects are connected according to the network structures in Figure 2-1. The structures $G_h$ and $G_l$ respectively support high and low network density (see Remark 2-2 later), with the degree of a player being either 2 or 3. An alternative approach would be to use random graphs like those developed by Bollobas (1985) and Erdos and Renyi (1960). We chose our current network design for its structural conciseness (The networks have only 2 degree types and 4 pairs of symmetric positions), which makes it intuitively easy for subjects to understand.

23 Each network is a statistically independent observation, hereafter referred as cohort. We form 5 cohorts for each treatment, and the resulting
20 cohorts labeled as \(N1, N2...N20\) respectively. Within each cohort subjects are randomly assigned over the 8 network positions (of either \(G_h\) or \(G_l\) depending on the treatment) in an equally likely manner at each game round. In the complete information treatments subjects are provided with pictures of the networks (same as shown in Figure 2-1) in the instruction as well as through the computer interface. In treatments with incomplete information, the network structures are hidden from subjects. Instead, each of the subjects is provided with private knowledge of her own degree and common knowledge of neighbor degree distribution (conditional on one’s own degree) as specified in Table 2-3. The set of neighbor degree distributions is induced by the random assignment of subjects over network positions. Subjects’ information structures are respectively consistent with the complete and incomplete information settings as described in Section 3.1.

Table 2-3. The neighbor degree distributions for \(G_h\) and \(G_l\)

<table>
<thead>
<tr>
<th>probability</th>
<th>(G_h)</th>
<th>(G_l)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. degree-3 neighbors</td>
<td>No. degree-2 neighbors</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1/2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

If the player is degree-2, probability: \(\frac{1}{2}\) for \(G_h\) and \(\frac{2}{3}\) for \(G_l\).

<table>
<thead>
<tr>
<th>probability</th>
<th>(G_h)</th>
<th>(G_l)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. degree-3 neighbors</td>
<td>No. degree-2 neighbors</td>
</tr>
<tr>
<td>1/2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1/2</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

If the player is degree-3, probability: \(\frac{1}{2}\) for \(G_h\) and \(\frac{1}{3}\) for \(G_l\).

**Remark 2-2.** \(G_h\) is denser than \(G_l\).

*Proof.* Remark 2-2 is obviously true when network information complete. Under incomplete information, the neighbor degree distributions conditional on one’s own degree, \(G_h(\cdot | k)\) and \(G_l(\cdot | k)\), are shown in Table 2-3. Therefore we have

\[
E_{G_h(\cdot |2)}[f_2] = \frac{1}{2} f_2(2,3) + \frac{1}{2} f_2(3,3) > \frac{2}{3} f_2(2,3) + \frac{1}{3} f_2(3,3) > \frac{2}{3} f_2(2,3) + \frac{1}{3} f_2(2,2) = E_{G_l(\cdot |2)}[f_2],
\]

As stated later, subjects play the base game for 20 rounds. They are randomized over network positions each round. Since the randomization is restrained within each cohort, the statistical independence across cohorts is maintained.
for any $f_k$ ($k = 2, 3$) non-decreasing in its argument. Applying the definition of network density (2-2) completes the proof. Q.E.D.

The experiment is constituted by a repetition of the base game for 20 rounds. (As stated earlier) In each round we randomize each cohort of 8 subjects over the network positions, so that partnerships constantly change and every round resembles an independent one-shot game. Repeating the play allows us to observe more experienced subject behavior. Meanwhile, it seems unlikely that subjects in the incomplete information treatments update their prior on networks in later rounds of the game, given the limited information they possess.

In the instruction to subjects, we labeled contribution as option A, and free ride as option B. Per the base game in Table 2-1, we explicitly told subjects that their optimal choice is to choose A (B) when there are less than (more than or equal to) 2 neighbors that execute the option A. These acts eliminate the framing effect regarding the description of choices, and rule out the noise in play attributed to not understanding the base game. After each round, players receive information on neighbors’ decisions (while anonymity is maintained for all players). Sample instructions are found in the Appendix.

In total, the data was collected from 160 subjects from February through July 2012, at the Laboratory for Economics Management and Auctions (LEMA) at Penn State University. Subjects are mainly undergraduate students at Penn State, recruited from an online information system. Cash is the only motivation offered for subject participation. The software is programmed in zTree (Fischbacher 2007). Sample software screenshots are found in the Appendix.

All that said, our first empirical question is: Can the actual play of this networked public goods game be explained as (or anyway close to) its equilibrium? This leads to the following hypothesis.

**Hypothesis 2-1.** Equilibrium play. In each treatment, the actual pattern of contribution is consistent with one of the theoretical equilibria prescribed in Table 2-4 below.
Table 2-4. Equilibrium structure

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Equilibria</th>
<th>( p_2 )</th>
<th>( p_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Hden_I )</td>
<td>1</td>
<td>0.094</td>
<td></td>
</tr>
<tr>
<td>( Lden_I )</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>Equilibria</td>
<td>( p_a, p_h )</td>
<td>( p_b, p_g )</td>
</tr>
<tr>
<td>( Hden_C )</td>
<td>0.6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Treatment</td>
<td>Equilibria</td>
<td>( p_a, p_f )</td>
<td>( p_b, p_e )</td>
</tr>
<tr>
<td>( Lden_C )</td>
<td>( \frac{4 - q}{1 + 4q} )</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: For deriving these equilibria see the Appendix.

**HYPOTHESIS 2-2. Network effects under incomplete information.**

**a. local network effect.** In a given network (with fixed density), players with more neighbors contribute with less probability.

**b. global network effect.** Players of each degree type contribute with higher probability in denser networks.

Hypothesis 2-2 is built on the theoretical results of Proposition 2-1 and -2 in Section 3, but differs from the theory in that it is constructed upon general, not-necessarily-equilibrium situations. Therefore, Hypothesis 2-1 and -2 are aimed at independent purposes (i.e. without any joint condition like equilibrium) and can be tested independently.

The background theory and our experimental setup manipulate two information scenarios. Does the network information make a difference in the actual behavior? In fact, recent experiments of Rosenkranz and Weitzel (2011) suggest that higher degree players perform more free riding under complete network information, coherent with the theoretical prediction with

---

25 To see that, notice that equilibrium is a prerequisite to the propositions, but it is no longer required in the statement of hypotheses other than Hypothesis 2-1.
incomplete network information (Hypothesis 2-2a). This suggests some potential analogy in the game play across different informational settings. To explore this issue, we introduce the following hypothesis.

**HYPOTHESIS 2-3. Information non-effect.** Subjects under complete network information behave similarly (on aggregate level) as they do under incomplete network information (as defined in Section 3.1).

Theories of network games are often remarkably complicated under full network information, since the complexity of network is supposed to be fully taken into account of the player strategy. With the assumption of incomplete network information, however, one can establish monotonic network effects on equilibrium (Proposition 2-1 and 2-2). If Hypothesis 2-3 is true, the simple network effects implied by Proposition 2-1 and 2-2 shall prevail in the actual decision making, even when the decision makers are exposed to full network information. This would suggest a way of seeking for simple behavioral patterns in complex networks.

5. Results

In this section, we report our findings from the data. A benchmark of system performance is efficiency – the maximal social welfare that is achievable by the system. In our experimental network games, all-contribution is the Pareto-optimal outcome (though not an equilibrium) in both informational settings.

**REMARK 2-3.** The efficient outcomes for the incomplete information games in Hden_I and Lden_I are that both degree types contribute with probability 1. The efficient outcomes for complete information games in Hden_C and Lden_C are that every network position contributes with probability 1.

The proof of Remark 3 is straightforward and omitted. As a consequence of Remark 2-3, we can conveniently measure the system performance by a single-dimensional metric, contribution rate, defined as the percentage of subjects who contribute in a given setting (e.g. a treatment).

The notations in Table 2-5 below apply to our dataset. Labels of treatments (e.g. Lden_I) and cohorts (e.g. N16) are used as their corresponding indicator variables. In order to support the dataset decomposed in multiple ways, we flexibly combine the notations to designate any
particular subset of data: For instance, \( L_{deg-I} \) denotes the data generated by low degree players \((L_{deg})\) under incomplete network information \((I)\).

Table 2-5. Notations for our dataset

<table>
<thead>
<tr>
<th>Notation</th>
<th>Interpretation</th>
<th>Notation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I, C )</td>
<td>Incomplete network information ((I)), Complete network information ((C))</td>
<td>( H_{den}, L_{den} )</td>
<td>high network density ((H_{den})), low network density ((L_{den}))</td>
</tr>
<tr>
<td>( X_i )</td>
<td>the dataset concerning degree type ( i ) in treatment ( X (i \in {2, 3}, X \in {L_{den-I}, H_{den-I}, L_{den-C}, H_{den-C}}) ), e.g. ( L_{den-I} )</td>
<td>( I_{info} )</td>
<td>( = 1 ) if network information is complete, ( = 0 ) if network information is incomplete.</td>
</tr>
<tr>
<td>( X_j )</td>
<td>the dataset concerning position ( j ) in treatment ( X (j \in {a, b, \ldots h}, X \in {L_{den-I}, H_{den-I}, L_{den-C}, H_{den-C}}) ), e.g. ( L_{den-C} )</td>
<td>( I_{degree} )</td>
<td>( = 1 ) if degree ( = 3 ) (high), ( = 0 ) if degree ( = 2 ) (low).</td>
</tr>
<tr>
<td>( action )</td>
<td>( = 1 ) if contribute, ( = 0 ) if free ride</td>
<td>( H_{deg} (d_3), L_{deg} (d_2) )</td>
<td>high degree ((H_{deg}, d_3)), low degree ((L_{deg}, d_2))</td>
</tr>
<tr>
<td>( period )</td>
<td>round of the game, taking values of ( 1, 2 \ldots 20 ).</td>
<td>( X_Y )</td>
<td>the interaction between two variables ( X ) and ( Y ) (e.g. ( I_{degree-L_{den-C}} )).</td>
</tr>
</tbody>
</table>

5.1. Network Effects at Treatment Level

**Observation 5.1.** **Network effects at treatment level.** The network effects on public goods contribution and welfare are not significant on treatment level in both information settings.

![Network effect at incomplete information](image1.png) ![Network effect at complete information](image2.png)

Figure 2-2. Treatment-level network effects
Contribution to the public goods does not significantly differ by the underlying network structures on treatment level (Figure 2-2). The corresponding two-sample Mann Whitney test (with cohort as unit) yields a p-value of 0.1161 for incomplete information setting and 0.1732 for complete information setting. As we shall see later, the effects of network become salient when we break down the data.

Recall that contribution rate reflects the level of social welfare (Remark 2-3). Therefore, our results also imply that social welfare is not significantly influenced by the underlying network layout. To be precise, define efficiency rate as the ratio of the actual welfare of a given cohort to its theoretical maximum. We conduct rank sum tests on the (cohort-wise) efficiency rates between high density and low density networks, which yield p-values >0.75 for both comparisons involved.

Treatment-level behavior is informative, yet not adequate to explain the rationale behind observations. As implied by Hypothesis 2-2, we disentangle the two sources of network effects under incomplete information: 1) We compare the game plays of different degree types within the same network, so as to separate out the effect of local network while controlling that of global network. 2) We analyze the data from the same degree type but in different networks, so as to control the local network effect while varying the global network. 26

5.2. Equilibrium

Observation 5.2. In none of the cases the exact pattern of free riding is consistent with the equilibrium prediction.

26 In the network notations introduced in Section 3.1, this paragraph means that the local network effect is tested by fixing $G$ while changing $k$, and the global network effect tested by fixing $k$ and altering $G$. 

43
Table 2-6. Sign rank tests on equilibria as null hypotheses

<table>
<thead>
<tr>
<th>Ho: contribution rate**</th>
<th>Hden_I</th>
<th>z-value</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>d2</td>
<td>-2.032</td>
<td>0.0422</td>
</tr>
<tr>
<td>0.094</td>
<td>d3</td>
<td>2.060</td>
<td>0.0394</td>
</tr>
</tbody>
</table>

Ho: contribution rate

<table>
<thead>
<tr>
<th>Ho: contribution rate</th>
<th>Lden_I</th>
<th>z-value</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>d2</td>
<td>-2.032</td>
<td>0.0422</td>
</tr>
<tr>
<td>0</td>
<td>d3</td>
<td>2.023</td>
<td>0.0431</td>
</tr>
</tbody>
</table>

* Two-tailed.
**The data unit is the contribution rate of each degree type in each cohort.

Table 2-7. Sign rank tests on equilibria as null hypotheses (cont.)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Ho: contribution rate</th>
<th>symmetric positions, $G_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>d</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Ho: contribution rate</th>
<th>symmetric positions, $G_l$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>g</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Ho: contribution rate</th>
<th>symmetric positions, $G_l$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\frac{4 - q}{1 + 4q}$</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>$q \in (\frac{3}{5}, 1)$</td>
<td>g</td>
</tr>
</tbody>
</table>

* Ho rejected at 5% level.
**Ho rejected because it violates the observation that contribution level decreases with degree under complete information (Later see Section 5.6).

The data unit of all tests in Table 2-7 is the contribution rate of each network position in each cohort.
As seen in Table 2-6 and Table 2-7, all the equilibria taken as null hypotheses are rejected by the data. Therefore, we conclude that theoretical equilibria are not exact predictors of the outcomes of our games, under both complete and incomplete network information. Next, we shall test the theoretical predictions on network effects decomposed to global and local network levels.

5.3. Incomplete Network Information: Local Network Effect

Observation 5.3. Under incomplete network information, the contribution level significantly decreases with player degree.

In both treatments of incomplete information we observe strong separation of actions by player degree types. Specifically, subjects with more neighbors tend to free ride more on their neighbors’ contribution (Figure 2-3). Dependent-sample sign rank tests 27 based on cohort-averaged contribution rates for both treatments produce p-values of 0.043 and z-values of 2.023. Thus the tests reject the null hypotheses of equal contribution levels across degree types and furthermore indicate that low degree type contributes with higher rate than does the high type. Therefore, our results endorse Hypothesis 2-2a.

For the case of complete network information, we cannot obtain in principle a decomposition of local network and global network that predicts the equilibrium behavior. Therefore, the analyses of local network effect as above and global network effect as in the next section do not apply to the complete network information case. Nevertheless on individual level, as we shall see in Section 5.6, player’s degree is still a significant factor in shaping the subject’s free riding behavior under complete network information.

27 The tests are drawn on dependent samples, since each subject generates data for different degree types throughout the game (due to subject’s random re-positioning within a given cohort after each round --- see Section 4). The tested term is the contribution rate of degree-2 players minus that of degree-3 players.
5.4. Incomplete Network Information: Global Network Effect

**Observation 5.4. Global Network Effect.** Under incomplete network information, the network density has no significant effect on player contribution.

![Figure 2-4. Global network effect under incomplete information](image1)

As shown in Figure 2-4, little evidence is found to support the comparative statics on global networks under incomplete information. Mann Whitney tests with cohort units on contribution rates of players located in networks with different density but having the same degree type favor the null hypotheses that contribution rates do not change with network density (p-values > 0.45 for both comparisons involved).

5.5. Information Effect

**Observation 5.5. Information Effect.** The information effect is not significant except for some high degree players in high density networks, where the contribution rate is higher when network information is incomplete than that under complete information.

![Figure 2-5. Information Effect: Aggregate level](image2)

On treatment level, enriching network information makes little change in aggregate-level public goods contribution (Figure 2-5). In fact, independent-sample rank sum tests on incomplete
vs. complete information treatments yield p-value of 0.3457 (0.2933) for high (low) density networks, both suggesting no statistically significant difference in contribution level across information settings, supporting Hypothesis 2-3. As we will see, the lack of information effect extends when we break down the data.

Figure 2-6. Information Effect: Position-wise

Figure 2-6 compares the contribution level of each network position in distinct information settings. In low density networks, low degree players contribute visibly more in the incomplete information scenario (though not statistically significant per the test in Table 2-8), while information does not make a difference for high degree players\textsuperscript{28}. The network information seems to have slightly opposite influences for different degree types in high density networks, of which the high degree’s contribution decision is statistically significantly impacted: Positions c and f contribute significantly less with complete information than they do under incomplete information. Therefore in this case, providing incomplete network information may help sustaining a high contribution level to the public goods. Nevertheless, the information of network structure in general does not seem to affect decision making in a strong and consistent manner.

\textsuperscript{28} For the low density treatments, there is little noticeable difference in contribution rate for the high degree positions (Figure 2-6, Table 2-8). To see why, notice in case of low network density, the information that high degree subjects receive about neighbor degrees does not differ by information settings: A degree-three subject knows that two of her neighbors are degree-2, and the third neighbor is degree-3 (c.f. Table 2-3) regardless of informational setups. Therefore, the cooperation rate of high degree players in low density networks should not differ by information settings, if player’s decision is primarily affected by their neighbor degrees.
Table 2-8. Information effect: independent-sample Mann Whitney tests

<table>
<thead>
<tr>
<th>density</th>
<th>position</th>
<th>z-value</th>
<th>p-value</th>
<th>density</th>
<th>position</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>a</td>
<td>-0.424</td>
<td>0.6714</td>
<td>a</td>
<td>1.616</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>0.973</td>
<td>0.3305</td>
<td>b</td>
<td>-1.719</td>
<td>0.0857</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>0.629</td>
<td>0.5296</td>
<td>c</td>
<td>2.319*</td>
<td>0.0204*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>-0.212</td>
<td>0.832</td>
<td>d</td>
<td>-1.814</td>
<td>0.0696</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>0.645</td>
<td>0.5192</td>
<td>e</td>
<td>1.273</td>
<td>0.2031</td>
<td></td>
</tr>
<tr>
<td></td>
<td>f</td>
<td>0</td>
<td>1</td>
<td>f</td>
<td>2.712*</td>
<td>0.0067*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>g</td>
<td>-0.426</td>
<td>0.6704</td>
<td>g</td>
<td>-0.964</td>
<td>0.3352</td>
<td></td>
</tr>
<tr>
<td></td>
<td>h</td>
<td>1.064</td>
<td>0.2873</td>
<td>h</td>
<td>-0.315</td>
<td>0.7526</td>
<td></td>
</tr>
</tbody>
</table>

*: significant at 5% level
The tested term is position-wise contribution rate under incomplete information minus that under complete information.
The unit of test is the contribution rate per cohort per network position.

5.6. Individual behavior

In testing the foregoing hypotheses on equilibrium, network effects, and information effect, we explain behavior on different levels of aggregation (i.e. treatment level, degree level, position level). In this section, we shall investigate behavior on individual level, which could serve as a “root cause” of aggregate level observations. Inspired by Berninghaus et al. (2002), we examine two types of behavior in laboratory networked games: 1) myopia, meaning that subjects best-respond to their neighbor decisions last round when they had the same degree type as they do in the present period, and 2) inertia, meaning that subjects adhere to their own action last round when their degree was same as it is now. As Figure 2-7 shows, player contribution rates generally decline with the number of their neighbors who contributed last round (when the player in question had the same degree as does she in the present period). This indicates the potential of myopia in playing the game. Meanwhile, Figure 2-8 shows that the contribution rate for players

29 The original concept of myopia (inertia) in Berninghaus et al. (2002) is about best response (adherence) to neighbor (own) actions last round. In our context, since subjects’ degrees alter over rounds, it is natural to relate subject’s present decision to her past experience under the same degree type. As will be seen, defined this way, myopia and inertia are significant behavioral factors.
of each degree type in every treatment is higher when the player contributed in the previous round under the same degree type, than it is when the player chose to free ride. Therefore, one’s propensity of contribution is positively related to her own contribution level in the previous round of game (under the same degree type), which implies inertia in game play.

In Section 5.5 we find that the network information plays a rather limited role in affecting free riding of subjects. Therefore, this section shall aim to explain behaviors under both complete and incomplete network information through a unified model on the individual level. Inspired by the finding (Section 5.3) that one’s degree significantly influences free riding under incomplete information (while the network density does not; Section 5.4), we shall attempt a model that

![Figure 2-7. Individual behavior: myopia](image1)

![Figure 2-8. Individual behavior: inertia](image2)
includes one’s degree (but not the network density) as a predictor of behaviors for both informational settings.

To proceed, denote by $lact$ and $lsna$ respectively one’s own action and the sum of neighbor actions (or the number of contributing neighbors), in the last period in which the subject had the same degree as she does now. Then we perform the following logistic regression:

$$Pr\{\text{action} = 1\} = \frac{1}{1 + \exp(-\beta_0 - \beta_{lsna} lsna - \beta_{lact} lact - \beta_{degree} degree)}$$ (2-8)

As suggested by Table 2-9, the free riding patterns under complete and incomplete network information demonstrate a remarkable level of analogy: Both $lsna$ and $lact$ are significant terms (p-values < 0.01), of whose coefficients the signs indicate respectively myopia and inertia in subject’s behavior. Besides those, the effect of one’s degree in public goods contribution is strongly present in both information settings (at 1% significance level), thereby extending the local network effect defined and observed for incomplete information treatments (Section 5.3).

Table 2-9. Individual behavior: logistic regression

<table>
<thead>
<tr>
<th>Dep. var. = action</th>
<th>incomplete information</th>
<th>complete information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>Significance</td>
</tr>
<tr>
<td>$lsna$</td>
<td>-0.570</td>
<td>**</td>
</tr>
<tr>
<td>$lact$</td>
<td>1.551</td>
<td>**</td>
</tr>
<tr>
<td>$I_{degree}$</td>
<td>-0.823</td>
<td>**</td>
</tr>
</tbody>
</table>

Note: * (**): 5% (1%) level of significance

The robust standard errors clustered on cohorts are used in the regression.

In principle, the local network effect is anticipated only when the network information is incomplete (c.f. Proposition 2-1). Nevertheless, the same effect remains striking when subjects are provided with complete network information. To understand this phenomenon, we need to recall an important property of our game: By Assumption 2-1, a low degree player faces a game

---

30 We have also run models involving network density as one of independent variables. However, it turned out that network density is statistically significant in neither informational setting. This in fact supports Observation 5.4 about the lack of global network effect in the incomplete information setting.

31 To help memorize the regression notations, $lact$ stands for “lag action”, while $lsna$ is literally short for “lag sum of neighbor actions”.

50
that is equivalent to a high degree player having one of her neighbors free rides with probability 1. Thus contribution of the low degree type is more robust to the variation of neighbor actions than that of the high type. This fact is likely to solicit a common belief among all players that contribution level declines with one’s degree. With negative neighbor affiliation, such belief can be reinforced and turn into the actual outcome in both complete and incomplete information settings (as what we have observed). Notice that the above reasoning only requires localized network information (neighbor degrees and the correlation of neighbor degrees), and is independent of whether exact information about the global network is available or not. Therefore, the reasoning applies to both information settings and induces the observed effect of one’s degree in both settings.

6. Conclusions and Discussion

We examine a laboratory public goods contribution game in which each player engages with an exogenous subset of the player population. The set of relationships indicating “whose action affects whose payoff” defines a social network. Our experiments shed light on how the network structure, as well as information about the network structure, affects and the extent of free riding. A theory that assumes incomplete network information suggests the following empirical hypothesis: 1) local network effect: Players with more connections in the network show higher tendency to free ride. 2) global network effect: Players conduct less free riding in denser network. The local network effect is strongly observed in our experiment, even in cases where players have complete network information. Global network effect, on the contrast, is not significantly observed. Our results also suggest the information about network architecture has only limited effect on the behavior, and in cases where such effect is significant, providing incomplete network information may assist reducing the level of free riding of highly connected players. On individual level, we find that both behaviors under complete information and incomplete

---

32 It is worth emphasized that, Assumption 2-1 alone cannot ensure equilibria to be non-increasing. To guarantee all equilibria are non-increasing, neighbor degrees have to be independent (Proposition 2, Galeotti et al. 2010). To see the intuition, note by Assumption 2-1 one can legitimately suppose that a player with degree $n$ is playing with $n$ real neighbors and 1 fictional neighbor who free rides with probability 1. Then, adding a new neighbor to the player is equivalent to turning the fictional neighbor to a real one who has positive probability to contribute. In this course, the player’s prior on the contribution from incumbent neighbors does not change, due to neighbor degree independence. Therefore, the player with increased degree must free ride with higher probability. In a behavioral sense, however, Assumption 2-1 might well induce belief on decreasing strategy.
information can be explained by a unified framework, which encompasses one’s degree and one’s tendency of reacting to past actions of neighbors and oneself.

Traditional laboratory studies on the cooperation for public goods contribution focus on the circumstances under which subjects interact globally (c.f. Chapter 2 of Kagel and Roth (1997), Hamburger, et al. 1975). That is equivalent to interactions upon a complete network, with everyone connecting to everyone else. In this case, agent decisions hinge on network size. Our paper concerns situations where one interacts with a connected subset of agent population. As a consequence, what matters to decision making is the connection structure. By the nature of localized interactions, agents in our game foremost react to their neighbors. Therefore, the most noticeable effect on game play comes from one’s local network. Changes outside one’s neighborhood (e.g. variation of the global network) may only indirectly affect player decisions, by altering one’s belief on neighbors’ tendency of free riding.

In theory, the introduction of incomplete information on network structure dramatically simplifies the models of network games and yields monotonic network effects on equilibrium (Galeotti et al. 2010, Sundararajan 2007). In our experiment we find, furthermore, some of these simple network effects even carry over to the game play with complete network information, where the theoretical predictions are hard to obtain. Therefore, this paper suggests that, while social / economic networks in real world are complex, strategic behavior in these networks may follow some simple and tractable pattern. The use of network information in actual decision makings may reflect the bounded rationality of players in reacting to the network institution.

The player payoff in our game exhibits a property that equates having a neighbor free riding for sure to not having that neighbor. Although this property (Assumption 2-1) fits in a broad class of applications and is important to shape the local network effect (as elaborated in Section 5.6), it remains an interesting project to inspect the robustness of the observed network effects to other scenarios without Assumption 2-1.

Social / Organizational networks in practice usually have large sizes. Will the same conclusion we draw upon laboratory networks hold in networks of much larger scale? Note when the network is large, the information that individuals have about the network is likely to be incomplete. Therefore in large networks, one should anticipate the network effects derived from the incomplete information theory, which are independent of network size (Proposition 2-1 and -2, Section 3). Of the two sources of network effects, the local network effect would appear more salient than is the global network effect in influencing free riding (see Section 5.3 and 5.4). These conjectures can be tested by replicating our experiment under various network sizes.
In practice, the level of network information that individuals have might fall between the two extreme cases we’ve studied (settings of complete information and incomplete information as defined in Section 3.1). For example, players in reality may know exactly their own degree and the degrees of their direct neighbors, but not the degrees of people beyond their neighborhood (knowledge of which could be represented by some probability distributions). Thus another extension of our experiment is to investigate these in-between information scenarios.

References


Roberts, Jennifer. *A brief introduction to social network analysis.*


Chapter 3

Innovation, Citation, and the Emergence of Knowledge Core

1. Introduction & Background Literature

Empirically, the citation networks of intellectual products often hold a core-periphery structure: A group of interconnected artifacts (e.g. patents or technical articles) contains the core knowledge to the field, which is commonly cited by peripheral artifacts (Jaffe and Trajtenberg (2002), Doreian (1985), Brass (1984), Kanter (1977), Mullins et al (1977)). As an example, Figure 3-1 sketches typical landscapes for citation networks of patents and research papers, each with a clear-cut core constituted by highly cited objects, and peripheral patents / papers connected to the core via citation links.

![RFID patent citation network, source: Fig. 2, Hung and Wang (2010)](image)

Figure 3-1. Exemplary citation networks

We develop a simple model that produces the core-periphery structure as a consequence of decentralized innovation and citation. To be specific, we consider the citation networks as formed from voluntary knowledge contributions from individual authors who, by doing so, maximize
their own utilities associated with knowledge. To output her artifact of new knowledge, the author has to learn from existing knowledge and thereby creates citation to existing artifacts.

Our paper is primarily motivated by the literature on games of information provision in networks. Viewing information as public goods, Bramoull and Kranton (2007) and Bramoull et al (2010) study public good games upon an exogenously given social network, while Galeotti and Goyal (2010) endogenizes the network structure. On the other direction, Hojman and Szeidl (2008) investigates a problem of network formation, but considering public goods as exogenously endowed to agents in the network. Cho (2009) models networked public good supply under a bargaining framework. We examine a payoff scheme similar to Galeotti and Goyal (2010). However, unlike Galeotti and Goyal (2010), our model features a dynamic horizon.


We are also aware of papers that study core-periphery network structures in other relevant contexts. Borgatti and Everett (1999) studies characteristics and measures of core-periphery topology in general networks. Weng et al (2009) presents statistical evidence for knowledge clustering in technological development networks. In a recent paper, Lovejoy and Sinha (2010) report that core-periphery social networks are efficient institutions for spreading innovation. The focus of Lovejoy and Sinha (2010) is given to the efficient diffusion of exogenous information, while our paper further incorporates the knowledge supply into the scope.

The rest of the paper is organized as follows. Section 2 sets up the model. Section 3 presents the analytical results. Section 4 concludes and discusses future works.
2. The Baseline Model

2.1. The citation network

Individual authors arrive to the knowledge field according to a prescribed arrival process, which keeps arrivals one at a time and has an increasing failure rate -- At each period the process continues with probability \( p(t) \) that decreases with time \( t \). Upon arrival the author produces a single artifact and then leaves, with her artifact remaining in the system (therefore citable by future authors). This way creates one-to-one mappings between authors, artifacts, and system time: At period \( i (i = 1,2,3...) \), author \( i \) composes an artifact \( i \). In producing her artifact, the author both creates knowledge (by innovation) and seeks knowledge (by citation). In the latter case a citation link is made between the knowledge artifact in question and an existing one from which the knowledge is sought. We use \( N_i = \{1,2\ldots\} \) as set of artifacts, and \( E_i = \{g_{jk} \in \{0,1\}; j,k \in N\} \) as the set of connections between artifacts. Hereafter we call an artifact that contains \( j \) units of knowledge \( j \)-artifact, and a \((j+1)\)-artifact larger or higher than a \( j \)-artifact. Throughout the paper we use innovation, knowledge creation, and knowledge production indifferently, refer citation as knowledge-seeking in appropriate contexts, and refer an author of knowledge artifact as a player.

2.2. Individual author’s problem

Let \( x_i \) be the new knowledge created by player \( i \), \( x_i \in \{1,2,\ldots X\} \). \( g_{ji} = 1 \) if artifact \( j \) is cited by artifact \( i \), and \( =0 \) otherwise.\(^{33} \) \( g_i = (g_{ji})_{j=1\ldots i-1} \cdot x_i \) and \( g_i \) are decision variables of player \( i \). The marginal cost for creating new knowledge, \( c(y) \), is negatively affected by the amount of knowledge cited from existing artifacts, \( y \).\(^{34} \) We further assume \( c'' \gg 0 \) to limit the extent that knowledge seeking improves innovation. Each citation incurs a cost \( k \) to the author who cites, which summarizes the costs for locating the targeted artifact, digesting its content, and transcribing the information into one’s own work, etc. Each player \( i \) faces the state of the system, represented by the collection of preceding knowledge production levels, \( s_t = (x_3,x_2,\ldots x_{t-1}) \).

(Notationally we write \( s_{t+1} = (s_t,x_t) \).) We allow an initial part of \( s_t \) to be exogenous and refer

\(^{33} \) The order of the subscript in \( g_{ji} \) indicates the direction of information flow from \( j \) to \( i \).

\(^{34} \) The literature of knowledge management usually considers innovation as recombining and transforming the learned knowledge to new knowledge (see Borgatti and Cross (2003) for example). This idea is reflected in our model through the functional form of \( c \) (“the more you learn, the more you could innovate”).
to it as *initial state*. The benefit for each author $i$ consists of two parts, an instantaneous reward by the amount of new knowledge she creates, $f(x_i)$, and expected rewards proportional to the number of citations artifact $i$ will receive in future, denoted by $D_i(s_i, x_i)$. We make the following assumptions in the model.

**ASSUMPTION 3-1.** $f' > 0, f'' < 0, f(0) = 0, c' < 0, c'' \gg 0$.

**ASSUMPTION 3-2.** The player has complete information on the amount of knowledge in existing artifacts.

Note $f(\cdot)$ is concave and increasing in its argument by Assumption 3-1. Assumption 3-2 is an approximate for many epistemic communities where informational products and citation lists are transparent and synchronized, e.g. patent systems, academic journals. Underlying the complete information setup, there is also an associate assumption that, once completed, the knowledge artifact will be made public, with explicit references to existing works. This assumption is compatible with many field practices.\(^\text{35}\)

**ASSUMPTION 3-3.** When the player is indifferent between any groups of preceding artifacts to cite, she will cite the group with the smallest sum of artifact indexes.

Simply put, Assumption 3-3 means that earlier created artifact shall be cited *in case of a tie in player preference*. In fact, Price (1965) suggests that older papers are frequently favored for citation in fields including geology and mathematics, which supports the above assumption.

Altogether, the player chooses innovation level $x_i$ and citation choice $g_i$ to maximize her payoff given as Equation (3-1) below.

$$\Pi_i(s_i; x_i, g_i) = \pi_i(s_i; x_i, \sum_{j<i} g_j x_j) - k \sum_{j<i} g_j,$$  \hspace{1cm} (3-1)

where

$$\pi_i(s_i; x_i, y_i) = f(x_i) - c(y_i)x_i + D_i(s_i, x_i).$$  \hspace{1cm} (3-2)

$$D_i(s_i, x_i) := R \sum_{j=i+1}^{\infty} n_i^j p(i) g_j^* \left( (s_i, x_i, x_{i+1}^*, x_{i+2}^*, \ldots, x_{i+j-1}^*)(s_{i-1}) \right),$$

$x_j^*(s_j), g_j^*(s_j)$ are the best response of player $j$ to state $s_j$. On the equilibrium path we can write

\(^\text{35}\) In many scenarios like academic publishing, the benefit to the author is only received after the artifact is made public. Therefore, it is reasonable to assume that the knowledge artifacts are publicly visible and thus quotable to later community members.
That said, we are ready for the results of the model. We shall first establish the equilibrium strategy for this game, and then inspect the network structure generated throughout the equilibrium play.

3. Equilibrium Analysis

3.1. General results

Denote by \( x^*_i(y_i) \) the optimal \( x_i \) for given cited knowledge amount \( y_i \), and \( \pi^*_i(s, y_i) := \pi_i(s; x^*_i(y_i), y_i) \). By the supermodular setup, it is easy to see that \( \pi^*_i(s, y_i) \) increases in \( y_i \), and is concave in \( y_i \) given \( c'' \) sufficiently large (Assumption 3-1). Therefore, the player's strategy is characterized by two sets of thresholds: \( m_{ij}(s_i) \) as the amount of cited knowledge needed by player \( i \) to optimally produce a \( j \)-artifact under state \( s_i \), and \( m^j_i(s_i) \) as the threshold amount of cited knowledge upon which player \( i \) should stop citing more \( j \)-artifacts with state \( s_i \).

For convenience, let \( s^j_i \) be the number of \( j \)-artifacts at period \( i \). Denote by \( X_i \) the highest non-empty class of artifacts at period \( i \), i.e. the largest \( j \) such that \( s^j_i > 0 \).

**Lemma 3-1. Player strategy.** The strategy of player \( i \) takes the following form: She keeps citing \( X_i \)-artifacts until either the threshold \( m^X_i(s_i) \) is reached or all such artifacts have been cited. In the former case stop citing. In the latter case, stop citing if \( m^X_i(s_i) \) is reached, otherwise keep citing \( X_i - 1 \) artifacts until either threshold \( m^{X-1}_i(s_i) \) is reached or all such artifacts have been cited... Repeat the process down to \( 1 \)-artifacts. The threshold \( m^1_i(s_i) \) is determined to be either equal to \( m_{it}(s_i) \) for some \( t \), or such that \( \pi^*_i(s, m^1_i(s_i) + j) - \pi^*_i(s, m^1_i(s_i)) = k \). \( ^{37} \) Player \( i \) produces \( x^*_i \) (cited knowledge amount).

**Lemma 3-2.** For any given player \( i \), \( m_{ij}(s_i) < m_{i,j+1}(s_i) \) and \( m^j_i(s_i) < m^{j+1}_i(s_i) \), \( \forall j \in \{1, 2 \ldots X - 1\} \).

\(^{36}\) It may be helpful to note that, given \( y_i \), a \( x^*_i(y_i) \) that is optimal to \( \pi_1(s_i; x_i, y_i) \) is also optimal to \( \Pi_1(s_i; x_i, g_i) \), assuming the mapping between \( g_i \) and \( y_i \) given \( s_i \) is one-to-one. \n
\(^{37}\) We assume the unity of knowledge is small enough so that the equalities can hold.
Unless otherwise specified, the proofs are found in the appendix. Lemma 3-1 establishes the optimal form of player strategy, and Lemma 3-2 orders the thresholds associated with the strategy for a given player. Based on those results, we will show that the core-periphery structure can form if the system begins with sufficiently large artifacts.

**PROPOSITION 3-1. Equilibrium.** Suppose the system begins with an initial state constituted by artifacts \(\{1, 2, \ldots, i - 1\}\). If \(\sum_{j=x-1}^{X} j^{\lambda}_t > m_{t+2}^{X-1}(s_{i+2}) \forall x_i, x_{i+1}\), the equilibrium with \(m_{t+1}(s_t) < m_{t+1, j}(s_{t+1}), m_t^j(s_t) > m_{t+1}^j(s_{t+1}), \forall t, x_t, j\) is subgame perfect. The knowledge production on the equilibrium path decreases over time since artifact \(i\). The network formed on the equilibrium path exhibits a core-periphery structure. It has every artifact since \(i + 2\) (periphery) connecting not with each other, but to a subset of artifacts \(\{1, 2, \ldots, i + 1\}\) (core).

**Proof.** The proof of Proposition 3-1 is done by induction. First observe that as time advances, there will be a player (indexed by \(l\)) that faces so large the probability of the field being terminated that the consideration of future citation reward becomes trivial. It follows from Lemma 3-3 below that \(m_t^l(s_{i-1}) < m_{t+1}^l(s_{i-1}), m_t^l(s_t) > m_{t+1}^l(s_{i+1}) \forall x_i, x_{i+1}\).

**LEMMA 3-3.** If \(D_t(s_t, x + 1) - D_t(s_t, x) > D_{t+1}(s_t + 1, x + 1) - D_{t+1}(s_t + 1, x), \forall x, x_{t+1}\), then \(m_{t+1}^l(s_t) < m_{t+1}^j(s_{i+1}), \text{ for all } x_t, x_{t+1}\). Then we hypothesize the same threshold order through player \(i + 1\): \(m_t^l(s_{i-1}) > m_{t+2}^l(s_{i-2}) \forall x_{i-2}, m_{t+2}^l(s_{i-2}) > m_{t+3}^l(s_{i-3}) \forall x_{i-3}, \ldots, m_{t+2}^j(s_{i+2}) > m_{t+1}^j(s_{i+1}) \forall x_{i+1}, \text{ and } m_{t+1}^l(s_{i-1}) < m_{t+2}^l(s_{i-2}) \forall x_{i-2}, \ldots, m_{t+2}^j(s_{i+2}) < m_{t+1}^j(s_{i+1}) \forall x_{i+1}\), and try to prove it for player \(i\). The condition \(\sum_{j=x-1}^{X} j^{\lambda}_t > m_{t+2}^{X-1}(s_{i+2}) \forall x_i, x_{i+1}\) suffices to imply that player \(i + 2\) (and players afterwards by hypothesis) will not have the capacity to cite all artifacts of the top two classes that were created before \(i\), regardless of the knowledge production levels of players \(i\) and \(i + 1\). Therefore, the decrease of \(m_t^i(\cdot)\) thresholds will lead to the decrease of the actual amount of cited knowledge over time (decreasing by \(t\)). By hypothesis however, the \(m_t^i(\cdot)\) thresholds are increasing over time. That is, players over time need to cite more existing knowledge to produce any given level of new knowledge. The changes of the two sets of thresholds together imply that the optimal knowledge production level, \(x_t^i\), is decreasing over
time for $t \geq i + 1$. Per the optimal player strategy (Lemma 3-1), any artifact from $i + 2$ onwards therefore only competes for future citation with the artifacts produced before itself. That allows us to separately study whether artifact $i + 1$ will be cited by each incoming player $t$ ($t \geq i + 2$) on the equilibrium path:

$$g_{i+1,t}((s_i,0,x)) - g_{i+1,t}((s_i,x_i,x)) =
\begin{cases}
0, & \text{if } x_i < x \\
1, & \text{if } \frac{\sum_j (s_i)^j}{\sum j x_j^i} < m^i_t((s_i,0,x)), \frac{\sum_j (s_i)^j + x_i}{\sum j x_j^i} > m^i_t((s_i,x_i,x)) \\
0, & \text{otherwise}
\end{cases}
$$

(3-3)

To see Equation (3-3) holds, notice any $x_i < x$ will not affect whether an $x$-artifact produced by player $i + 1$ will be cited by the incoming player $t$. The case $x_i \geq x$ (equality included due to the first-come-first-cite tie breaking rule in Assumption 3-3) will however make a difference, if the introduction of the $x_i$-artifact causes player $t$ no longer able to finish citing all $x$-artifacts up to period $i + 1$.

That provided, it suffices to show, for any $x_i$,

$$g_{i+1,t}^*((s_i,0,x + 1)) - g_{i+1,t}^*((s_i,0,x)) \geq g_{i+1,t}^*((s_i,x_i,x + 1)) - g_{i+1,t}^*((s_i,x_i,x))$$

(3-4)

if

$$\sum_{j=x-1}^X js_i^j > m_{i+2}^{X-1}(s_{i+2}), \forall x_i, x_{i+1}$$

(3-5)

Taking expectation to both sides of (3-4) over $t$ yields, under (3-5),

$$D_{i+1}((s_i,0),x + 1) - D_{i+1}((s_i,0),x) \geq D_{i+1}((s_i,x_i,x + 1) - D_{i+1}((s_i,x_i,x)), \forall x_i$$

(3-6)

Since $p(j) > p(j + 1) \forall j$, we have $D_i(s_i,x + 1) - D_i(s_i,x) > D_{i+1}((s_i,0),x + 1) - D_{i+1}((s_i,0),x)$. Combining it and (3-6) results in the desired relationship $D_i(s_i,x + 1) - D_i(s_i,x) > D_{i+1}(s_{i+1},x + 1) - D_{i+1}(s_{i+1},x), \forall x_i$, which by Lemma 3-3 suggests that $m_{ij}(s_i) < m_{i+1,j}(s_{i+1})$ and $m_{ij}(s_i) > m_{i+1,j}(s_{i+1}), \forall x_i, j$. Therefore the hypothesis is proved true for player $i$, and we have $X \geq x_i > x_{i+1} > x_{i+2} > x_{i+3} > \cdots$. With the unity of knowledge small enough, $x_{i+2} \leq X - 1$. Under (3-5), player $i + 2$ produces less or equal knowledge than the lowest class of artifacts he can cite. That the capacity for citation shall decrease in future guarantees that $i + 2$ and artifacts afterwards will not have any future citations and thus form the periphery.

Proposition 3-1 qualifies the initial state needed to generate the core-periphery structure of the citation network. Will such state be surely reached on a ``complete" evolutionary path that
commences at an empty state? In general, it is difficult to construct an equilibrium with increasing capacity for citation and decreasing threshold amount of cited knowledge needed to upgrade knowledge production, before the system reaches the threshold state defined by (3-5). Nevertheless, the restriction on the initial state vanishes in a long horizon, when we narrow down the model into binary knowledge production ($X = 2$).

### 3.2. A special case with binary knowledge production

In this section, we consider a mode of binary knowledge production. That is, an artifact contains either $H$ or $L$ units of knowledge, and we normalize $L = 1$. This captures settings where knowledge products in a field can be roughly classified as high-value products or low-value ones.

**Corollary 3-1. Binary knowledge production.** In a system with $X = 2$, the core-periphery structure as specified in Proposition 3-1 will emerge with probability 1, conditional on that the system has survived long enough for (3-5) to be satisfied.

The proof of Corollary 3-1 is obvious and omitted. Since there are only two categories of artifacts, the top two artifact classes involved in (3-5) cannot be both empty. As time advances, the qualification (3-5) will be satisfied with probability 1, regardless of the player strategy. The remaining question is how the innovation level changes (before it begins to strictly decline as implied by Proposition 3-1). The system may have to start with $L$-artifacts because of the lack of citable knowledge, then upgrade into $H$-artifacts as it moves on with a good prospect of future citations; and eventually reach the threshold state that triggers the afterwards decrease of knowledge creation.

To illustrate, Figure 3-2 depicts a typical core-periphery citation network that could be possibly generated by our model. Artifacts 1 to 8 consist of the knowledge core, which is surrounded by artifacts created later, representing the "periphery" of knowledge.
4. Conclusions and Discussion

Through a simple analytical model, we explain the emergence of core-periphery architecture in knowledge citation networks as an accumulative process, driven by innovation and citation of individual forward looking authors. We characterize the subgame perfect equilibrium for the game played by the authors. Specifically, the core-periphery structure forms on the equilibrium path if the initial state contains sufficient amount of knowledge. If the artifacts have binary knowledge levels, the citation network will take on the core-periphery shape for sure regardless of its initial state, once it has evolved for adequately long time.

Citation network is usually considered as a “window on the knowledge economy” (Jaffe and Trajtenberg, 2002). Our paper reveals several factors that may drive the evolution of a citation network, and in particular, shape its core-periphery landscape: knowledge creation, knowledge citation, and anticipated credits from future citations. These findings may potentially assist a policy maker with the design and development of epistemic / scientific community to sustain contributions of new knowledge and facilitate knowledge sharing.

We consider knowledge creation and citation as simultaneous and complementary choices from an individual author. It remains an interesting question whether the core-periphery network layout is robust to sequential decisions on innovation and citation. For instance, an author may...
first cite existing artifacts, then create her own knowledge at a cost negatively affected by the volume of cited knowledge.

The core-periphery configuration modeled in this paper is a high abstraction of the actual structure of citation networks. While many practical measures have been proposed (see Borgatti and Everett, 1999) to gauge core-peripher-y-ness, our model only reflects its ideal form. Besides core-peripher-y-ness, there are some other statistical properties of citation networks which our model has not yet captured, e.g. degree distribution, clustering, density (c.f. Bonacich 1987 for discussion). To explore this issue, we must evaluate actual networks generated under specific and representative parameterization of the model. This will be an important direction for future works.

References


Appendix A

Background Theory for Chapter 1

We first introduce notations. Unless otherwise noted, the theoretical work in this section is attributed to Galeotti et al. (2010). The set of feasible degree values in a size $N$ network is denoted by $\kappa := \{1, 2, \ldots, N - 1\}$. Recall $\underline{k}$ ($\overline{k}$) is the lowest (highest) degree value in a given network. Let $G(k|k)$ be the probability that neighbor degrees are ($k$-dimensional vector) $k$, conditional on one’s own degree being $k$. If player’s degrees are independent, we drop the condition on distribution and refer to $G(k)$ as the probability that a single neighbor’s degree equals to $k$. Describe network density $F$ as the collection of conditional neighbor degree distributions $G(\cdot | k)$ for every degree type $k$, i.e. $F := \{G(k|k)\}_{k \in \kappa}$. Define $f$ as a non-decreasing mapping $f: k^k \rightarrow \mathbb{R}$. Converting any neighbor degree vector to a scalar, $f$ can be thought as a scoring mechanism: Any collective uplift of (any subset of) neighbor degrees will lead to higher scores. Let $E_{G(\cdot | k)}[f] := \sum_{k \in \kappa} G(k|k)f(k)$, which can be understood as an ex ante score of neighborhood of a given degree type $k$. Based on that, we introduce a concept that describes how neighbor degrees are correlated: A network exhibits positive neighbor affiliation, if

$$E_{G(\cdot | k')}[f] \geq E_{G(\cdot | k)}[f],$$

for all $k' > k$ and any non-decreasing $f$.\(^{38}\) The neighbor affiliation is negative when the inequality is reversed.

Positive neighbor affiliation means that higher own degree will raise the distribution of neighbor degrees (in the way defined by the $f$ operator). Besides positive neighbor affiliation, the other case we are focusing on is neighbor degree independence, in which the neighbor degree distribution can be defined for single neighbor (a joint distribution is no longer needed). The independence or correlation of neighbor degrees are measures of one’s local network. Next we shall discuss the measure on global network – network density. Define a network $F'$, $F' := \{G'(k|k)\}_{k \in \kappa}$. The network $F'$ is said to have higher density than does another network $F$, if for any non-decreasing $f$

$$E_{G'(|k)}[f] > E_{G(\cdot | k)}[f], \forall k.$$

---

\(^{38}\) Note that in the left side hand of (A-1), $f$ operates on a subset of $k$ neighbors out of the total $k'$ neighbors.
That is, if one network allows, for every degree type, higher neighbor degree distribution than does the other, then the former network is denser. Next we restate and prove the theoretical results in the main text.

**PROPOSITION 1-1. Equilibrium structure. a. degree-pooling equilibria.** There exists a pooling equilibrium where all players defect. If \( C < R \) another pooling equilibrium exists with everyone coordinating. **b. degree-separating equilibria.** In some cases there exists separating equilibrium increasing in degree. Moreover, such equilibrium involves a threshold \( t \), such that a player coordinates if having more than \( t \) neighbors, defects if having less than \( t \) neighbors, and (only if \( t \) is integer) coordinates with probability \( p_t \in (0,1) \) if having exactly \( t \) neighbors. When neighbor degrees are independent, the degree threshold is solved from the following equation.

\[
 t = \frac{c}{R \left( 1 - \sum_{k=1}^{t} G(k) + p_t(t) \right)}
\]  

(A-3)

**Proof.**

The proof on the existence of degree pooling equilibria is straightforward and omitted. Galeotti et al. (2010) has shown that there exists a separating equilibrium non-decreasing in degree (Proposition 1, Galeotti et al. 2010). In a binary action coordination context, any increasing equilibrium (denoted by \( \sigma \)) entails a threshold, which we will solve explicitly when neighbors degrees are independent. Let \( k_{N_i} \) be the \((k_i\text{-dimensional})\) vector of degrees of player \( i \)'s neighbors and \( k_{N_i} = (k_1, k_2, \ldots, k_{k_i}) \). Notice

\[
 E \left( \sum_{j=1}^{k} x_j \mid k \right) = \sum_{k_1, k_2, \ldots, k_k} E \left( \sum_{j=1}^{k} \sigma(k_j) \right) G(k_1, k_2, \ldots, k_k \mid k)
\]

\[
 \equiv \sum_{x_{k+1}=0}^{x_{k+1}} \sum_{k_1, k_2, \ldots, k_k} E \left( \sum_{j=1}^{k} \sigma(k_j) + x_{k+1} \right) G(k_1, k_2, \ldots, k_k \mid k + 1)
\]

\[
 \leq \sum_{x_{k+1}=0}^{x_{k+1}} \sum_{k_1, k_2, \ldots, k_k, k_{k+1}} E \left( \sum_{j=1}^{k} \sigma(k_j) + x_{k+1} \right) G(k_1, k_2, \ldots, k_k, k_{k+1} \mid k + 1)
\]

\[
 \leq \sum_{k_1, k_2, \ldots, k_k, k_{k+1}} E \left( \sum_{j=1}^{k+1} \sigma(k_j) \right) G(k_1, k_2, \ldots, k_k, k_{k+1} \mid k + 1)
\]

\[
 = E \left( \sum_{j=1}^{k+1} x_j \mid k + 1 \right)
\]
The second equality above is induced by introducing one defecting neighbor to a degree-$k$ player (who takes action 0 and thus does not change the expected sum of neighbor actions). The first $\leq$ holds because $E(\sum_{j \in N_i} \sigma(k_j))$ is nondecreasing in $k_i$ (given $\sigma$ increasing) and yields a scalar, so that the property of positive neighbor affiliation applies. The second $\leq$ holds because $\sigma$ is equilibrium. All that said, $E(\sum_{j \in N_i} x_j | k_i)$ increases in $k_i$. According to equation (1-1) in the main text, player $i$ will coordinate (defect; be indifferent) if the expected number of coordinating neighbors exceeds (does not exceed; equals to) $C/R$. When $E(\sum_{j \in N_i} x_j | k_i)$ crosses with $C/R$ line in the valid degree range $[k_i, k]$, the separating equilibrium exists and takes the above threshold structure.

When neighbor degrees are independent, the number of coordinating neighbors of player $i$ follows Binomial distribution. For player $i$ suppose her neighbors play $t$-strategy, the probability of a single neighbor coordinating is $1 - \sum_{k=1}^{t} G(k) + p_t G(t)$. The expected payoff for player $i$ to coordinate is $Rk_i \left(1 - \sum_{k=1}^{t} G(k) + p_t G(t)\right)$ while that for her to defect is $C$. Player $i$ is indifferent between coordinating and defecting if $Rk_i \left(1 - \sum_{k=1}^{t} G(k) + p_t G(t)\right) - C = 0$. For the equilibrium to be self-enforcing, degree threshold $t$ must be the solution to the foregoing equation, that is, $t = \frac{C}{R\left(1 - \sum_{k=1}^{t} G(k) + p_t G(t)\right)}$. Then the desired result follows. Q.E.D.

**Proposition 1-2. Comparative statics of equilibria.** If the network density is increased, there exists an equilibrium that weakly payoff-dominates ($\succeq$) the current one, with lowered equilibrium threshold and every degree type coordinating with weakly higher ($\succeq$) probability.

**Proof.**

Suppose the neighbor degree distribution (network density) increases from $F(\cdot)$ to a denser network $F'(\cdot)$. The (increasing) equilibrium changes from $\sigma$ to $\sigma'$, and the degree threshold from $t$ to $t'$. 39 We will show that it is possible (though not necessary) that $t' < t$. Assuming $t' < t$, we have

---

39 The expression of threshold equilibrium also applies to degree pooling equilibria, for which the thresholds are either lower or higher than all the degree types in the network, such that all players take the same action in equilibrium.
\[
\frac{C}{R} = E_\sigma \left( \sum_{j=1}^{t} x_j | t \right) = \sum_{k_1, k_2, \ldots, k_t} E \left( \sum_{j=1}^{t} \sigma(k_j) \right) G(k_1, k_2, \ldots, k_t | t) < \sum_{k_1, k_2, \ldots, k_t} E \left( \sum_{j=1}^{t} \sigma(k_j) \right) G'(k_1, k_2, \ldots, k_t | t) \leq \sum_{k_1, k_2, \ldots, k_t} E \left( \sum_{j=1}^{t} \sigma'(k_j) \right) G'(k_1, k_2, \ldots, k_t | t) = E_{\sigma'} \left( \sum_{j=1}^{t} x_j | t \right).
\]

Given \( E \left( \sum_{j=1}^{k} \sigma(k_j) \right) \) increasing in \( k \), the first "<" results from the definition of density (A-2). The \( \leq \) results from the hypothesis that \( t' < t \) or \( \sigma' \) increasing in degree.

The facts that \( E_{\sigma'} \left( \sum_{j=1}^{t} x_j | t \right) > \frac{C}{R} \) and \( E_{\sigma'} \left( \sum_{j=1}^{k} x_j | k \right) \) increasing in \( k \) (a result in the proof of Proposition 1-1) imply \( t' < t \), which reinforces the foregoing assumption. Thus it is easy to adapt the fixed point theorem to find an equilibrium threshold \( t' \) that is smaller than \( t \). Moreover, the lowered coordination threshold \( (t') \) together with the increased network density \( (F') \) implies that every degree type raises its coordination probability in the new equilibrium \( (\sigma') \), and the \textit{ex ante} social welfare is increased. Then the intended payoff-dominance relationship follows. Q.E.D.

Appendix B

Experimental Implementation for Chapter 1

Figure 1-1 (duplicated from the main text). Network structure for a single cohort
This section shows how to derive equilibria upon the networks given in Figure 1-1 and how we inform subjects information about these networks. Attributed to the random assignment of subjects over network locations in an equally likely manner (see Section 1.2.3), players in network \( G_h \) face the following neighbor degree distribution.

For a degree-2 player (he),

- With 1/2 chance, One of his neighbors has 2 neighbors while the other has 3 neighbors.
- With 1/2 chance, each of his neighbors has 3 neighbors.

For a degree-3 player (she),

- With 1/2 chance, One of her neighbors has 2 neighbors while the other two neighbors have 3 neighbors each.
- With 1/2 chance, One of her neighbors has 3 neighbors while the other two neighbors have 2 neighbors each.

Likewise, players in the network \( G_t \) face the following neighbor degree distribution,

For a degree-2 player (he),

- With 2/3 chance, One of his neighbors has 2 neighbors while the other has 3 neighbors.
- With 1/3 chance, each of his neighbors has 2 neighbors.

For a degree-3 player (she),

- Two of her neighbors have 2 neighbors each, and the third neighbor has 3 neighbors.

Claim B-1. \( G_h \) is denser than \( G_t \).

Next we prove Claim B-1. Given the neighbor degree distributions described above (for \( G_h \) and \( G_t \)), For any \( f \) increasing in its argument, we have

\[
E_{G_h[\{2\}]}[f] = \frac{1}{2} f(2,3) + \frac{1}{2} f(3,3) > \frac{2}{3} f(2,3) + \frac{1}{3} f(3,3) > \frac{2}{3} f(2,3) + \frac{1}{3} f(2,2) = E_{G_t[\{2\}]}[f].
\]

\[
E_{G_h[\{3\}]}[f] = \frac{1}{2} f(2,3,3) + \frac{1}{2} f(2,2,3) > f(2,2,3) = E_{G_t[\{3\}]}[f].
\]

Applying the definition of network density (A-2) completes the proof of Claim B-1.

Let \( \Delta \pi \) be the expected difference between payoff of coordination and that of defection, i.e.,

\[
\Delta \pi = \pi_{\text{coordination}} - \pi_{\text{defection}}.
\]

Substituting the above neighbor degree distribution for \( G_h \) yields

\[
E(\Delta \pi | G_h, degree = 2) = 1/2 (p_2^2 (2R - C) + 2p_3 (1 - p_3) (R - C) + (1 - p_3)^2 (-C)) + 1/2 (p_3 p_4 (2R - C) + (1 - p_2) p_3 + p_2 (1 - p_3)) (R - C) + (1 - p_2) (1 - p_3) (-C)
\]

(B-1)
where recall that \( p_i \) is the probability that degree-\( i \) player assigns to coordination. The expressions (B-1) and (B-2) can be reproduced for \( G_1 \) in a similar way.

To work out the equilibria for \( G_h \) we need to solve the systems of inequalities that are built on (B-1) and (B-2). For example, when the low degree type is indifferent while the high type prefers coordination, the system for network \( G_h \) is as follows:

\[
\begin{align*}
E(\Delta \pi | G_h, \text{degree} = 2) &= 0, \\
E(\Delta \pi | G_h, \text{degree} = 3) &> 0, \\
0 &< p_2 < 1, p_3 = 1
\end{align*}
\]

The (multiple) solutions to the systems of inequalities (involving those for both \( G_h \) and \( G_l \)) are the equilibria in Table 1-3 (in the main text).

We find it ineffective for subjects to understand and utilize the stochastic information on neighbor degrees when it is given as the above-described, conditional joint probability distribution. To overcome this issue, we provide subjects an unconditional distribution on neighbor degrees, obtained by integrating the conditional neighbor degree distributions over the distribution of one’s own degree (the latter given as network density, ex. for \( G_h \): probability ½ for being degree-2 and probability ½ for being degree-3). Given that subjects are assigned with degree types according to the prescribed network density, the unconditional neighbor degree distribution we provide is accurate aggregated over all the rounds of game. Also to ease subjects understanding, the degree distribution we provide is defined on single neighbor, which is a close approximation of multivariate degree distribution given the two lead to almost the same equilibrium structure. Notice that network density is the degree distribution of a randomly chosen player in the network, while neighbor degree distribution is the degree distribution of a randomly selected player’s neighbors. In our case, the difference between the two concepts in deriving equilibria is negligible for behavioral purposes. All that said, we use network density (which is a unconditional and uni-variate probability distribution) as the distribution of neighbor degrees, and

\[40\] It is reasonable to assume subjects treat neighbor degrees independently in our experiment, given they have no knowledge on either the detailed connectivity or the size of the global network they play in, neither the subjects know that they are organized into separate cohorts (so that the networks to the subjects are potentially large enough to induce degree independence). Therefore, we present subjects independent neighbor degree distributions (defined on single neighbor) that approximate the exact distributions.
our instruction regarding the global network information takes the following simple form (The example is given for high density networks; see the full instruction in Appendix D):

“...Each of your neighbors has either two neighbors or three neighbors (including yourself). There is 50% of the chance that he or she has two neighbors, and a 50% of chance that he or she has three neighbors...”

Appendix C.

The Heterogeneity of Coordination in Treatment \textit{LrcLden}, Chapter 1

\textbf{Observation C-1.} In low density networks with low economic return (Treatment \textit{LrcLden}), coordination evolves differently in distinct network cohorts.

Figure C-1 plots the degree-specific coordination rates of each network cohort in the first and second half of the game. As the figure suggests, in two cohorts out of five (N16 and N20), coordination rates dramatically drop for both degree types as the game moves to later rounds (The decrements are 38.3% (35.0%) for low (high) degree of N16, and 23.3% (30.0%) for low (high) degree of N20.). The overall coordination level of N20 (21.25%) is considerably lower than that of the N16 (43.75%). In the rest of cohorts the plays separate on degree types. High-degree subjects all perfectly coordinate in these cohorts (coordination rates being 100%), whereas low-degree players behave differently: In two cohorts (N17 and N19) their coordination level ascends as the game advances, with increments of 5.0% for N17 and 23.3% for N19. In one network cohort (N18) their coordinate rate shrinks in the later rounds, changing from 51.7% to 36.7%.
Figure C-1. Coordination trends for each network cohort in Treatment *LrcLden*

Table C-1 reproduces the information that Figure C-1 delivers, but on numerical scales. A logit model is executed on separate datasets of two degree types with dummy variables for cohorts. This generates the output below in Table C-1.

<table>
<thead>
<tr>
<th>Logit Regression</th>
<th>low degree data</th>
<th>high degree data***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dep. var = action</td>
<td>Coef.</td>
</tr>
<tr>
<td>N16</td>
<td>1.032802</td>
<td>2.80892543</td>
</tr>
<tr>
<td>N17</td>
<td>1.887444</td>
<td>6.602474</td>
</tr>
<tr>
<td>N18</td>
<td>-0.64965</td>
<td>0.5222285</td>
</tr>
<tr>
<td>N19</td>
<td>-1.19159</td>
<td>0.3037395</td>
</tr>
<tr>
<td>N20</td>
<td>-0.5268</td>
<td>0.5904943</td>
</tr>
<tr>
<td>period_N16</td>
<td>-0.18604 **</td>
<td>0.830241</td>
</tr>
<tr>
<td>period_N17</td>
<td>0.266339 *</td>
<td>1.305177</td>
</tr>
<tr>
<td>period_N18</td>
<td>0.126575 *</td>
<td>1.134935</td>
</tr>
<tr>
<td>period_N19</td>
<td>0.360273 **</td>
<td>1.43372</td>
</tr>
<tr>
<td>period_N20</td>
<td>-0.14229</td>
<td>0.8673701</td>
</tr>
</tbody>
</table>

Note: * (**): 5% (1%) level of significance

***In logit regression on high degree data, since all actions are equal to 1 in N17, N18 and N19, we cannot obtain any legitimate estimation.
We see evidence of full coordination in three of the cohorts (N17, N18 and N19) with high degree players (See the note for Table C-1). The low-degree ones in each cohort display qualitatively different behaviors: The coefficients for time effects (measured by period interacting cohort variables) range from negative (-0.186 for period_N16) to positive (0.36 for period_N19), implying that in some cohort (N16) coordination of low-degree subjects generally deteriorates over time (p-value < 0.01) while in some others (N17, N18, N19) low-degree coordination level is increasing (All p-values < 0.05) as the game proceeds. The Wald test in Table C-2 evaluates the significance of cohort-related variables in the preceding logit model. It yields both of the p-values < 0.01 for data of the two degree types. Therefore, we claim significant heterogeneity in game play across cohorts in LrcLden.

Table C-2. Wald test for cohort heterogeneity in LrcLden dataset

<table>
<thead>
<tr>
<th>Tested terms*</th>
<th>low degree</th>
<th>high degree</th>
<th>Wald test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N17, N18, N19, N20, N17_period, N18_period, N19_period, N20_period</td>
<td>N20, N20_period</td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>124.34 (8**)</td>
<td>9.30 (2**)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0095</td>
<td></td>
</tr>
</tbody>
</table>

Note: * We test whether these terms have zero coefficients in the preceding logit model.

**Degree of freedom

We conclude that, in Treatment LrcLden where coordination is least attractive in all four scenarios, coordination levels begin to fluctuate across cohorts. Qualitative differences in the evolution of coordination are observed for each individual network cohorts. This instability in behavior results from lack of robustness of coordination, which we explained in the main text --- Coordination from both degree types become sensitive to neighbor actions, and thereby exhibits heterogeneity across various cohorts.
Appendix D.

Experimental Instruction for Chapter 1

We motivate our experiments with the product adoption scenario discussed in the main text. Coordination means purchasing the product (cell phone), defection means not buying the product.

************************************************

General. Welcome and thank you for participating in this experiment. In this experiment you will earn money. The actual amount depends on your decisions and the decisions of other participants. From now on until the end of the experiment, please do not communicate with other participants. If you have any question, please raise your hand. An experimenter will come to your place and answer your question privately.

The Game. In the Experiment we use ECU (Experimental Currency Unit) as the monetary unit. The profits you make during the experiment will be added to this account in ECU. At the end of the experiment, the balance of the account will be converted from ECUs into dollars according to the conversion rate stated below, and paid out in cash after the experiment.

The experiment lasts for 20 rounds. In each round, participants will be organized in a network. In this network, you are connected to either two or three people, who are your neighbors. Every round your neighbors will be different people. You know the number of neighbors whom you are connected, but you will not know their identity. Each of your neighbors has either two neighbors or three neighbors (including yourself). There is 50% of the chance that he or she has two neighbors, and a 50% of chance that he or she has three neighbors.

At the beginning of each round, each person in the network decides whether to purchase a (fictional) cell phone. Each does so without any knowledge of what any other person decides. The profit you earn depends on whether you buy a cell phone and, if so, how many of your neighbors buy a phone. Specifically:

If you don’t buy the cell phone,
• Your profit is 100 ECU.

If you buy the cell phone and you have two neighbors then
• If both neighbors buy the cell phone, your profit is 160 ECU.

We only present the experimental instruction for treatment HrcHden in this appendix. This instruction can be applied to other treatments with straightforward adjustments.
• If exactly one neighbor buys the cell phone your profit is 80 ECU.
• If neither neighbor buys the cell phone your profit is 0 ECU.

This is summarized in the Payoff Table as below, also shown on your computer screen during your play.

<table>
<thead>
<tr>
<th>The number of My Neighbors who Buy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Buy</td>
</tr>
<tr>
<td>My Choice</td>
</tr>
<tr>
<td>Don't Buy</td>
</tr>
</tbody>
</table>

If you buy the cell phone and have three neighbors, then
• If all three neighbors buy the cell phone your profit is 240 ECU.
• If exactly two neighbors buy the cell phone, your profit is 160 ECU.
• If exactly one neighbor buys the cell phone your profit is 80 ECU.
• If no neighbor buys the cell phone your profit is 0 ECU.

This is summarized in the Payoff Table as below, also shown on your computer screen during your play.

<table>
<thead>
<tr>
<th>The number of My Neighbors who Buy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Buy</td>
</tr>
<tr>
<td>My Choice</td>
</tr>
<tr>
<td>Don't Buy</td>
</tr>
</tbody>
</table>
All the scenarios and associated profits are summarized in the Payoff Table as shown below. The payoff table is also shown on your computer screen during your play. The conversion rate is 1ECU=0.003 dollars. Before the game starts, you are required to complete a quiz, which covers the important knowledge about the game. The quiz will be shown on your computer screen. You will start to play the game only after you correctly answer all the questions in the quiz.

Consent Forms. Please read the consent form that is delivered to you before the start of the experiment.

Appendix E.

The Experimental Software Interface for Chapter 1

This section provides snapshots for the experimental software. The software is programmed with zTree (Fischbacher, 2007).

All the relevant snapshots are contained in Figure E-1. Subjects begin with a quiz testing their understanding of the game, with no earning accumulated to the game. The quiz screens are shown in Figure E1-a. The actual decision making interfaces are found in Figure E1-b. For better references, the interfaces for high-degree players and low-degree players are separately presented.
### Quiz (non-paying period)

The number of my neighbors who buy

<table>
<thead>
<tr>
<th></th>
<th>Three</th>
<th>Two</th>
<th>One</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buy</td>
<td>240</td>
<td>160</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Don’t Buy</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

1) Suppose one of your neighbors buys and the others do not buy. Then your profit from buying is

2) You have the same number of neighbors as any other does.

3) Your neighbors change in every period.

4) What is your profit when you don’t buy and all your neighbors buy?
### Quiz (non-paying period)

#### The number of My Neighbors who buy

<table>
<thead>
<tr>
<th></th>
<th>Two</th>
<th>One</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buy</td>
<td>160</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>Don't buy</td>
<td>100</td>
<td>110</td>
<td>100</td>
</tr>
</tbody>
</table>

1. Suppose one of your neighbors buys and the others do not buy. Then your profit from buying is **[ ]**
2. You have the same number of neighbors as any other does. **[ ]**
3. Your neighbors change in every period. **[ ]**
4. What is your profit when you don’t buy and all your neighbors buy? **[ ]**

E1-a. Quiz
You have 2 neighbors.

The number of My Neighbors who Buy

<table>
<thead>
<tr>
<th></th>
<th>Two</th>
<th>One</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Buy</strong></td>
<td>160</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td><strong>Don’t Buy</strong></td>
<td>100</td>
<td>110</td>
<td>100</td>
</tr>
</tbody>
</table>

Action I choose this period:
- Bar
- Don’t Buy

<table>
<thead>
<tr>
<th>Period</th>
<th>Your choice</th>
<th>Your Profit</th>
<th>Your first neighbor’s choice</th>
<th>Your second neighbor’s choice</th>
<th>Your third neighbor’s choice</th>
<th>Your Total Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no Buy</td>
<td>100</td>
<td>Bar</td>
<td>Buy</td>
<td>NA</td>
<td>100</td>
</tr>
</tbody>
</table>
E1-b. Decision Making

Figure E-1. The snapshots for the experimental software interface

Appendix F.

The Derivation of Equilibria for Chapter 2

Table 2-4 in the main text contains equilibria of games in all treatments. This section illustrates how these equilibria are derived. Let the base game be represented by the following
matrix, $V := \begin{bmatrix} 100 & 175 & 225 & 260 \\ 0 & 100 & 275 & 335 \end{bmatrix}$. The components of $V$ are denoted by $V_{i,j}$ ($i = 1,2; j = 1,2,3,4$).

**F.1. Incomplete information**

Recall that the players in networks $G_h$ and $G_l$ face the following neighbor degree distributions as shown in Table F-1 (reproduced from Table 2-3 in the main text) under incomplete information.

<table>
<thead>
<tr>
<th>Probability</th>
<th>No. degree-3 neighbors</th>
<th>No. degree-2 neighbors</th>
<th>Probability</th>
<th>No. degree-3 neighbors</th>
<th>No. degree-2 neighbors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>1</td>
<td>1</td>
<td>2/3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1/2</td>
<td>2</td>
<td>0</td>
<td>1/3</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Applying the notations defined in (2-6) and (2-7), we have for $G_h$:

$U(1, \sigma, 2) - U(0, \sigma, 2) = 1/2 \left( p_2^2(V_{1,3} - V_{2,3}) + 2p_3(1 - p_3)(V_{1,2} - V_{2,2}) + (1 - p_3)^2(V_{1,1} - V_{2,1}) \right) + 1/2 \left( p_2 p_3(V_{1,3} - V_{2,3}) + \left( (1 - p_2)p_3 + p_2(1 - p_3) \right)(V_{1,2} - V_{2,2}) + (1 - p_2)(1 - p_3)(V_{1,1} - V_{2,1}) \right)$.

$U(1, \sigma, 3) - U(0, \sigma, 3) = 1/2 \left( p_2^3(V_{1,4} - V_{2,4}) + 2p_3(1 - p_3)(V_{1,3} - V_{2,3}) + (1 - p_3)^2(V_{1,2} - V_{2,2}) \right) + 1/2 \left( p_2^2(V_{1,3} - V_{2,3}) + 2p_3(1 - p_3)(V_{1,2} - V_{2,2}) + (1 - p_3)^2(V_{1,1} - V_{2,1}) \right) + 1/2 \left( p_2^3(V_{1,4} - V_{2,4}) + 2p_3(1 - p_3)(V_{1,3} - V_{2,3}) + (1 - p_3)^2(V_{1,2} - V_{2,2}) \right) + 1/2 \left( p_2^2(V_{1,3} - V_{2,3}) + 2p_3(1 - p_3)(V_{1,2} - V_{2,2}) + (1 - p_3)^2(V_{1,1} - V_{2,1}) \right)$.

The expressions (F-1) and (F-2) above are for network $G_h$, and can be reproduced for $G_l$ in a similar way. From (F-1) and (F-2), one can construct the systems of inequalities for $G_h$ and $G_l$, the solution to which will lead to equilibria. For example, suppose the low degree type is indifferent while the high type prefers free riding, then the system is as follows:
\[
\begin{align*}
(U(1, \sigma, 2) - U(0, \sigma, 2) &= 0 \\
(U(1, \sigma, 3) - U(0, \sigma, 3) &< 0, \\
0 < p_2 < 1, p_3 = 0
\end{align*}
\] (F-3)

Solving these systems gives us the following Bayesian Nash equilibria (consistent with Table 2-4):

Table F-2. Bayesian Nash equilibria for incomplete information treatments

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Equilibria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_2$</td>
</tr>
<tr>
<td>$Hden. I$</td>
<td>1</td>
</tr>
<tr>
<td>$Lden. I$</td>
<td>1</td>
</tr>
</tbody>
</table>

F.2. Complete information

Figure F-1 (duplicated from Figure 2-1). Network structures used in the experiment
When the network information is complete, recall that we focus on Nash equilibria assigning the same action to symmetric network positions. To be specific, the equilibria in network $G_h$ should satisfy:

$$p_a = p_h, p_b = p_g, p_c = p_f, p_d = p_e,$$  \hspace{1cm} (F-4)

and the corresponding conditions for $G_l$ are

$$p_a = p_f, p_b = p_e, p_c = p_d, p_g = p_h.$$  \hspace{1cm} (F-5)

In the complete information setting, let $U_i(x_i) = \sum_{x_{N_i}} v_k(x_i; x_{N_i}) \phi'(x_{N_i}, i)$, where $\phi'(x_{N_i}, i)$ is the probability that the actions of network position $i$’s $(i \in \{a, b, ..., h\})$ neighbors are $x_{N_i}$. $U_i(x_i)$ represents the expected payoff of the player at position $i$ when she takes action $x_i$.

For $G_h$ we have

$$U_a(1) - U_a(0) = p_h p_f (V_{1,4} - V_{2,4}) + \left(1 - p_h\right) p_f \left(1 - p_f\right) V_{1,3} - V_{2,3} + (1 - p_h)(1 - p_f) V_{1,2} - V_{2,2} + (1 - p_h)(1 - p_f) V_{1,1} - V_{2,1}.$$

For $G_l$ we have

$$U_c(1) - U_c(0) = p_h p_d (V_{1,4} - V_{2,4}) + \left(1 - p_h\right) p_d \left(1 - p_d\right) V_{1,3} - V_{2,3} + (1 - p_h)(1 - p_d) V_{1,2} - V_{2,2} + (1 - p_h)(1 - p_d) V_{1,1} - V_{2,1}.$$

Likewise, $U_i(1) - U_i(0)$ can be calculated for other position $i \in \{b, d, e, f, g, h\}$. Based on that, we now show an example to compute the equilibria for the network $G_h$. Consider the case in which the high degree nodes $(a, c, f, h)$ free ride while the low degree ones (positions $b, d, e, g$) being indifferent. In this case we have the following inequalities system

$$\begin{align*}
U_i(1) - U_i(0) &< 0 \\
U_j(1) - U_j(0) &= 0 \\
0 < p_j < 1, p_i &= 0
\end{align*}
\quad \forall i \in \{a, c, f, h\}, j \in \{b, d, e, g\}$$  \hspace{1cm} (F-8)

Solving simultaneously (F-4) and (F-8) yields the equilibria for the case considered above. One can write down the systems to solve for network $G_l$, in a similar fashion. Altogether, this gives the full set of equilibria for complete information treatments (consistent with Table 2-4):
Table F-3. Symmetric Nash equilibria for complete information treatments

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Equilibria</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{den_C}$</td>
<td>$p_a\cdot p_h$ 0.6 1 0 1</td>
</tr>
<tr>
<td>$L_{den_C}$</td>
<td>$p_a\cdot p_f$ $p_{b\cdot p_e}$ $p_{c\cdot p_d}$ $p_{g\cdot p_h}$ $\frac{4-q}{1+4q}$ 0 1 $q \in \left[\frac{3}{5},1\right]$</td>
</tr>
</tbody>
</table>

Appendix G.

Experimental Instruction for Chapter 2

We present the experimental instructions for treatments $H_{den\_C}$ and $H_{den\_I}$ in this appendix. The instructions can be applied to low density treatments with only adjustments of the neighbor degree distribution information or the picture of network.

General. Welcome and thank you for participating in this experiment. In this experiment you will earn money. From now on until the end of the experiment, please do not communicate with other participants. If you have any question, please raise your hand. An experimenter will come to your place and answer your question privately.

The Game. In the Experiment we use ECU (Experimental Currency Unit) as the monetary unit. The profits you make during the experiment will be added to this account in ECU. At the end of the experiment, the balance of the account will be converted from ECUs into dollars according to the conversion rate stated below, and paid out in cash after the experiment.

The experiment lasts for 20 rounds. In each round, participants will be organized in a network shown in Figure 1 below. In this network, the people connected to you are your neighbors. Every round your neighbors will be different people. Your location in the network is determined by an ID number ranging from 1 to 8 (as marked on Figure 1). A new ID number will be assigned to you prior to each round of experiment.
Figure 1. The network

At the beginning of each round, each person in the network chooses one of the two options: A, B. Each does so without any knowledge of what any other person decides. The profit you earn depends on your location in the network, the option you choose, and how many of your neighbors choose A. Specifically,

*If you choose A and you have two neighbors then*
- If both neighbors choose A, your profit is 225 ECU.
- If exactly one neighbor chooses A your profit is 175 ECU.
- If neither neighbor chooses A your profit is 100 ECU.

*If you choose B and you have two neighbors then*
- If both neighbors choose A, your profit is 275 ECU.
- If exactly one neighbor chooses A your profit is 100 ECU.
- If neither neighbor chooses A your profit is 0 ECU.

This is summarized in Table 1 as below, also shown on your computer screen during your play.

<table>
<thead>
<tr>
<th>Your Profit</th>
<th>Number of Your Neighbors who Choose A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Your</td>
<td>A</td>
</tr>
<tr>
<td>Choice</td>
<td>B</td>
</tr>
</tbody>
</table>

Table 1. Two-Neighbor Payoff Table

*If you choose A and you have three neighbors, then*
- If all three neighbors choose A your profit is 260 ECU.
- If exactly two neighbors choose A, your profit is 225 ECU.
• If exactly one neighbor chooses A your profit is 175 ECU.
• If no neighbor chooses A your profit is 100 ECU.

If you choose B and you have three neighbors, then
• If all three neighbors choose A your profit is 335 ECU.
• If exactly two neighbors choose A, your profit is 275 ECU.
• If exactly one neighbor chooses A your profit is 100 ECU.
• If no neighbor chooses A your profit is 0 ECU.

This is summarized in Table 2 as below, also shown on your computer screen during your play.

<table>
<thead>
<tr>
<th>Your Profit</th>
<th>Number of Your Neighbors who Choose A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Your Choice</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>B</td>
</tr>
</tbody>
</table>

Table 2. Three-Neighbor Payoff Table

There are a few tips to keep in mind, which will help you earn more profits:
• I earn more profit by choosing B if two or more than two of my neighbors choose A.
• I earn more profit by choosing A if none or one of my neighbors chooses A.

The conversion rate is 1ECU=0.002 dollars. Before the game starts, you are required to complete a quiz, which covers the important knowledge about the game. The quiz will be shown on your computer screen. You will start to play the game only after you correctly answer all the questions in the quiz.

Consent Forms. Please read the consent form that is delivered to you before the start of the experiment.

********************************************************* Instruction: Hden_I*********************************************************
the experiment, the balance of the account will be converted from ECU into dollars according to the conversion rate stated below, and paid out in cash after the experiment.

The experiment lasts for 20 rounds. In each round, participants will be organized in a network. In this network, you are connected to either two or three people, who are your neighbors. Every round your neighbors will be different people. You know the number of neighbors whom you are connected, but you will not know their identity. Specifically,

If you have two neighbors,
- With 1/2 chance, One of your neighbors has 2 neighbors (including yourself) while the other has 3 neighbors (including yourself).
- With 1/2 chance, each of your neighbors has 3 neighbors (including yourself).

If you have three neighbors,
- With 1/2 chance, One of your neighbors has 2 neighbors (including yourself) while the other two neighbors have 3 neighbors each (including yourself).
- With 1/2 chance, One of your neighbors has 3 neighbors (including yourself) while the other two neighbors have 2 neighbors each (including yourself).

At the beginning of each round, each person in the network chooses one of the two options: A, B. Each does so without any knowledge of what any other person decides. The profit you earn depends on how many neighbors you have, the option you choose, and how many of your neighbors choose A. Specifically,

If you choose A and you have two neighbors then
- If both neighbors choose A, your profit is 225 ECU.
- If exactly one neighbor chooses A your profit is 175 ECU.
- If neither neighbor chooses A your profit is 100 ECU.

If you choose B and you have two neighbors then
- If both neighbors choose A, your profit is 275 ECU.
- If exactly one neighbor chooses A your profit is 100 ECU.
- If neither neighbor chooses A your profit is 0 ECU.

This is summarized in Table 1 as below, also shown on your computer screen during your play.
If you choose A and you have three neighbors then

- If all three neighbors choose A your profit is 260 ECU.
- If exactly two neighbors choose A, your profit is 225 ECU.
- If exactly one neighbor chooses A your profit is 175 ECU.
- If no neighbor chooses A your profit is 100 ECU.

If you choose B and you have three neighbors then

- If all three neighbors choose A your profit is 335 ECU.
- If exactly two neighbors choose A, your profit is 275 ECU.
- If exactly one neighbor chooses A your profit is 100 ECU.
- If no neighbor chooses A your profit is 0 ECU.

This is summarized in Table 2 as below, also shown on your computer screen during your play.

<table>
<thead>
<tr>
<th>Your Profit</th>
<th>Number of Your Neighbors who Choose A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Your Choice</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>B</td>
</tr>
</tbody>
</table>

Table 2. Three-Neighbor Payoff Table

There are a few tips to keep in mind, which will help you earn more profits:

- I earn more profit by choosing B if two or more than two of my neighbors choose A.
- I earn more profit by choosing A if none or one of my neighbor chooses A.

The conversion rate is 1ECU=0.002 dollars. Before the game starts, you are required to complete a quiz, which covers the important knowledge about the game. The quiz will be shown on your computer screen. You will start to play the game only after you correctly answer all the questions in the quiz.
**Consent Forms.** Please read the consent form that is delivered to you before the start of the experiment.

*********************************************************************************

Appendix H.

**The Experimental Software Interface for Chapter 2**

This section provides snapshots of the experimental software. The software is programmed with zTree (Fischbacher, 2007). Subjects begin with a quiz testing their understanding of the game, with no earning accumulated to the game. The quiz screen is shown in Figure H-1. The actual decision making interfaces are found in Figure H-2. Figure H-3 is the feedback screen displayed after one’s decision. For better references, the interfaces for complete information game and incomplete information game are separately presented. For the complete (incomplete) information games, we respectively present the screen faced by a high (low) degree type in a high (low) density network. The other cases not covered in the screenshots are presented to the subjects in an analogous way.
### Quiz (non-paying period)

**Number of Your Neighbors who choose A**

<table>
<thead>
<tr>
<th>Choose A</th>
<th>Three</th>
<th>Two</th>
<th>One</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your Choice</td>
<td>260</td>
<td>225</td>
<td>175</td>
<td>100</td>
</tr>
<tr>
<td>Choose B</td>
<td>235</td>
<td>275</td>
<td>130</td>
<td>0</td>
</tr>
</tbody>
</table>

1. Suppose one of your neighbors chooses A and the others do not. Then your profit from choosing A is

2. You have the same number of neighbors as any other does.

3. Your neighbors change in every period.

4. What is your profit when you choose B while two of your neighbors choose A?

5. I earn more profit by choosing B if two or more than two of my neighbors choose A.

6. I earn more profit by choosing A if none or one of my neighbor chooses A.

---

Figure H-1. Quiz
You are Player No. 8. The network you play in is shown at the right, with your location highlighted.

Number of Your Neighbors who choose A

<table>
<thead>
<tr>
<th>Three</th>
<th>Two</th>
<th>One</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>280</td>
<td>225</td>
<td>175</td>
<td>100</td>
</tr>
</tbody>
</table>

Choose A

Your Choice

Choose B

335   275   100   0

Choose your action for this period:

- A
- B

Period | Your choice | Your Profit | Your first neighbor’s choice | Your second neighbor’s choice | Your third neighbor’s choice | Your Total Profit
---|-------------|-------------|-----------------------------|-----------------------------|-----------------------------|---------------------
1      | Choice B   | 275         | Choice A                   | Choice A                   | Choice B                   | 275                
You have 2 neighbors.
With 2/3 chance, one of your neighbors has 2 neighbors (including yourself) while the other has 3 neighbors (including yourself).
With 1/3 chance, each of your neighbors has 2 neighbors (including yourself).

<table>
<thead>
<tr>
<th>Number of Your Neighbors who Choose A</th>
<th>Two</th>
<th>One</th>
<th>Zero</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose A</td>
<td>225</td>
<td>175</td>
<td>100</td>
</tr>
<tr>
<td>Choose B</td>
<td>275</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

Choose your action for this period: 
A
B

<table>
<thead>
<tr>
<th>Period</th>
<th>Your choice</th>
<th>Your Profit</th>
<th>Your first neighbor's choice</th>
<th>Your second neighbor's choice</th>
<th>Your third neighbor's choice</th>
<th>Your Total Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Choose B</td>
<td>275</td>
<td>Choose A</td>
<td>Choose B</td>
<td>Choose A</td>
<td>275</td>
</tr>
</tbody>
</table>

Figure H-2. Decision making
<table>
<thead>
<tr>
<th>Player</th>
<th>You</th>
<th>First Neighbor</th>
<th>Second Neighbor</th>
<th>Third Neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player’s Choice</td>
<td>Choose B</td>
<td>Choose A</td>
<td>Choose A</td>
<td>Choose A</td>
</tr>
</tbody>
</table>
I.1. Proof of Lemma 3-1

Given any fixed number of citations, it is the player’s dominant strategy to place these citations first to largest artifacts, and then to second largest ones, … until the given number of citations is exhausted. A player not executing this order can strictly gain by switching to this strategy regardless of others’ strategy, because by doing so she obtains more knowledge with the same citation cost. The issue yet to address is what marks the point that the citations should
optimally cease. Notice by setup, $f(x_i^*(y_i)) - c(y_i)x_i^*(y_i)$ is concave (in a discrete analogy, i.e. exhibiting decreasing increment) in the amount of cited knowledge, $y_i$. Therefore, $\pi_i^*(s_i, y_i)$ exhibits piecewise concavity in $y_i$ in any range where the optimal knowledge production is inelastic to $y_i$. When $x_i^*(y_i)$ is being shifted by $y_i$ to a higher level say $t$ (the cited knowledge amount corresponding to $m_{it}(s_i)$), there will be a jump in $\pi_i^*(s_i, y_i)$ that breaks concavity due to the instantaneous change in the value of $D_i(s_i, x_i)$. Therefore, citation should optimally stop at either a $m_{it}(s_i)$ bound, or at an interior point within one of the concavity ranges where the benefit of citing one additional artifact equals to its cost. This is how $m_i^j(t)$ is defined in Lemma 3-1.

I.2. Proof of Lemma 3-2

The relationship $m_{ij}(s_i) < m_{i,j+1}(s_i)$ is easily determined from the inverse function of $x_i^*(y_i)$, which increases in $y_i$ in our setup. Since $\pi_i^*(s_i, y_i)$ increases in $y_i$, we have $\pi_i^*(s_i, m_i^j(s_i) + (j + 1)) - \pi_i^*(s_i, m_i^j(s_i)) > \pi_i^*(s_i, m_i^j(s_i) + j) - \pi_i^*(s_i, m_i^j(s_i)) = k$. Thus the player has still capacity to cite $(j + 1)$-artifacts at $m_i^j(s_i)$. This fact joint with the decreasing increment property (concavity) of $\pi_i^*(s_i, y_i)$ in $y_i$ implies $m_i^j(s_i) < m_i^{j+1}(s_i)$.

I.3. Proof of Lemma 3-3

We can write $\pi_i(s_i; x + 1, y_i) - \pi_i(s_i; x, y_i) = f(x + 1) - f(x) - c(y_i) + D_i(s_i, x + 1) - D_i(s_i, x)$, and $\pi_i(s_i; x + 1, y_i) - \pi_i(s_i; x, y_i)$ therefore increases in $y_i$. Notice that $D_i(s_i, x + 1) - D_i(s_i, x) > D_i + 1(s_i + 1, x + 1) - D_i + 1(s_i + 1, x)$ $\forall x_i$ implies $\pi_{i+1}(s_{i+1}; j, m_{ij}(s_i)) - \pi_{i+1}(s_{i+1}; j - 1, m_{ij}(s_i)) < \pi_i(s_i; j, m_{ij}(s_i)) - \pi_i(s_i; j - 1, m_{ij}(s_i))$, while the RHS is close enough to 0 by the definition of $m_{ij}(s_i)$ and sufficiently small knowledge units. As a result, $\pi_{i+1}(s_{i+1}; j, m_{ij}(s_i)) - \pi_{i+1}(s_{i+1}; j - 1, m_{ij}(s_i)) < 0$. For player $i + 1$ to optimally switch from producing $j - 1$-artifacts to producing $j$-artifacts, she has to cite more than $m_{ij}(s_i)$. Therefore, $m_{ij}(s_i) < m_{i+1,j}(s_{i+1})\forall x_i$. 

98
Next we establish the conditions for \( m^j_i(s_i) > m^j_{i+1}(s_{i+1}) \forall x_i \). Notice the inequality \( m_{ij}(s_i) < m_{i+1,j}(s_{i+1}) \forall x_i \) indicates that \( x^*_i(y) > x^*_{i+1}(y) \) for given \( y \). So
\[
\frac{\partial \pi^*_i(s,y)}{\partial y} = -c'(y)x^*_i(y) > -c'(y)x^*_{i+1}(y) = \frac{\partial \pi^*_{i+1}(s_{i+1},y)}{\partial y} > 0.
\]
Hence for \( y \) such that \( y > m^j_i(s_i) \) and \( \pi^*_i(s_i, y + j) - \pi^*_i(s_i, y) = k \), we have \( \pi^*_i(s_i + 1, y) - \pi^*_{i+1}(s_{i+1}, m^j_i(s_i)) < \pi^*_i(s_i, y) - \pi^*_i(s_i, m^j_i(s_i)) \). Notice by the definitions of \( y \) and \( m^j_i(s_i) \), \( \pi^*_i(s_i, y) - \pi^*_i(s_i, m^j_i(s_i)) < \) the associated citation cost difference. Therefore, \( \pi^*_{i+1}(s_{i+1}, y) - \pi^*_{i+1}(s_{i+1} + 1, m^j_i(s_i)) < \) the associated citation cost difference. So it cannot be the case that player \( i + 1 \) has a larger capacity for citation than player \( i \) does. \( m^j_i(s_i) > m^j_{i+1}(s_{i+1}) \forall x_i \).

---

\(^42\) Note that by definition \( y \) does not coincide with \( m^j_i(s_i) \), since \( \pi^*_i(s_i, y) \) is piecewise concave but not globally concave. So it may well be that \( y \) and \( m^j_i(s_i) \) are respectively local and global optimums for player \( i \).
Vita
Yang Zhang

EDUCATION
Ph.D. The Pennsylvania State University, Business Administration & Operations Research (Dual title), Dec 2013, GPA: 3.91/4
Bachelor. Industrial Engineering, Xi’an Jiaotong University (China), 2007

SELECTED PAPERS
Zhang, Y., G. E. Bolton. “Network Effects on Coordination: A Laboratory Investigation”
Zhang, Y. “The Emergence of Core-Periphery Structures in Information Networks”
Refereed Proceedings, INFORMS Conference on Information Systems and Technology (CIST) 2011

SELECTED RESEARCH PRESENTATIONS
“Network Effects on Coordination and Technology Adoption”
Sponsored Session, INFORMS 2012 Phoenix.
Sponsored Session, DSI 2012 San Francisco.
BOM (Behavioral Operations Management) 2012 Washington DC.
Invited Session, POMS 2012 Chicago.
Sponsored Session, INFORMS 2010 Austin.

“Innovation, Citation, and the Emergence of Knowledge Core”
INFORMS Conference on Information Systems and Technology (CIST) 2011 Charlotte.
Operations Research Colloquium Fall 2011, Penn State.

“Network, Information, and Free Riding”
Invited Session, INFORMS 2013 Minneapolis (forthcoming).
Invited Seminar Speaker, Fall 2012, University of Texas at Dallas.

SELECTED AWARDS
National Science Foundation Grant for Doctoral Dissertation,
Decision, Risk & Management Sciences Program, $8,700, 2011-2013
Summer Research Stipend, Smeal College
The Pennsylvania State University, $2,000, 2012
Doctoral Research Grant, Smeal College
The Pennsylvania State University, $1,000, 2010