# ECONOMIC ACTIVITIES AND NETWORKS OF RELATIONSHIPS

by

**Yong-Ju Lee**

B.A. in Economics, SungKyunKwan University, 2000

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

**Doctor of Philosophy**

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Yong-Ju Lee

It was defended on

April 9th, 2007

and approved by

John Duffy, Professor, Department of Economics, University of Pittsburgh

Ted Temzelides, Associate Professor, Department of Economics, University of Pittsburgh

Utku Unver, Assistant Professor, Department of Economics, University of Pittsburgh

Esther Gal-Or, Professor, Katz Graduate School of Business, University of Pittsburgh

Dissertation Director: John Duffy, Professor, Department of Economics, University of Pittsburgh

**ECONOMIC ACTIVITIES AND NETWORKS OF RELATIONSHIPS**

Yong-Ju Lee, PhD

University of Pittsburgh, 2007

The fundamental question I address in the dissertation is how the behavior of economic agents interacts with networks of relationships which underlie a wide set of economic situations. In Ch. 2, entitled "Decentralized Information Sharing in Oligopoly," I analyze the incentives of firms for information sharing in a decentralized environment when firms face a stochastic demand. In order to do that, I develop a two stage model of strategic network formation, where a cooperative network formation stage is played in the first stage and a noncooperative Bayesian Cournot game is played in the second stage. I derive pure strategy mixed cooperative and noncooperative equilibria that are subgame perfect and stable, and characterize the resulting network structures. Ch. 3, entitled "A War of Attrition in Network Formation," investigates the strategic behavior of agents when they face a decision on the formation of relationships. I apply a war of attrition to the dynamic network formation process when links among agents have characteristics of public goods. Agents are randomly but exogenously matched in each stage. Based on Bala and Goyal's (2000) two-way flow model, I characterize the subgame-perfect equilibrium outcomes and discuss their efficiency. Finally, Ch. 4, entitled "Social Norms and Trust among Strangers," (with Huan Xie) studies the development of trust and reciprocity among strangers in the indefinitely repeated trust game with random matching. If reputation is attached to the community as a whole and if a single defection leads to the destruction of the cooperative social norm through contagious punishments, the cooperative social norm can be sustained by the self-interested community members in the equilibrium. We provide sufficient conditions that support the social norm of trust and reciprocity as a sequential equilibrium.

## TABLE OF CONTENTS

## LIST OF FIGURES

# 1.0 INTRODUCTION

Networks of relationships play a critical role in a wide set of economic and social situations. For instance, personal contacts play important roles in obtaining information about job opportunities. Such networks of relationships also underlie the trade and exchange of goods in non-centralized markets, the provision of mutual insurance in developing countries, R&D and collusive alliances among corporations, and international alliances and trade agreements. Given the prevalence and importance of network structures, the literature on the formation of networks among agents addresses various questions: Some examples are how such network structures are important in determining the outcome of economic interaction, how we predict which networks are likely to form when agents have the strategic discretion to choose their connections, and how efficient the networks are.

The fundamental question I address in this dissertation is how the behavior of economic agents interacts with networks of relationships which underlie a wide set of economic situations. In order to do that, I categorize the network structures into the purely endogenous networks, the purely exogenous networks, and mixed endogenous and exogenous networks, and then analyze the behavior of individual agents, oligopolistic firms, and the community as a whole.

The purpose of the first paper (Ch. 2) is to investigate firms' incentives to form a network in order to share information in an oligopolistic market where firms face an uncertain demand. Many papers have analyzed the existence of incentives to share private information in stochastic market environments. (See, Clarke (1983), Gal-Or (1985, 1986), Novshek and Sonnenschein (1982), Raith (1996), and Vives (1984).) These papers show that it is unclear whether the exchange of information about an uncertain world has a profitable effect on the firms when they compete against each other as Nash competitors in the product market.

This is because the results crucially depend on the market's random variables, their distribution, the nature of competition (Cournot or Bertrand), and the nature of information. On the other hand, it is well known that, with an unknown common demand, information sharing is the unique Nash equilibrium outcome under Bertrand competition and concealing information is the unique equilibrium outcome under Cournot competition (Vives (1984) and Gal-Or (1985)). And, with unknown private costs, information sharing is the unique Nash equilibrium outcome under Cournot competition and concealing information is the unique equilibrium outcome under Bertrand competition (Gal-Or (1986)). Here we show that the results on information sharing vary depending on whether the decision is made in a centralized or decentralized environment.

The second paper (Ch. 3) investigates the strategic behavior of agents when they face a decision on the formation of relationships. Many social and economic relationships that benefit both the corresponding parties are efficiently established by the effort of a single initiative party. Inviting new neighbors for a dinner and links between politicians and businessmen can be examples. Under this consideration, we observe that links among economic agents in Bala and Goyal's (2000) two-way flow model have such properties of public goods among linked agents. Once relationships are established by someone, others can use this network of relationships freely as long as they are parts of the given network. However, the costs of link formation are incurred only by the agents who initiate the links. So each agent strongly wants to wait for others to initiate a link to him. This kind of a waiting situation allows us to analyze network formation in the framework of a war of attrition where the strategy of an agent is the specification of the set of waiting times among agents with whom he plays the game. The main goal of this paper is to understand how and which network structures emerge as equilibrium outcomes when links among agents have such properties of a public good and all agents want to wait. Also we investigate who volunteers to initiate the link and pays for link formation in the network formation process, and who would be a center if the star network is formed.

Finally, my third paper (Ch. 4) asks how to sustain cooperative equilibrium when players have an incentive to deviate from cooperation, since they are completely anonymous and they meet at randomly determined times. Economists have long recognized "reputation" as an effective means of enforcing cooperative behavior, but these personal enforcement mechanisms are effective only if quick and substantial retaliations are available. However, many important transactions in reality are infrequent in nature and many transactions happen among essentially anonymous players. Electronic transactions are done between strangers who have no contact except through cyberspace. In this case only partial information about a stranger's reputation is available at best, and, therefore, the effectiveness of reputation is far less certain. This observation raises important question about economic behavior. What factors drive the emergence of trust and reciprocity in economic transactions? In order to analyze this question, previous papers consider a random matching model under the most extreme information restriction. We theoretically extend Kandori's (1992) arguments to the trust game.

## 2.0 DECENTRALIZED INFORMATION SHARING IN OLIGOPOLOY

### 2.1 MOTIVATION

The purpose of this paper is to investigate firms' incentives to form a network in order to share information in an oligopolistic market where firms face an uncertain demand. Many papers analyzed the existence of incentives to share private information in stochastic market environments. (See, Clarke (1983), Gal-Or (1985, 1986), Novshek and Sonnenschein (1982), Raith (1996), and Vives (1984).) These papers show that it is unclear whether the exchange of information about an uncertain world has a profitable effect on the firms when they compete against each other as Nash competitors in the product market. This is because the results crucially depend on the market's random variables, their distribution, the nature of competition (Cournot or Bertrand), and the nature of information. On the other hand, it is well known that, with an unknown common demand, information sharing is the unique Nash equilibrium outcome under Bertrand competition and concealing information is the unique equilibrium outcome under Cournot competition (Vives (1984) and Gal-Or (1985)). And, with unknown private costs, information sharing is the unique Nash equilibrium outcome under Cournot competition and concealing information is the unique equilibrium outcome under Bertrand competition (Gal-Or (1986)). Here we show that the results on information sharing vary depending on whether the decision is made in a centralized or decentralized environment.

Network structures play a significant role in determining the outcome of many important economic relationships.[1] Although the previous papers analyze the possibility of collabora-

[1] For instance, personal contacts play important roles in obtaining information about job opportunities. Such networks of relationships also underlie the trade and exchange of goods in non-centralized markets, the

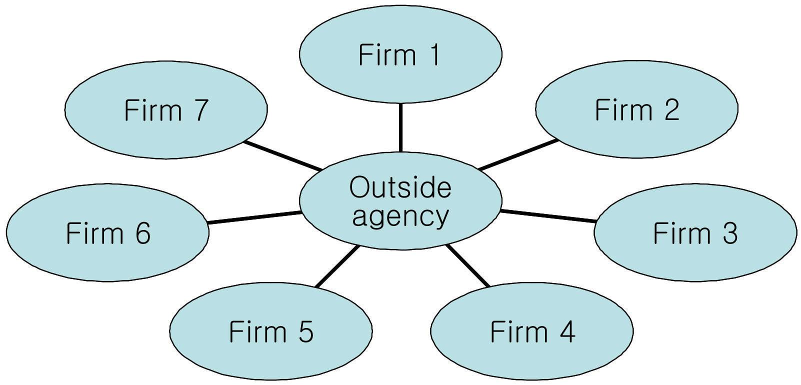


Figure 1: Centralized Information Sharing

tion or cooperation among firms, they impose severe restrictions on the interaction structure among the firms. The literature explicitly or implicitly assumes the existence of an "outside agency" such as a trade association. This agency collects private information from firms and disseminates it throughout the industry. In the terminology of the network literature, they assume that the network structure is a *star network* where the central node is the trade association and the periphery nodes are the firms. This is equivalent to assuming the *complete network* among the firms without an outside agency.[2] This interpretation is without loss of generality since the literature has not treated the outside agency as a player in the game, and, instead, the outside agency is modeled as a purely mechanical part of the environment, with no decisions to make. Therefore, once a firm decides to reveal its information on its private signal, the firm has to share it with all the other firms at the same time. In this sense, the previous literature assumes that each firm faces an *exogenously* given *centralized* market structure and makes an *industry-wide* decision.

In this paper, we reexamine the incentives for information sharing in a decentralized setting. For this purpose, we endogenize the network formation process explicitly in the model. Each firm selects with which firms to collaborate and share information with the understanding that firms engage in noncooperative competition in the product market. And,

provision of mutual insurance in developing countries, R&D and collusive alliances among corporations, and international alliances and trade agreements. (Refer to Jackson (2004) for an excellent survey.)

[2] By the complete network we mean the graph in which each firm has a direct link with every other firm.

only linked firms can share information bilaterally. That is, the present paper studies the information sharing problem in oligopoly where each firm faces a *decentralized* market structure and engages in a *pair-wise* desision.[3] Collaboration and cooperation among firms are common in oligopolistic markets (Goyal and Joshi (2003)). As Goyal and Moraga-Gonzalez (2001) indicate, a distinctive feature of collaboration or cooperation among firms is that they are often bilateral interactions that are embedded in a broader network. Even in situations where firms $i$ and $j$, and $j$ and $k$ have a cooperative relationship, respectively, firms $i$ and $k$ may not have such a relationship. These structural features justify incorporating strategic network formation in the existing information sharing literature.

Another innovative aspect of our model resides in the choice of the degree of information sharing. In much of the previous literature, firms decide the degree of information sharing by choosing the level of variance of a message (which is a random variable). Given that the network structure is complete as mentioned above, firms share information uniformly with all the other firms at the same time. For example, they reveal information completely if the variance in messages is zero, while they don't share information if the variance is infinity. In our model, the firms choose the degree of information sharing by selecting the set of firms with which they want to form collaborative links. It is assumed that a link involves a commitment on bilateral and truthful information sharing between the linked firms. The firms share information completely if the resulting graph is a complete one, while they don't share information if the resulting graph is an empty one.[4] Therefore, we can say that the previous literature measures the degree of information sharing by "*depth*," while the present paper measures it by "*width*." In this sense, we might view the current paper as complementary to the existing literature on information sharing.

We consider a simple two stage game. In the first stage, firms strategically form pair-wise links which allow them to obtain the others' private signals regarding the stochastic market demand. We assume that only the directly linked firms can share information bilaterally,

[3] We argue that pair-wise information sharing is more natural and realistic than any other forms in the world. First, pair-wise information sharing captures informal (or, private) and local behaviors among firms in the market. Second, it seems to be necessary at least as a pre-step to another forms of coalitions such as the cost reduction alliance, joint venture formation, and R&D agreement. Before those decisions are made, firms need to know and have an incentive for communicating about their market situations. The author thanks Esther Gal-Or for pointing this out.

[4] By the empty network we mean the graph in which there is no link between any two firms.

and that they share information truthfully if there is a collaborative link between them. After the link formation, each firm observes its own private signal and they transmit their private information to the linked firms simultaneously. In the second stage, after information transmission, each firm chooses its level of output in the product market. That is, a Bayesian Cournot game is played in the second stage. Following the spirit of d'Aspremont and Jacquemin (1988), we derive pure strategy mixed cooperative and noncooperative equilibria that are subgame perfect and (pairwise and strongly) stable, and characterize the resulting graphs. That is, given a collaborative network structure, firms compete with each other as Nash competitors in the product market in order to maximize their own profits, and firms cooperate bilaterally in the first stage so as to overcome market uncertainty and, hence, to maximize their own profits with the understanding that they engage in noncooperative competition in the second stage. The main questions we address are what is the incentive of firms to collaborate and what is the resulting network structure, what are the effects of strategic network formation on individual and industry-wide performance, and why does a firm's incentive to share information depend on whether the network structure is centralized or decentralized.

We start by characterizing the unique equilibrium decision rule of firms in the product market. We show that the results of the existing literature easily extend to endogenous network settings. That is, we show that the equilibrium decision rule is affine in the vector of observations available to the firm. We then examine the relationship between network structure and strategic information exchange. We apply both *pairwise stability* and *strong stability* as solution concepts to the network formation game. We demonstrate that, in contrast to the results of the existing literature on centralized information sharing, complete information sharing, no information sharing, and partial (and asymmetric) information sharing emerge as pairwise stable equilibria when firms face the decentralized market structure and make pair-wise decisions. Also we show that this result extends to the asymmetric environment. This result is interesting in that pairwise stability is a natural solution concept in the analysis of the decentralized oligopolistic market, since, as Roth and Sotomayor (1990, p.156) argue, "identifying and organizing large coalitions may be more difficult than making private arrangements between two parties." However, if we allow for a broader level of coordination and cooperation among the groups of the firms, no information sharing emerges as the unique equilibrium outcome. This illustrates that information sharing among firms facing an uncertatin environment critically depends on the coalition structure. Finally, by introducing heterogeneity across firms, we can clarify the features of decentralized information sharing: In any equilibrium, information is shared among firms with a similar accuracy (amount) of information. Otherwise, a firm with inaccurate (less) information refuses to collaborate.

Our paper can be seen as a contribution to the study of group formation and cooperation in oligopolies. Modeling strategic network formation is inspired by Bala and Goyal (2000), Dutta et al (1995), and Jackson and Wolinsky (1996). This paper attempts to study the relationship between strategic networks and stochastic market environments. Recently several papers have addressed similar issues to the ones we address in the present paper. Some examples are Goyal and Moraga-Gonzalez (2001), Goyal and Joshi (2003), and Goyal et al (2004). These papers highlight the relationship between the firms' incentives for R&D and network formation. However, here we study the relationship between the incentives for strategic information exchange in an uncertain market and network formation. This paper, to the best of our knowledge, is the first to analyze information sharing among oligopoly firms in the context of decentralized network formation. Interestingly, the R&D literature and the information sharing literature frequently study very different incentives for network formation even in exactly the same environments. For example, under a homogeneous product oligopoly assumption and a small link formation cost, Goyal and Joshi (2003) and Goyal and Moraga-Gonzalez (2001) show that the complete network is a stable network under Cournot competition. Our paper has a very different prediction.

The remainder is organized as follows. Section 2 studies the model, which consists of the cooperative network formation game and the noncooperative oligopoly game with a stochastic common demand. In section 3 we derive equilibria of the game and characterize equillibrium network structures. Section 4 introduces heterogeneity to the basic model by assuming that the accuracy of private information is different across firms. Section 5 discusses possible extensions and future research. Finally, section 6 concludes.

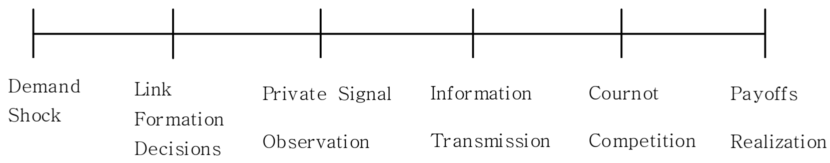


Figure 2: Timing of the game

## 2.2 THE MODEL

In this section we establish the benchmark model where the game involves a symmetric environment and firms produce homogeneous goods. Section 4 extends the basic model by introducing heterogeneity across firms.

We consider a two stage game. The timing of the game is shown in Figure 2. In the first stage, firms strategically form pair-wise links to obtain information about the stochastic market demand. After link formation each firm observes its own private signal. Then firms simultaneously transmit their information to the linked firms. We call this stage game the *network formation game*. In the second stage, after information transmission, each firm chooses its level of output in the market. We call this second stage game the *oligopoly game*. Our goal is to derive the pure strategy mixed cooperative and noncooperative equilibria of the game. Firms coordinate bilaterally in the first stage so as to overcome market uncertainty and, hence, to maximize their own profits with the understanding that they engage in noncooperative competition in the second stage. We examine the incentives of firms for network formation and the resulting information sharing network structures, and analyze how completely firms collaborate. Finally, we analyze the effects of the network structure on both individual and market outcomes. We now develop the notation and define our notions of stability and efficiency.

### 2.2.1 The network formation game

Let $\mathrm{N} = \{1, 2, \cdots, n\}, n \geqq 3$ be the set of ex ante identical firms. In the first stage firms form pair-wise links which represent commitments that both firms must honour regarding bilateral information transmission.[5] For any pair of firms $i,\ j \in \mathrm{N}$, a pair-wise relationship between the two firms is represented by a binary variable $g_{ij} \in \{0, 1\}$. When $g_{ij} = 1$, this means that the two firms are linked at cost $\gamma$ respectively, while $g_{ij} = 0$ refers to the case of no link. In our model, $g_{ij} = 1$ also means that the two linked firms $i$ and $j$ share their information bilaterally. To avoid reaching a conlusion about collaborative network formation that critically depends on the link formation cost $\gamma$, we assume that the link formation cost $\gamma$ is negligibly small. The number of pairwise links represents the degree of information sharing among firms in the industry. A network $g$ is a collection of links, i.e., $g = \{g_{ij}\}_{i,j \in \mathrm{N}}$. Let $g_i = \{g_{ij} | g_{ij} = 1\}_{j \neq i}$ be the set of links involving firm $i$, where $g_{-i} = \{g_k\}_{k \neq i}$ is the set of all the firms except firm $i$. The set of all possible graphs on N is denoted by $G$. Let $\mathrm{N}_i(g) = \{j \in \mathrm{N} \backslash \{i\} | g_{ij} = 1\}$ be the set of firms with which firm $i$ has a link in $g$, and let $\eta_i(g)$ be the cardinality of the set $\mathrm{N}_i(g)$. Let $g - g_{ij}$ denote the network obtained by severing an existing link between firms $i$ and $j$ from network $g$, while $g + g_{ij}$ is the network obtained by adding a new link between firms $i$ and $j$ in network $g$. A path in $g$ connecting firms $i$ and $j$ is a set of distinct firms $\{i_1, i_2, \cdots, i_k\}$ such that $g_{ii_1} = g_{i_1 i_2} = \cdots = g_{i_k j} = 1$. We say that a network is connected if there exists a path between any pair $i, j \in \mathrm{N}$. A network, $g' \subset g$, is a *component* of $g$ if for all $i, j \in g'$, $i \neq j$, there exists a path in $g'$ connecting $i$ and $j$, and for all $i \in g'$ and $j \in g$, $g_{ij} = 1$ implies $g_{ij} \in g'$. The profits of firm $i$ in network $g$ are denoted by $\pi_i(g)$, which will be specified in the next subsection.

We shall say that a network $g$ is *pairwise stable* if and only if for all $i, j \in \mathrm{N}$ :

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

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

This definition of stability is taken from Jackson and Wolinsky (1996). These conditions indicate that agents need a bilateral (mutual) agreement to form a link, while agents can sever

[5] The present model has a different kind of commitment from the one in the previous literature. In the previous literature there is the commitment between the outside agency and each firm, while in this model commitment is between the pair of corresponding firms. Because of this commitment problem we choose the cooperative approach in modeling network formation.

an existing link unilaterally. To the extent that larger groups can coordinate their actions to make changes in a network, a stronger solution concept might be needed. Nevertheless, pairwise stability is a natural solution concept in our model, since, as Roth and Sotomayor (1990, p.156) argue, "identifying and organizing large coalitions may be more difficult than making private arrangements between two parties."

Alternatives to pairwise stability that allow for larger coalitions than just pairs of firms to deviate were first considered by Dutta and Mutuswami (1997). The following definition is modified from Jackson and van den Nouweland (2005).

A network $g' \in G$ is *obtainable from $g \in G$ via deviation by $S$* if

(i) $g'_{ij} = 1$ in $g'$ and $g_{ij} = 0$ in $g$ implies $i, j \in S$, and

(ii) $g_{ij} = 1$ in $g$ and $g'_{ij} = 0$ in $g'$ implies $\{i, j\} \cap S \neq \emptyset$.

The above definition identifies changes in a network that can be made by a coaltion $S$ without the need of consent of any firms outside of $S$. (i) requires that any new links that are added can only be between firms in $S$. This reflects the fact that consent of both firms is needed to add a link. (ii) requires that at least one firm of any deleted link be in $S$. This reflects the fact that either firm with a link can unilaterally sever the relationship.

A network $g$ is *strongly stable* if for any $S \subset \mathrm{N}$, $g'$ that is obtainable from $g$ via a deviation by $S$, and $i \in S$ such that $\pi_i(g') > \pi_i(g)$, there exists $j \in S$ such that $\pi_j(g') < \pi_j(g)$.

The definition of strong stability allows for a deviation to be valid if some firms are strictly better off and others are weakly better off, while the definition in Dutta and Mutuswami (1997) considers a deviation valid only if all firms of a coalition are strictly better off. This stronger notion implies pairwise stability. Strong stability provides a powerful refinement of pairwise stability. The concept of strong stability mainly makes sense in smaller network situations where agents have substantial information about the overall network structure and the potential payoffs.

We now define some networks that play aprominent role in our analysis. A network is said to be *symmetric* if every firm has the same number of links. Otherwise it is asymmetric. In a symmetric network $\eta_i(g) = \eta_j(g) = \Delta$ for any two firms $i$ and $j$. We will denote a symmetric network of degree $\Delta$ by $g^{\Delta}, \Delta = 0, 1, \cdots, n-1$. In particular, if $\Delta = 0$, the nework is called the *empty* network, while it is called the *complete* network if $\Delta = n - 1$.

It can be shown that if the number of firms is even, then there is always a symmetric network of degree $\Delta$, $\Delta = 0, 1, \cdots, n-1$ (Goyal and Moraga-Gonzalez (2001)). Among the asymmetric networks, the *dominant group architecture*, $g^{d(k)}$, is characterized by one complete non-singleton component with $k \geq 2$ firms and $n-k$ singleton firms. Thus, there is a set of firms $\mathrm{N}^d \subset \mathrm{N}$ with the property that $g_{ij} = 1$ for every pair $i, j \in \mathrm{N}^d$ while for any $m \in \mathrm{N} \backslash \mathrm{N}^d, g_{ml} = 0, \forall l \in \mathrm{N} \backslash \{m\}$ (Goyal and Joshi (2003)). A network is said to be *component-symmetric* if every firm in a component has the same number of links.

### 2.2.2 The oligopoly game

The oligopoly game is based on the existing information sharing literature, especially Gal-Or (1985, 1986).

The oligopoly consists of $n$ firms producing a product at no cost. We assume the same type of demand uncertainty as in Gal-Or (1985, 1986). The demand function facing this industry is linear and stochastic:

$$p = a - bQ + u, \qquad a, b > 0 \tag{2.1}$$

The prior distribution of $u$ is normal with mean zero and variance $\sigma_u$, $p$ is the price and $Q$ is the aggregate quantity produced. Before deciding its output quantity, each firm observes a noisy signal for $u$, and then transmits it to the linked firms. The signal observed by firm $i$ is $x_i$. We assume:

$$x_i = u_i + e_i, \qquad u_i \sim N\ (0, \sigma), \qquad e_i \sim N\ (0, m) \tag{2.2}$$

where $Cov(e_i, e_j) = 0, \quad i \neq j$; $Cov(u_i, e_j) = 0 \ \ \forall\ i, j$; $Cov(u_i, u_j) = 0, \quad i \neq j$; and $u = (\sum_i u_i)\ /\ n$, hence $u \sim N\ (0, \frac{\sigma}{n})$.[6] $e_i$ is called the signal error. The normal distribution assumption for these random variables, together with the linearity of demand function, enables us to derive explicit forms for their conditional and unconditional expected values. The private signals might be positively correlated, but here we simply assume that they are

[6] This environment is called a "common values" problem in the auction literature. (Gal-Or (1986))

independent.[7] Hence a firm cannot make any inference about the signals observed by the other firms based on its own signal. This fact gives the firms strong incentives for strategic link formation. Since, by assumption, every firm transmits its private signal to other linked firms simultaneously, the transmitted information (signals) from others cannot be used to generate a firm's own message. However, note that there are indirect network effects here. For instance, consider a situation where firms $i$ and $j$, and $j$ and $k$ have collaborative links to one another. We can obviously conclude that $x_k$ $(x_i)$ is unknown information to $i$ $(k)$, since there is no direct link between $i$ and $k$. But firm $i$ does know that firm $k$ will use information transmitted from $j$, $x_j$, (which is also known to $i$) to make the optimal decision in the product market in the second stage, and vice versa. That is, under rational expectations there is an indirect network effect even though there are no direct spillovers in the model.

We denote by $X$ the vector of true signals observed by all firms. $X_{-i}$ designates the vectors of true signals excluding those of firm $i$.

After information transmission, the transmitted information is subsequently used by each firm to choose its output. The output choice depends on the information available to the firm. For firm $i$ this information consists of network $g$, its private information $x_i$, the reported information from other firms $\{x_j|\ g_{ij} = 1\}_{j\neq i}$, and the known values of the parameters $m$ and $\sigma$. We denote the information available to $i$ by $y_i = (g, x_i, \{x_j|\ g_{ij} = 1\}_{j\neq i}, m, \sigma)$. For example, if $n = 5$ and we have a wheel (circle) network $g = \{g_{12} = g_{23} = g_{34} = g_{45} = g_{51} = 1\}$, then $y_1 = (g, x_1, (x_2, x_5), m, \sigma)$. The vector of information that is available to all firms is denoted by $y$, and the information available to all firms except firm $i$ is $y_{-i}$.

The oligopoly game is a Bayesian Cournot game in which each firm decides its product level based on the information available at the beginning of the second stage. We derive the symmetric equilibrium decision rule for the game beginning at the second stage. Note that, although we allow asymmetric networks throughout the paper, the unique Nash equilibrium decision rule in the game beginning at the second stage is "symmetric" in the limited sense that, given network $g$, each firm has the same form of decision rule even though its realizations

[7] If private signals are partially correlated, a firm can make partial inferences about the competitors' signals based on its own signal. This lowers firm's incentive for strategic link formation and information sharing. Therefore this assumption only leads to weakening our results since, as analyzed below, there already exist strong negative effects of information sharing under Cournot competition.

are different. That is, the firm's decision rule is affine in the vector of signals available to the firm. Since all firms are ex ante identical, labelling in the network is not important, and the only thing that matters is the number of links each firm retains given the network structure.

The (stategy) choice of firm $i$ in the whole game is a pair $((g_{ij})_{j\neq i}, q_i(y_i))$ where $g_{ij}$ : $\mathrm{N} \longrightarrow \{0,1\}$, and $q_i : G \times \mathrm{R} \times \mathrm{R}^{\eta_i(g)-1} \longrightarrow \mathrm{R}$. We denote by $Q(y)$ the vector of decision rules used by all firms, and by $Q_{-i}(y_{-i})$ the vector of decision rules of all firms except firm $i$.

The payoff of firm $i$ as a function of the stragegies chosen is:

$$\pi_i(g_i, g_{-i}, q_i(y_i), Q_{-i}(y_{-i})) = \mathrm{E}_{y,u}\{q_i(y_i)[a - b\sum_k q_k(y_k) + u]\} - \eta_i(g)\gamma \tag{2.3}$$

where E is the expected value operator. The Nash equilibrium of the above oligopoly game (played in the second stage) is $Q^*(y)$ which satisfies the following property:

$$J_i(q_i^*(y_i), Q_{-i}^*(y_{-i}) \mid g) \geq J_i(q_i(y_i), Q_{-i}^*(y_{-i}) \mid g) \quad \forall\ i, \tag{2.4}$$

where $q_i^*(\cdot),\ q_i(\cdot) : G \times \mathrm{R} \times \mathrm{R}^{\eta_i(g)-1} \longrightarrow \mathrm{R}$.

## 2.3 DERIVATION OF THE EQUILIBRIA

At the second stage, firm $i$ chooses its decision rule $q_i(\cdot)$ to maximize:

$$W = \mathrm{E}_{y_{-i},u}\{q_i(y_i)[a - b\sum_{j=1}^{n} q_j(y_j) + u] \mid y_i\} \tag{2.5}$$

In (2.5) profits are conditioned on the realization of $y_i$. Variables that remain unobserved at the beginning of the second stage are the values of the signals observed by all the unlinked firms $\{x_k | \ g_{ik} = 0\}_{k\neq i}$, the value of random variable $u$. Equation (2.5) may be rewritten:

$$W = q_i(y_i)[a - bq_i(y_i) - b\sum_{j\neq i} \mathrm{E}_{y_j}(q_j(y_j)|y_i) + \mathrm{E}_u(u|y_i)] \tag{2.6}$$

Firm $i$ chooses its decision rule $q_i(\cdot)$ to maximize (2.6) given the decision rule chosen by the other firms. When the private signals are independent, Proposition 1 displays the solution for any finite number of firms.

**Proposition 1.** *With independent signals, for given $g$, the following decision rule forms the unique Nash equilibrium of the game beginning at the second stage:*

$$q_i(y_i) = A_0^i + \sum_{j \neq i} A_j^i g_{ij} x_j + A_i^i x_i \ , \forall \ i \tag{2.7}$$

*where* $A_0^i = \frac{a}{b(n+1)}$, $A_j^i = \frac{\sigma}{(\eta_j(g)+2)bn(m+\sigma)}$, $A_i^i = \frac{\sigma}{(\eta_i(g)+2)bn(m+\sigma)}$.

*Proof.* To maximize (2.6) while taking $q_j(y_j)$ as given, set:

$$\frac{\partial W}{\partial q_i} = a - b \sum_{j \neq i} \mathrm{E}_{y_j}(q_j(y_j)|y_i) + \mathrm{E}_u(u|y_i) - 2bq_i(y_i) = 0$$

and

$$\frac{\partial^2 W}{\partial q_i^2} = -2b < 0.$$

Hence

$$q_i(y_i) = \frac{a - b \sum_{j \neq i} \mathrm{E}_{y_j}(q_j(y_j)|y_i) + \mathrm{E}_u(u|y_i)}{2b}, \ \forall \ i \tag{2.8}$$

Equation (2.8) is a necessary and sufficient condition for the decision rule $q_i(y_i), \forall \ i$ to be a Nash equilibrium. Using the posterior distribution of $u$, $\mathrm{E}(u|y_i) = \frac{1}{n} \frac{\sigma}{m+\sigma} \{ \sum_{j \neq i} g_{ij} x_j + x_i \}$, since $\mathrm{E}(u|y_i) = \frac{1}{n} \sum_{k=1}^{n} \mathrm{E}(u_k|y_i) = \frac{1}{n} [\mathrm{E}(u_i|x_i) + \sum_{j \neq i} g_{ij} \mathrm{E}(u_j|x_j)]$, and $\mathrm{E}(u_i|x_i) = \frac{\sigma}{m+\sigma} x_i$. Using the suggested solution of the Proposition,

$$\mathrm{E}_{y_j}(q_j(y_j)|y_i) = \mathrm{E}\{A_0^j + \sum_{k \neq j} A_k^j g_{jk} x_k + A_j^j x_j \mid y_i\} = A_0^j + \sum_{k \neq j} A_k^j g_{jk} g_{ik} x_k + A_j^j g_{ij} x_j. \tag{2.9}$$

Using these in condition (2.8) and requiring (2.8) to be satisfied for every possible $y_i$ and $y_j$, yields a system of equations with the same number of unknowns. Solving this equation system yields the unique solution specified in (2.7). According to Radner (1962) it is sufficient to restrict attention to decision rules of the generic form expressed by (2.7), since the decision rules must be affine in the vector of observations available to the firm. □

Since all firms are ex ante identical (they face the same technology and observe signals of the same precision), only the cardinality of the set $\mathrm{N}_i(g)$ matters in characterizing the decision rule. From this we observe that the firm imposes the same weights on signals transmitted from other firms as on its own signal as long as they have the same number of links. Without loss of generality, we assume that $q_i(y_i)$ is nonnegative. From equation (2.7) we can see various effects of information sharing. For example, suppose firm $i$ forms a link to firm $l$. There are two conflicting direct effects. First, there emerges an $A_l^i x_l$ term in the equation (since $x_l$ is now available information to the firm $i$), which makes $q_i(y_i)$ increasing. Second, firm $i$ imposes less weight in its own information $x_i$ (since the $\eta_i(g)$ component is in the denominator of $A_i^i$), which makes $q_i(y_i)$ decreasing. Also there are indirect effects embedded in equation (2.9). Expecting the decision rules of other firms connected to firm $l$, firm $i$ can deduce that any firm $k$ also use $x_l$ (now known to firm $i$) as information in its decision, if $g_{lk} = 1$. We view these effects as a kind of network externality. Next we will see whether negative or positive effects are bigger in terms of payoff. By considering the graph structure explicilty, we can unambiguously capture the effects of information sharing on the decision rules regarding quantity. This decision rule is reduced to the formula of Theorem 1 in Gal-Or (1985) when the complete network structure is exogenously given and each firm truthfully transmits its signal to the linked firms.

Now we use this equilibrium decision rule to derive the payoffs of the subgame that starts at the second stage, denoted $w_i(y_i; g)$ for firm $i$. The payoff function in this subgame starting at the second stage is then used to derive the payoff in the game that starts at the first stage. Denote this last function by

$$\pi_i(g) = \pi_i(q_i(g), q_{-i}(g)) = \mathrm{E}_{y_i}[w_i(y_i; g)] - \eta_i(g)\gamma = b\mathrm{E}_{y_i}[q_i(y_i)]^2 - \eta_i(g)\gamma$$

The last equality follows directly from the payoff function expressed in (6) and the form of the Nash equilibrium decision rule expressed in (8). We can rewrite $\pi_i(g)$ explicitly as follows:

$$
\begin{aligned}
\pi_i(g) &= b\mathrm{E}_{y_i}[A_0^i + \sum_{j \neq i} A_j^i g_{ij} x_j + A_i^i x_i]^2 - \eta_i(g)\gamma \\
&= \frac{a^2}{b(n+1)^2} + \sum_{j \neq i} g_{ij} \frac{\sigma^2}{(\eta_j(g)+2)^2 bn^2(m+\sigma)} \\
&\quad + \frac{\sigma^2}{(\eta_i(g)+2)^2 bn^2(m+\sigma)} - \eta_i(g)\gamma
\end{aligned} \tag{2.10}
$$

Equation (2.10) follows since $Var(x_k) = m+\sigma, \forall k$, and $Cov(x_i, x_j) = 0, \forall i, j, i \neq j$. Note that this derivation of the payoff functions is possible only under the assumptions of the linearity of market demand fuction and the normality of the signals. We can check the conflicting effects of network formation and information sharing on the payoff function in equation (2.10). With additional link formation (deletion), firm $i$ receives a negative (positive) effect on the third term in equation (2.10), and experiences a positive (negative) effect from the second term at the same time. The marginal benefit and marginal cost from an additional link with $j$ are $\frac{\sigma^2}{((\eta_j(g)+1)+2)^2 bn^2(m+\sigma)}$ and $\frac{\sigma^2(2\eta_i(g)+5)}{(\eta_i(g)+2)^2(\eta_i(g)+3)^2 bn^2(m+\sigma)}$ respectively, while the marginal benefit and marginal cost from severing an existing link to $j$ are $\frac{\sigma^2(2\eta_i(g)+3)}{((\eta_i(g)+1)^2((\eta_i(g)+2)^2 bn^2(m+\sigma)}$ and $\frac{\sigma^2}{(\eta_j(g)+2)^2 bn^2(m+\sigma)}$ respectively, where $\eta_i(g)$ and $\eta_j(g)$ are the cardinalities of $\mathrm{N}_i(g)$ and $\mathrm{N}_j(g)$ before link formation(deletion).

### 2.3.1 Pairwise stable networks

We characterize the stable collaborative networks under quantity competition. It turns out that there exist multiple (symmetric and asymmetric) pairwise stable networks. Recall that we assume that the link formation cost, $\gamma$, is negligibily small since our main goal is to study the incentives for information sharing (equivalently, the benefit and cost of information sharing) when firms face uncertainty. The interesting findings, as we will see, are that even in the case where $\gamma = 0$, firms are not willing to form as many links as they possibly can. This indicates that information sharing itself has a nonneglibibly negative effect on the payoff of the firm.

We think it is natural and important to start by checking the commonly known results of the existing information sharing literature. Recall that, under unknown common demand, no information sharing is the unique Nash equilibrium outcome under Cournot competition (Vives (1984) and Gal-Or (1985)).

**Proposition 2.** *Under Assumptions (1), (2) and $\sigma > 0$, both the empty and the complete network are pairwise stable when $n \geq 3$.*

*Proof.* First we show that the empty network is pairwise stable. Notice that the stability condition (i) is trivially satisfied. Thus we only need to check whether condition (ii) is satisfied. Suppose that firms $i$ and $j$ form a link. The resulting network is $g^0 + g_{ij}$. We next check whether firms $i$ and $j$ find such a deviation profitable. Using equation (10), we have $\pi_i(g^0) - \pi_i(g^0 + g_{ij}) = \frac{\sigma^2}{36bn^2(m+\sigma)} + \gamma > 0$. So condition (ii) is satisfied. Now we show that the complete network is pairwise stable. By a similar reasoning, we only need to check whether pairwise stability condition (i) is satisfied. Consider again that one of the firms $i$ and $j$ sever a link. The resulting network is $g^{n-1} - g_{ij}$. Using equation (10), we have $\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij}) = \pi_j(g^{n-1}) - \pi_j(g^{n-1} - g_{ij}) = \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{n^2-2n-1}{(n+1)^2n^2}\} - \gamma > 0$ if $n \geq 3$. So condition (i) is satisfied. Hence the complete network is also pairwise stable. $\square$

This result indicates that complete information sharing is also a pairwise stable equilibrium. Note that centralized (and industry-wide) decision making, by definition, is exactly the same as decentralized (and pairwise) decision making if $n = 2$. So we reasonably expect that both produce the same result. The following result shows that the empty network (no information sharing) is the unique equilibrium outcome in a duopoly as in the centralized setting.

**Remark** In a duopoly (i.e, $n = 2$), the empty network (no information sharing) is a unique equilibrium outcome. From Proposition 2, $\pi_i(g^0) - \pi_i(g^0 + g_{ij}) = \frac{\sigma^2}{36bn^2(m+\sigma)} + \gamma > 0$, and $\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij}) = \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{n^2-2n-1}{(n+1)^2n^2}\} - \gamma < 0$ when $n = 2$. $\|$

With this, we can derive the following result directly from Proposition 2. Recall that the dominant group architecture, $g^{d(k)}$, is characterized by one complete non-singleton component with $k \geq 2$ firms and $n-k$ singleton firms. Thus, there is a set of firms $\mathrm{N}^d \subset \mathrm{N}$ with the property that $g_{ij} = 1$ for every pair $i, j \in \mathrm{N}^d$ while for any $m \in \mathrm{N}\backslash\mathrm{N}^d, g_{ml} = 0, \forall l \in \mathrm{N}\backslash\{m\}$ (Goyal and Joshi (2003)).

**Proposition 3.** *Under Assumptions (1), (2) and $\sigma > 0$, the dominant group architecture, $g^{d(k)}$, is pairwise stable when $k \geq 3$.*

*Proof.* Consider any firm $i \in \mathrm{N}^d$. From Proposition 2, this firm has no incentive to sever its existing links. That is, $\forall\ i,\ \pi_i(g^{k-1}) - \pi_i(g^{k-1} - g_{ij}) > 0$ when $k \geq 3$. So the stability condition (i) is easily satisfied. We need to check whether this firm has any incentive to form a link to an isolated firm $m$.

$$\begin{aligned}
\pi_i(g) - \pi_i(g + g_{im}) &= \{\frac{a^2}{b(n+1)^2} + k\frac{\sigma^2}{b(k+1)^2 n^2 (m+\sigma)}\} - \{\frac{a^2}{b(n+1)^2} \\
&\quad +(k-1)\frac{\sigma^2}{b(k+1)^2 n^2(m+\sigma)} + \frac{\sigma^2}{b(k+2)^2 n^2 (m+\sigma)} + \frac{\sigma^2}{9bn^2(m+\sigma)}\} + \gamma \\
&< 0 \ \text{ when } k \geq 2.
\end{aligned}$$

This means that firm $i$ wants to form an additional link to the isolated firm $m$. The remaining thing to check is the incentives of firm $m$ :

$$\begin{aligned}
\pi_m(g) - \pi_m(g + g_{im}) &= \{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{4bn^2(m+\sigma)}\} \\
&\quad -\{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{9bn^2(m+\sigma)} + \frac{\sigma^2}{b(k+2)^2 n^2(m+\sigma)}\} - \gamma \\
&> 0 \ \text{ when } k \geq 2.
\end{aligned}$$

So, the isolated firm wants to remain isolated. Also, following from the previous Corollary, firm $m$ does not have any incentives to form a link to $l \in \mathrm{N}\backslash\{m\}$. Therefore, the stability condition (ii) is satisfied. □

This shows that partial and asymmetric information sharing appears as an equilibrium if firms make a decentralized decision, even though the firms are ex ante identical. In addition, there exist other (symmetric and asymmetric) equilibrium structures specified below. The following result incorporates all findings studied above as special cases.

**Theorem 4.** *Let* $\mathcal{F}_1(\mathrm{N}) = \{\mathrm{N_1}, \mathrm{N_2}, \cdots, \mathrm{N_p}\}$ *be a partition of* $\mathrm{N}$ *such that* $\forall\ \mathrm{i} \in \{1, \cdots, \mathrm{p}\}$, $|\mathrm{N_i}| \neq 2$, *and* $\forall\ \mathrm{i} \in \{1, \cdots, \mathrm{p}-1\}$, $|\mathrm{N_{i+1}}| > \frac{(|\mathrm{N_i}|+1)(|\mathrm{N_i}|+2)}{\sqrt{2|\mathrm{N_i}|+3}} - 2$. *Let* $g^{|\mathrm{N_i}|-1}$ *denote the complete network over* $\mathrm{N_i}$ *for all* $\mathrm{i} \in \{1, \cdots, \mathrm{p}\}$. *Then*

*(1)* $g(\mathcal{F}(\mathrm{N})) = \bigcup_{\mathrm{i}=1}^{\mathrm{p}} g^{|\mathrm{N_i}|-1}$ *is a pairwise stable network,*

*(2) If* $g$ *is component-symmetric and pairwise stable, then* $g \in g(\mathcal{F}(\mathrm{N})) = \bigcup_{\mathrm{i}=1}^{\mathrm{p}} g^{|\mathrm{N_i}|-1}$.

*Proof. For the proof of (1)*: Take any network $g(\mathcal{F}(\mathrm{N})) = \bigcup_{\mathrm{i}=1}^{\mathrm{p}} g^{|\mathrm{N_i}|-1}$ satisfying the conditions specified. Without loss of generality, take any three firms $i, j$, and $k$ such that $i \in \mathrm{N_i}$, $j \in \mathrm{N_j}$, $k \in \mathrm{N_{j+1}}$ where $|\mathrm{N_i}| = 1$, $|\mathrm{N_j}| \geq 3$. First, we show that the firms $j$ and $k$ have no incentive to sever their existing link. In Proposition 2 we have shown that $\pi_j(g^{|\mathrm{N_j}|-1}) - \pi_j(g^{|\mathrm{N_j}|-1} - g_{jl}) > 0$ if $|\mathrm{N_j}| \geq 3, l \in \mathrm{N_j}$. This is true for the case of firm $k$. So condition (i) of the definition of pairwise stability is satisfied. Let's check whether condition (ii) of pairwise stability is satisfied. Obviously, the isolated firm $i$ has no incentive for link formation since $\pi_i(g) - \pi_i(g + g_{ij}) > 0$ and $\pi_i(g) - \pi_i(g + g_{ik}) > 0$. Then we only need to investigate the incentives of firms $j$ and $k$ by checking

$$\begin{aligned}
\pi_j(g) - \pi_j(g + g_{jk}) &= \{\frac{a^2}{b(n+1)^2} + |\mathrm{N_j}|\frac{\sigma^2}{b(|\mathrm{N_j}|+1)^2 n^2(m+\sigma)}\} - \{\frac{a^2}{b(n+1)^2} \\
&\quad +(|\mathrm{N_j}|-1)\frac{\sigma^2}{b(|\mathrm{N_j}|+1)^2 n^2(m+\sigma)} + \frac{\sigma^2}{b(|\mathrm{N_j}|+2)^2 n^2(m+\sigma)} \\
&\quad +\frac{\sigma^2}{b(|\mathrm{N_{j+1}}|+2)^2 n^2(m+\sigma)}\} - \gamma \\
&= \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{1}{(|\mathrm{N_j}|+1)^2} - \frac{1}{(|\mathrm{N_j}|+2)^2} - \frac{1}{(|\mathrm{N_{j+1}}|+2)^2}\} - \gamma \\
&> 0 \text{ if } |\mathrm{N_{j+1}}| > \frac{(|\mathrm{N_j}|+1)(|\mathrm{N_j}|+2)}{\sqrt{2|\mathrm{N_j}|+3}} - 2.
\end{aligned}$$

This means that firm $j$ does not agree to form an additional link with $k$ even if firm $k$ tries to form a link to $j$. This implies that $j$ does not form a link to any other firm with more than $|\mathrm{N_{j+1}}| - 1$ links. Therefore condition (ii) of pairwise stability is also satisfied.

*For the proof of (2)*: It suffices to show that each component $g' \in g$ is complete. We denote a symmetric component of degree $\Delta$ by $g'^{\Delta}$, $\Delta = 2, 3, \cdots, |\mathrm{N_i}| - 1$. Suppose that $|\mathrm{N_i}| \geq 3$. Then $\forall\ i, j \in \mathrm{N_i}$, $\pi_i(g'^{\Delta}) - \pi_i(g'^{\Delta} + g_{ij}) = \frac{\sigma^2\{(\Delta+3)^2 - 2(\Delta+2)^2\}}{bn^2(m+\sigma)(\Delta+3)^2(\Delta+2)^2} < 0$ for $\Delta = 2, 3, , \cdots, |\mathrm{N_i}| - 2$,

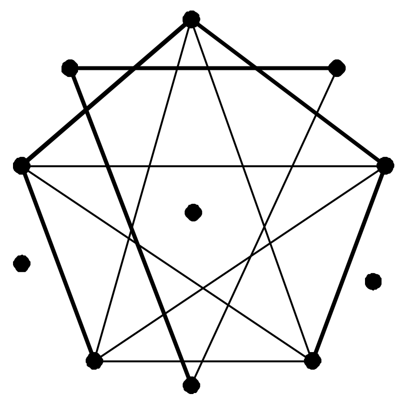

Figure 3: An example of graph $\mathcal{F}_1(\mathrm{N})$

and $\forall\ i, j \in \mathrm{N_i}$, $\pi_i(g'^{|\mathrm{N_i}|-1}) - \pi_i(g'^{|\mathrm{N_i}|-1} - g_{ij}) > 0$. Therefore, if $g$ is component-symmetric and pairwise stable, each component $g' \in g$ must be complete. $\square$

Theorem 4 illustrates the basic feature of pairwise stable network structures. The more links a firm has, the stronger is its incentive additional information. This property becomes clear when we introduce heterogeneity into the model in section 4. From this we can summarize our main finding: Unlike the result of previous literature on "centralized" information sharing, there emerges a broader level of information sharing as pairwise stable equilibria.

**Example** Suppose $n = 4$. Then $\{g^0, g^{d(3)}, g^3\}$ is the set of *all* pairwise stable networks.

Here we feel it necessary to interpret the results of the present model qualitatively. When firms behave as Nash competitors in the market, the effects of pooling private information

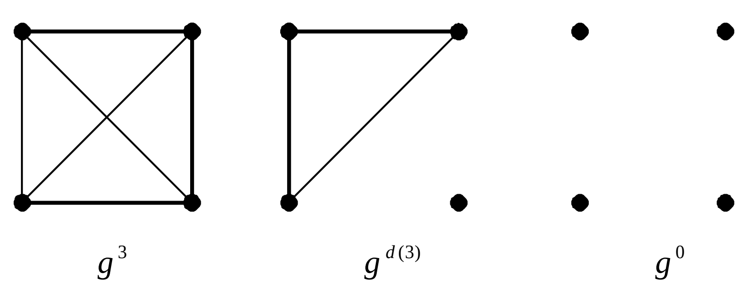


Figure 4: All pairwise stable networks when $n = 4$

on the firms profits are unclear. When more accurate information is available, the strategies can be chosen more accurately. Increased accuracy has an unambiguously positive effect on the payoff of the firm. However, the pooling of private information increases the correlation among the firms' decision rules. This increased correlation has ambiguous effects on the firm. Suppose that a firm observes a signal of low demand and shares it with others. Then it reduces the likelihood that its competitors overproduce. But when it observes a signal of high demand and reveals it to others, it reduces the likelihood that its competitors underproduce. While the first effect raises the profits of the firm, the second effects reduces them. Hence, it is unclear whether firms will transmit their private information to rival firms. In the previous literature, as Gal-or (1985) indicates, the benefits from pooling information and obtaining a more accurate measure of demand are dominated by the losses from increasing the correlation of firms' output decisions. Due to the underlying complete network structure, the models in the previous literature have much higher correlation among the decision rules than the present model. Thus, the negative effect dominates and no information sharing is the equilibrium outcome. In the present model, due to the (decentralized) pairwise interaction structure, the correlation in the decision rules is lower than in the previous literature. But there still exists a strong negative effect from correlation of firms' output decision.

Until now we have considered pairwise stablity as a solution concept which allows unilateral and pairwise deviation of the firms. As is well known in the network formation literature, individual or pairwise based solution concepts often lead to multiple stable networks, so that they provide broad predictions. Nevertheless we would like to emphasize this result, since we think that pairwise stability is a relevant and natural equilibrium concept in the analysis of decentralized oligopolistic markets. We can imagine and observe that, in an oligopoly which consists of Nash competitors, forming a large and credible coalition is quite difficult. In our context, however, we cannot exclude the possibility of communication among firms that may allow a number of them to coordinate their choices of links. It is also probable that information sharing depends on the coaltion structure. Given such coalitional considerations, we now study strongly stable networks as a means of making tighter predictions.

### 2.3.2 Strongly stable networks

Strong stability of networks is a very demanding property, since it means that no set of firms could benefit through any rearranging of the links that they are involved in (including those linking to firms outside the coaltion). Such networks are essentially impossible to destabilize, as there is no possible reorganization that would be improving for all firms whose consent is needed. We now characterize the strongly stable collaboration networks under quantity competition.

**Remark** Under Assumptions (1), (2) and $\sigma > 0$, neither the complete network, $g^{n-1}$, nor the *dominant group architecture,* $g^{d(k)}$, are strongly stable.

We can show this by way of counterexample. First, consider a certain firm, $i$, in the complete network $g^{n-1}$, and suppose that firm $i$ alone deviates by severing all its links at once. The strong stability concept captures this kind of deviation while the pairwise stability does not. Then, firm $i$ is the only singleton firm and the remaining $(n-1)$ firms form one complete non-singleton component. Formally, the dominant group architecture $g^{d(n-1)} \in G$ is obtainable from $g^{n-1} \in G$ via a deviation by $S = \{i\}$, and we can check that

$$\begin{aligned}
\pi_i(g^{d(n-1)}) - \pi_i(g^{n-1}) &= \{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{4bn^2(m+\sigma)}\} - \{\frac{a^2}{b(n+1)^2} + \frac{n\sigma^2}{(n+1)^2bn^2(m+\sigma)} - (n-1)\gamma\} \\
&= \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{(n-1)^2}{4(n+1)^2}\} + (n-1)\gamma > 0
\end{aligned}$$

That is, this kind of deviation is profitable. Therefore $g^{n-1}$ is not strongly stable. Similar arguments hold in the case of the dominant group architecture $g^{d(k)} \in G$. Consider again a certain firm $j$ in the complete component, (i.e., $j \in \mathrm{N}^d$), and suppose firm $j$ severs all its links at once. Then we can find that another dominant group architecture $g^{d(k-1)} \in G$ is obtainable from $g^{d(k)} \in G$ via a deviation by $S = \{j\}$, and we can check that

$$\begin{aligned}
\pi_j(g^{d(k-1)} - \pi_j(g^{d(k)}) &= \{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{4bn^2(m+\sigma)}\} - \{\frac{a^2}{b(n+1)^2} + \frac{k\sigma^2}{(k+1)^2bn^2(m+\sigma)} - (k-1)\gamma\} \\
&= \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{(k-1)^2}{4(k+1)^2}\} + (k-1)\gamma > 0 \text{ when } k \geq 2
\end{aligned}$$

Therefore $g^{d(k)} \in G$ is not strongly stable.

This example says that complete information sharing is not an equilibrium outcome if we allow a broad range of coordination and cooperation among the group of firms in the industry. Also it captures the idea that a firm would not benefit from severing any single link but would benefit from severing several links simultaneously which is not accounted for under the concept of pairwise stability. A some discussion of this issue will follow in section 5.

**Theorem 5.** *Under Assumptions (1), (2) and $\sigma > 0$, no information sharing (and the resulting empty network) is the unique strongly stable equilibrium outcome.*

*Proof.* Suppose that the empty network $g^0$ is not strongly stable. Consider the deviating coaltion $S \ni i, j$, and the obtainable graph $g' \in G$ from $g^0 \in G$ via a deviation by $S$. Needless to say, $g'$ must be nonempty. Suppose $\eta_i(g') = 1$, $g_{ij} = 1$. Then

$$\begin{aligned}\pi_i(g^0) - \pi_i(g') &= \{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{4bn^2(m+\sigma)}\} - \{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{9bn^2(m+\sigma)} \\ &\quad + \frac{\sigma^2}{(\eta_j(g')+2)^2 bn^2(m+\sigma)} - \gamma\} \\ &= \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{5(\eta_j(g')+2)^2 - 36}{36(\eta_j(g')+2)^2}\} + \gamma > 0\end{aligned}$$

That is, forming a single link is not profitable. Hence, every deviating firm in $S$ must form at least two links simultaneously. Now suppose that every firm in $S$ forms at least two links. We can show this kind of deviation is not profitable either. That is,

$$\begin{aligned}\pi_i(g^0) - \pi_i(g') &= \{\frac{a^2}{b(n+1)^2} + \frac{\sigma^2}{4bn^2(m+\sigma)}\} - \{\frac{a^2}{b(n+1)^2} \\ &\quad + \frac{\sigma^2}{(\eta_i(g')+2)^2 bn^2(m+\sigma)} + \sum g_{ij} \frac{\sigma^2}{(\eta_j(g')+2)^2 bn^2(m+\sigma)} - \eta_i(g')\gamma\} \\ &= \frac{\sigma^2}{bn^2(m+\sigma)}\{\frac{1}{4} - \frac{1}{(\eta_i(g')+2)^2} - \sum g_{ij}\frac{1}{(\eta_j(g')+2)^2}\} + \eta_i(g')\gamma \\ &> 0 \text{ where } \eta_i(g') \geq 2,\ \eta_j(g') \geq 2 \text{ if } g_{ij} = 1, \forall\ j\end{aligned}$$

Therefore, a profitable deviation does not exist. This contradicts the assumption that $g^0$ is not strongly stable. Also, this implicitly proves the uniqueness of strongly stable networks. If there is a nonempty graph $g$ where a typical firm $i$ has a single link, then this firm will benefit from a unilateral deviation by severing the link. Suppose that there is another graph $g'$ where each firm has at least two links, or remains isolated. Then each firm with more than two links will benefit from severing all links simultaneously. This completes the proof. □

We reach the same no-information-sharing result as in the previous literature by allowing all firms to deviate and coordinate their choices simultaneously. It is not a coincidence that both approaches produce the same result. The concept of strong stability moves the analysis closer to that of the existing literature; by allowing a wider level of coalition and cooperation among firms, we allow firms to make an industry-wide decision. The wider the level of coaltion that is allowed, the more similar are the effects produced. In this sense, our result extends earlier findings about the incentives of oligopolistic firms, and can be regarded as complementary to the existing literature. Also, by charaterizing strongly stable networks, we have demonstrated that information sharing critically depends on the coaltion and cooperation structure as well. Note that strong stability is a very demanding property, in the sense that once formed such networks are essentially impossible to destabilize, as there is no possible reorganization that would be improving for all of the firms whose consent is needed. Therefore, this no-information-sharing result cannot understate the pairwise stable equilibrium outcomes of the model analyzed above. Furthermore, we can carefully state that the results of the existing literature could be obtained in very restricted situations, since they reach such results under the exogenously given grand coalition assumption.

#### 2.3.3 Efficient networks

For a given network $g$, social welfare, $V(g)$, is defined as the sum of the consumer surplus and the aggregate profit of the $n$ firms. We shall say that a network $g^*$ is efficient if $V(g^*) \geq V(g)$, for all $g \in G$. In this subsection we confine our attention to symmetric networks $g^{\Delta}$. When firms operate in the homogeneous product oligopoly, social welfare is

$$\begin{aligned}
V(g^{\Delta}) &= \frac{1}{2}b\mathrm{E}[Q]^2 + \sum_{i=1}^{n}\pi_i(g^{\Delta}) - \sum_{i=1}^{n}\eta_i(g^{\Delta})\gamma \\
&= \frac{1}{2}b\mathrm{E}[\sum_{i}^{n} q_i(g^{\Delta})]^2 + b\sum_{i=1}^{n}\mathrm{E}[q_i(g^{\Delta})]^2 - \sum_{i=1}^{n}\eta_i(g^{\Delta})\gamma \\
&= (\frac{1}{2}n^2 + n)\{\frac{a^2}{b(n+1)^2} + \sum_{j\neq i} g_{ij}\frac{\sigma^2}{(\eta_j(g^{\Delta})+2)^2 bn^2(m+\sigma)} + \frac{\sigma^2}{(\eta_i(g^{\Delta})+2)^2 bn^2(m+\sigma)}\} \\
&\quad -n\eta_i(g^{\Delta})\gamma, \text{ where } \eta_i(g^{\Delta}) = \eta_j(g^{\Delta}) = \Delta
\end{aligned}$$

It is easily seen that $V(g^{\Delta})$ is decreasing over $\Delta = 0, 1, 2, \cdots, n-1$, since the decrease of the term $\frac{\sigma^2}{(\eta_i(g)+2)^2 bn^2(m+\sigma)}$ from adding links dominates the appearance of the term $g_{ij}\frac{\sigma^2}{(\eta_j(g)+2)^2 bn^2(m+\sigma)}$. Therefore, the empty network is the unique efficient outcome among symmetric networks. Recall that if the number of firms is even, there always exists a symmetric network of degree $\Delta$, $\Delta = 0, 1, \cdots, n-1$ (Goyal and Moraga-Gonzalez (2001)).

**Lemma 6.** *Suppose that $n$ is even. The empty network (no information sharing) is the unique (pairwise and strongly) stable and efficient outcome among all symmetric networks.*

## 2.4 HETEROGENEOUS FIRMS

In this section we introduce heterogeneity across firms by assuming that firms observe a noisy signal for $u$ with different degree of accuracy. We maintain all other assumptions. This environment enables us to analyze which firm has the larger incentive to share information when the accuracy of private information is different across firms. While our goal remains that of characterizing the set of stable network structures, we can check the robustness of the basic model by perturbing the model a little. The signal observed by firm $i$ is $x_i$. We now assume:

$$x_i = u_i + e_i, \qquad u_i \sim N\ (0, \sigma_i), \qquad e_i \sim N\ (0, m_i) \tag{2.11}$$

where $Cov(e_i, e_j) = 0,\quad i \neq j;\ Cov(u_i, e_j) = 0\quad \forall\ i, j;\quad Cov(u_i, u_j) = 0,\quad i \neq j;\ \sigma_i < \infty,\ m_i < \infty, \forall\ i$; and $u = (\sum_i u_i)\ /\ n$, hence $u \sim N\ (0, \frac{1}{n^2}\sum \sigma_i)$. With a little calculation, the following decision rule forms the unique Nash equilibrium of the game beginning at the second stage:

$$q_i(y_i) = A_0^i + \sum_{j \neq i} A_j^i g_{ij} x_j + A_i^i x_i\ , \forall\ i \tag{2.12}$$

where $A_0^i = \frac{a}{b(n+1)}$, $A_j^i = \frac{\sigma_j}{(\eta_j(g)+2)bn(m_j+\sigma_j)}$, $A_i^i = \frac{\sigma_i}{(\eta_i(g)+2)bn(m_i+\sigma_i)}$. The payoff function starting at the first stage is given by:

$$\pi_i(g) = \frac{a^2}{b(n+1)^2} + \sum_{j \neq i} g_{ij}\frac{\sigma_j^2}{(\eta_j(g)+2)^2 bn^2(m_j+\sigma_j)} + \frac{\sigma_i^2}{(\eta_i(g)+2)^2 bn^2(m_i+\sigma_i)} - \eta_i(g)\gamma \tag{2.13}$$

We can now characterize the set of stable networks among heterogeneous firms.

**Proposition 7.** *Under Assumptions (1), (11) and $\sigma_i > 0 \ \forall\ i$, the empty network is always pairwise stable, and the complete network is pairwise stable if $n \geq 3$ and $\frac{2n+1}{n^2} < \frac{\sigma_j^2(m_i+\sigma_i)}{\sigma_i^2(m_j+\sigma_j)} < \frac{n^2}{2n+1}$, $\forall\ i, j$.*

*Proof.* First we show that the empty network is pairwise stable. Using equation (13), we have $\pi_i(g^0) - \pi_i(g^0 + g_{ij}) = \frac{5\sigma_i^2(m_j+\sigma_j)-4\sigma_j^2(m_i+\sigma_i)}{36bn^2(m_i+\sigma_i)(m_j+\sigma_j)} + \gamma$. If $5\sigma_i^2(m_j + \sigma_j) - 4\sigma_j^2(m_i + \sigma_i) > 0$, the proof is complete. Suppose $5\sigma_i^2(m_j + \sigma_j) - 4\sigma_j^2(m_i + \sigma_i) < 0$. Then we have $\pi_j(g^0) - \pi_j(g^0 + g_{ij}) = \frac{5\sigma_j^2(m_i+\sigma_i)-4\sigma_i^2(m_j+\sigma_j)}{36bn^2(m_i+\sigma_i)(m_j+\sigma_j)} + \gamma > 0$. So condition (ii) is satisfied. Similarly, we can show that $\pi_i(g^{n-1}) - \pi_i(g^{n-1} - g_{ij}) > 0$ and $\pi_j(g^{n-1}) - \pi_j(g^{n-1} - g_{ij}) > 0$ if $n \geq 3$ and $\frac{2n+1}{n^2} < \frac{\sigma_j^2(m_i+\sigma_i)}{\sigma_i^2(m_j+\sigma_j)} < \frac{n^2}{2n+1}$, $\forall\ i, j$. □

We can draw many implications from Proposition 7. There are two points we want to make about this result. First, the empty network is stable since a firm with the less accurate information of the two declines. It is still true that for any given $g$, if link formation fails between any two firms, it is the firm with inaccurate information that refuses to collaborate. Second, the complete network is stable only when firms with the similar level of accuracy comprise the network. When firms are symmetric, the complete network is pairwise stable if $n \geq 3$ (Proposition 2).

**Remark** In any equilibrium, information is shared among firms with a similar accuracy of information. If information sharing fails between any two firms for any network structure, it is the firm with inaccurate information that refuses to collaborate.

**Example** Under Assumptions (1), (11) and $\sigma_i > 0 \ \forall\ i$, the dominant group architecture, $g^{d(k)}$, is pairwise stable if (1) $k \geq 3$, (2) $\frac{2k+1}{k^2} < \frac{\sigma_j^2(m_i+\sigma_i)}{\sigma_i^2(m_j+\sigma_j)} < \frac{k^2}{2k+1}$, $\forall\ i, j \in \mathrm{N}^d$, and (3) for any $i \in \mathrm{N}^d$ and any $h \in \mathrm{N}\backslash\mathrm{N}^d$, $\frac{\sigma_h^2(m_i+\sigma_i)}{\sigma_i^2(m_h+\sigma_h)} > \frac{36}{5(k+1)^2}$ or $\frac{\sigma_h^2(m_i+\sigma_i)}{\sigma_i^2(m_h+\sigma_h)} < \frac{9(2k+3)}{(k+1)^2(k+2)^2}$. ||

$g^{d(k)}$ is pairwise stable when the complete nonsingleton component is formed by the $k$ firms with similar accuracy (represented by conditions (1) and (2)), and the dominant group of $k$ firms and $n - k$ singlton firms are significantely differentiated in their private information (represented by condition (3)). Also this shows that partial or asymmetric information sharing emerges as an equilibrium when heterogeneous firms make a decision in the decentralized way. The following result characterizes all pairwise stable networks, and incorporates all findings studied above as special cases.

**Theorem 8.** *Let* $\mathrm{N} = \{1, 2, \cdots, n\}, n \geqq 3$ *be the set of rearranged firms such that* $\frac{\sigma_i^2}{(m_i+\sigma_i)} > \frac{\sigma_{i+1}^2}{(m_{i+1}+\sigma_{i+1})}, \forall\, i \in \{1, 2, \cdots, n-1\}$. *So, firm* 1 *is the one with the least accurate information, while firm* $n$ *is the one with the most accurate information. Let* $\mathcal{F}_2(\mathrm{N}) = \{\mathrm{N}_1, \mathrm{N}_2, \cdots, \mathrm{N}_\mathrm{p}\}$ *be a partition of* N *such that* $\forall\ \mathrm{i} \in \{1, 2, \ldots, \mathrm{p}-1\}$, $\forall\ i \in \mathrm{N}_\mathrm{i},\ j \in \mathrm{N}_{\mathrm{i}+1},\ \frac{\sigma_i^2}{(m_i+\sigma_i)} > \frac{\sigma_j^2}{(m_j+\sigma_j)}$. *Suppose that* $g(\mathrm{N}_\mathrm{i})$ *is the network over* $\mathrm{N}_\mathrm{i}$ *for all* $\mathrm{i} \in \{1, \cdots, \mathrm{p}\}$. *Then* $g(\mathcal{F}_2(\mathrm{N})) = \bigcup\limits_{\mathrm{i}=1}^{\mathrm{p}} g(\mathrm{N}_\mathrm{i})$ *is a pairwise stable network iff*

*(1) If* $\mathrm{N}_\mathrm{i}$ *is nonsingleton,* $|\mathrm{N}_\mathrm{i}| \geq 3$ *and* $\eta_i(g(\mathrm{N}_\mathrm{i})) \geq 2,\ \forall\ i \in \mathrm{N}_\mathrm{i}$

*(2) For* $g_{ii+h} = 1,\ \frac{\sigma_i^2(m_{i+h}+\sigma_{i+h})}{\sigma_{i+h}^2(m_i+\sigma_i)} < \frac{(\eta_i(g)+1)^2(\eta_i(g)+2)^2}{(\eta_{i+h}(g)+2)^2(2\eta_i(g)+3)}$, *and*
*for* $g_{ii+h} = 0,\ \frac{\sigma_i^2(m_{i+h}+\sigma_{i+h})}{\sigma_{i+h}^2(m_i+\sigma_i)} > \frac{(\eta_i(g)+2)^2(\eta_i(g)+3)^2}{(\eta_{i+h}(g)+3)^2(2\eta_i(g)+5)},\ \forall\ i, i+h \in \mathrm{N}_\mathrm{i}$

*(3)* $\forall\ \mathrm{i} \in \{1, 2, \ldots, \mathrm{p}-1\},\ \forall\ i \in \mathrm{N}_\mathrm{i},\ j \in \mathrm{N}_{\mathrm{i}+1},\ \frac{\sigma_i^2(m_j+\sigma_j)}{\sigma_j^2(m_i+\sigma_i)} > \frac{(\eta_i(g)+2)^2(\eta_i(g)+3)^2}{(\eta_j(g)+3)^2(2\eta_i(g)+5)}$.

*Proof.* Just use the definition of pairwise stability. Condition (1) allows the existence of singleton components, condition (2) indicates stability within a component, and condition (3) guarantees stability between any two firms from distinct components. □

This is somewhat abstract, so here we take a special example which captures all the basic features of the model.

**Example: Assortative Networks** Let $\mathcal{F}_2(\mathrm{N}) = \{\mathrm{N}_1, \mathrm{N}_2, \cdots, \mathrm{N}_\mathrm{p}\}$ be a partition of N such that $\forall\ \mathrm{i} \in \{1, 2, \ldots, \mathrm{p}-1\},\ \forall\ i \in \mathrm{N}_\mathrm{i},\ j \in \mathrm{N}_{\mathrm{i}+1},\ \frac{\sigma_i^2}{(m_i+\sigma_i)} > \frac{\sigma_j^2}{(m_j+\sigma_j)}$. So, $\mathrm{N}_1$ is the component which consists of firms with the least accurate information, while $\mathrm{N}_\mathrm{p}$ is the component which consists of firms with the most accurate information. Suppose that $g^{|\mathrm{N}_\mathrm{i}|-1}$ is the complete network over $\mathrm{N}_\mathrm{i}$ for all $\mathrm{i} \in \{1, \cdots, \mathrm{p}\}$. Then $g(\mathcal{F}_2(\mathrm{N})) = \bigcup\limits_{\mathrm{i}=1}^{\mathrm{p}} g^{|\mathrm{N}_\mathrm{i}|-1}$ is a pairwise stable network if

(1) If $\mathrm{N}_\mathrm{i}$ is nonsingleton, $|\mathrm{N}_\mathrm{i}| \geq 3$ and $\frac{2|\mathrm{N}_\mathrm{i}|+1}{|\mathrm{N}_\mathrm{i}|^2} < \frac{\sigma_h^2(m_i+\sigma_i)}{\sigma_i^2(m_h+\sigma_h)} < \frac{|\mathrm{N}_\mathrm{i}|^2}{2|\mathrm{N}_\mathrm{i}|+1},\quad \forall\ i, h \in \mathrm{N}_\mathrm{i}$,

(2) $\forall\ \mathrm{i} \in \{1, 2, \ldots, \mathrm{p}-1\},\ \forall\ i \in \mathrm{N}_\mathrm{i},\ j \in \mathrm{N}_{\mathrm{i}+1},\ \frac{\sigma_i^2(m_j+\sigma_j)}{\sigma_j^2(m_i+\sigma_i)} > \frac{(|\mathrm{N}_\mathrm{i}|+1)^2(|\mathrm{N}_\mathrm{i}|+2)^2}{(|\mathrm{N}_\mathrm{j}|+2)^2(2|\mathrm{N}_\mathrm{i}|+3)}$.

This is similar to Theorem 4. Condition (1) guarantees the completeness of a component. Condition (2) requires that signals should be differentiated enough that firms from the different components do not form a link to each other. Partition $\mathcal{F}_2(\mathrm{N})$ exactly corresponds to partition $\mathcal{F}_1(\mathrm{N})$. The size of the component in partition $\mathcal{F}_1(\mathrm{N})$ has the same role as the accuracy of firms' information in the component in partition $\mathcal{F}_2(\mathrm{N})$. ‖

From this we can check the basic features of decentralized information sharing among heterogeneous firms: In any equilibrium, information is shared among firms with a similar accuracy of information. And if information sharing fails between any two firms for any network structure, it is the firm with inaccurate information that refuses to collaborate. Also this implies partial and asymmetric information sharing results.

**Lemma 9.** *Under Assumptions (1), (11) and $\sigma > 0$, no information sharing (and the resulting empty network) is the unique strongly stable equilibrium outcome.*

The results of this section reinforce previous findings in that partial and asymmetric information sharing emerges when firms make a decision in a decentralized way. In this section we introduced heterogeneity across firms by assuming that firms observed a noisy signal for $u$ with different degree of accuracy. We show that, unlike the result of the previous literature on the centralized information sharing, there emerges a broader level of information sharing as a pairwise stable equilibrium.

## 2.5 DISCUSSION OF FURTHER APPROACHES

### 2.5.1 A noncooperative game of network formation

Myerson (1991) suggests a noncooperative game of network formation. For every firm $i$, the strategy set is an $n-1$ tuple of 0 and 1, $G_i = \{0,1\}^{n-1}$. Let $g_{ij}$ denote the $j$th coordinate of $g_i$. If $g_{ij} = 1$, firm $i$ indicates its willingness to form a link with firm $j$. Given the strategy profile $g$, an undirected network $g$ is formed by letting firms $i$ and $j$ linked if and only if $g_{ij}g_{ji} = 1$. In words, the formation of a link requires the consent of both firms.

A strategy profile $g$ is a *Nash equilibrium* if and only if, for all $i$, all strategies $g_i'$ in $G_i$, $\pi_i(g) \geq \pi_i(g_i', g_{-i})$, and $g$ is called Nash stable. It is easy to see that the concept of Nash stability is too weak a concept for modeling network formation when links are bilateral. For instance, the empty network is always a Nash network, regardless of the payoff structure. Moreover, any network where no player could gain by severing some links is a Nash network, regardless of how attractive it might be to add additional links (Bloch and Jackson (2005)). Thus, some refinement is necessary, and different directions must be proposed.

### 2.5.2 Adding uncertainty about unknown private demands

Until now we have examined how incentives for oligopolists to share information change depending upon the network structure when there is uncertainty about an unknown common demand intercept. As is well known, if uncertainty reflects unknown private demands, our results will be significantly affected. Let's consider this case. The game is played in the same way as in the previous sections. Now a market consists of the firms, each producing a heterogeneous product and each firm faces an individual demand shock. The market demand is still linear, namely

$$p_i = a - b_i q_i - \sum_{j \neq i} b_j q_j + u_i, \qquad a, b_i > 0,\ b_i > |b_j|, \forall\ j \neq i \tag{2.14}$$

where $p_i$ is the price and $q_i$ the amount of product $i$ produced. Since $b_j$ can be positive or negative the two products can be substitutes or complements, and since $b_i > |b_j|, \forall\ j \neq i$, "cross effects" are dominated by "own effects". The closer the coefficients $b_i$ and $b_j$ are to each other, the less differentiated the two products are. Assuming a constant $a$ is without loss of generality, since $u_i$ captures heterogeneity in the intercept of the function.

The market demand is stochastic, i.e., $u_i$ is normally distributed with mean zero and variance $\sigma_i$.[8] The signal observed by firm $i$ is $x_i$. We continue to assume that $x_i = u_i + e_i$, $e_i \sim N\ (0, m_i)$, where $Cov(e_i, e_j) = 0, \quad i \neq j;\ Cov(u_i, e_j) = 0 \ \ \forall\ i, j; \quad Cov(u_i, u_j) = 0, \quad i \neq j.$ We can also allow asymmetry in network structure. We can reasonably expect that the analysis requires the use of numerical rather than analytical methods since asymmetry in network structure and heterogeneity in oligopoly cause great complexity and the calculations are likely to become cumbersome.

### 2.5.3 Some other issues

Another interesting issue concerns spillovers. Our model does not accommodate direct spillovers across the collaborative links of firms, since, by assumption, information transmissions happen simultaneously. We can study the incentives for information sharing under a correlated signals assumption. By doing so, we may vary the amount of initial correlation

[8] This environment is called a "private values" problem in the auction literature. (Gal-Or (1986))

among the signals to investigate how various degrees of initial correlation affect the incentives for network formation and information sharing.

Another extension would be to investigate whether the incentives for network formation and information exchange are affected by other sources of uncertainty in the market. In particular, stronger incentives may arise if technology rather than demand was stochastic, or if prices rather than quantities are chosen. However, generalizing the analysis to a different class of demand function may require the use of numerical rather than analytical methods.

In the present paper, we have restricted our analysis to an ex ante symmetric environment. We have shown that the resulting outcome may be ex post asmmetric. It would be another interesting direction to analyze an ex ante asymmetric environment where one firm has access to more precise information or enjoys a superior technology.

## 2.6 CONCLUSION

We have developed a simple two stage model of strategic network formation in order to analyze the incentive of firms to share information in an oligopolistic market where firms face an uncertain demand. Before firms observe a private signal they decide whether to form links to other firms in order to exchange information on market situation. After link formation each firm observes its own private signal, and then they transmit their private information to the linked firms simultaneously. After the network formation and information transmission stage, the firm chooses its level of output. Our interest has been in the interaction between the incentives of firms to collaborate for information sharing (and the resulting network structure) and the market uncertainty. We have derived pure strategy mixed cooperative and noncooperative equilibria that are subgame perfect and stable, and characterized the resulting graphs.

Our analysis has attempted to clarify the nature of collaboration structures that are stable under market uncertainties and different scopes of cooperation or coordination among the oligopolistic firms in the market. An important finding is that even in the setting where firms face an unknown common demand, complete information sharing, no information sharing, and partial-asymmetric information sharing emerge as subgame perfect and pairwise stable equilibrium outcomes under Cournot competition, if firms make a pairwise decision in the decentralized environment. This result contrasts with the existing literature on centralized information sharing. The result is both interesting and important, since pairwise stability captures a relevant and natural equilibrium state in the analysis of the decentralized oligopolistic market. Finally, the unique strongly stable equilibrium outcome involves no information sharing and the resulting empty network. This illustrates that information sharing among firms facing an uncertatin environment critically depends on the coaltion structure.

# 3.0 A WAR OF ATTRITION IN NETWORK FORMATION

## 3.1 MOTIVATION

"$\cdots$ *the individual, when isolated, is not self-sufficing; and therefore he is like a part in relation to the whole. But he who is unable to live in society, or who has no need because he is sufficient for himself, must be either a beast or a god* $\cdots$"

(Aristotle, "Politics")

The purpose of this paper is to investigate the strategic behavior of agents when they face a decision on the formation of relationships. Networks of relationships play a critical role in the wide set of economic and social situations. For instance, personal contacts play important roles in obtaining information about job opportunities. Such networks of relationships also underlie the trade and exchange of goods in non-centralized markets, the provision of mutual insurance in developing countries, R&D and collusive alliances among corporations, and international alliances and trade agreements.[1] Given the prevalence and importance of network structures, the literature on the formation of networks among agents addresses various questions: Some examples are how such network structures are important in determining the outcome of economic interaction, how we predict which networks are likely to form when agents have the strategic discretion to choose their connections, and how efficient the networks are. These are fundamentally and theoretically important, but there still remain essential questions to be addressed: who initiates the link of relationship and pays for it, and what behaviors agents show in the process. The current paper addresses these

[1] Refer to Jackson (2004) for an excellent survey.

questions, as well as the traditional questions specified above. For this purpose, we consider a somewhat specific situation where links or relationships among agents have characteristics of public goods.

Public goods and activities give a positive value to everyone, but their cost is borne entirely by the individual agent performing them. Therefore everyone has strong incentives to free ride by letting someone else do it. Many researchers have shown that this kind of economic environment is well described and analyzed by applying a war of attrition among agents. (See, Bilodeau, M. and Slivinski, A. (1996) and Bliss, C and Nalebuff, B (1984)). Many social and economic relationships that benefit both the corresponding parties are efficiently established by the effort of a single initiative party. Inviting new neighbors for a dinner and links between politicians and businessmen can be examples. Under this consideration, we observe that links among economic agents in Bala and Goyal's (2000) two-way flow model have such properties of public goods among linked agents. Once relationships are established by someone, others can use this network of relationships freely as long as they are parts of the given network. However, the costs of link formation are incurred only by the agents who initiate the links. So each agent strongly wants to wait for the others to initiate a link to him. This kind of a waiting situation allows us to analyze network formation in the framework of a war of attrition where the strategy of an agent is the specification of the set of waiting times among agents with whom he plays the game. The main goal of this paper is to understand how and which network structures emerge as equilibrium outcomes when links among agents have such properties of a public good and all agents want to wait. Also we investigate who volunteers to initiate the link and pays for link formation in the network formation process, and who would be a center if the star network is formed.

There are many other papers that provide theoretical models of network formation in strategic contexts. This paper is related to Jackson and Wolinsky (1996) in that we pursue the same questions they seriously began to ask. Both studies examine the incentive and the behavior of agents to form a link and characterize the resulting equilibrium structures and their efficiency. While Jackson and Wolinsky's (1996) analysis depends on the cooperative and static model, we choose the noncooperative and dynamic approach. By being 'cooperative' we mean that there needs to be mutual agreement of two corresponding agents to form a link, while by 'noncooperative' we mean that the link can be formed and severed unilaterally and the cost of link formation is incurred only by the agent who initiates the link. Jackson and Watts (2002) and Watts (2001) develop Jackson and Wolinsky's (1996) model into the dynamic network formation situation. In particular, Jackson and Watts (2002) studies the stochastic evolution of networks where agents occasionally form or delete links by mistake, thus they characterized the set of stochastically stable networks as limiting networks. However, our model is noncooperative, applies a war of attrition between agents to dynamic network formation, and excludes the possibility of mistakes. This paper is most closely related to the Bala and Goyal's (2000) noncooperative model in which individual agents can unilaterally form new links and sever the existing links. The possibility of unilateral link formation is quite important in methodology because this allows us to use Nash equilibrium concept in the network formation game unlike the other papers which use the concept of pairwise stablity or strongly stability. Although their model is basically static, Bala and Goyal (2000) also examine a repeated game which focuses on myopic agents' learning as a way to identify equilibria. They study both one-way and two-way flow of benefits (or information). We are especially interested in the two-way flow model where benefits are nonrival and nonexcludable and costs are incurred only by the one who forms a link.

The main contribution of this paper is in the modeling and analysis of the network formation in the context of the war of attrition among agents. Much of the previous literature implicily assumes a bargaining process between corresponding agents. It is a natural human behavior to free ride to enjoy as much utility as they could. It becomes clear the link has properties of public goods. Hence it is plausible to apply a war of attrition to network formation process in order to understand how agents behave when they face a free-riding situation and they have an option to wait. An analysis of this scenario must ask and answer to two simple questions: who wins the waiting game and when does the war of attrition end? We address these questions by deriving a unique equilibrium outcome. The second contribution is in overcoming the coordination problem. The existing literature on the formation of networks including the above papers focuses on the architecture of the equilibrium outcome and discuss the interaction between efficiency and stability. Of course, these questions are fundamental and important. However most of them suffer from the multiplicity of equilibria. Some avoid this problem by choosing the cooperative approach and others are silent at this trivially by adopting the symmetry assumption. Bala and Goyal's (2000) noncooperative model has only two nonempty network structures that are strict Nash equilibria. These are the circle network in the one-way flow model and the center-sponsored star network in the two-way flow model. If the benefit (or information) flows both ways through the link, the center-sponsored star is the unique nonempty strict Nash Equilibrium structure. Note that the center-sponsored star network is an asymmetric equilibrium, while the circle is a symmetric equilibrium where every agent chooses the same action. This raises a question of who initiates the links, pays for link formation, and who should be the center if the asymmetric star network is formed. We can analyze such a coordination problem although the analysis is conducted in the restricted environment. The third contribution is in the introduction of heterogeneity of agents to the network formation literature. The heterogeneity assumption is reasonable even in the network formation settings, and this assumption greatly helps to resolve the problem of multiplicity of equilibria.

Our results are as follows. We find that in each subgame the agent with the higher value of benefit-cost ratio forms a link immediately. Unlike the results of Bala and Goyal's (2000) two-way flow model in which either the center-sponsored star or the empty network is the strict Nash equilibrium, we show that as the set of agents grows bigger, a variety of the equilibrium structures which are minimally connected spanning trees emerge as subgame-perfect equilibrium structures. And we demonstrate that these equilibrium outcomes are usually inefficient. Also we observe that the cost payment for link formation does not depend on whether the agent is the central one or not even when star network emerges as an equilibrium outcome.

The remainder of the paper is organized as follows. Section 2 introduces the model of a war of attrition in network formation context, which allows for a unique subgame-perfect equilibrium outcome. Section 3 analyzes the dynamic network formation game as a finitely repeated war of attrition, characterizes the subgame-perfect equilibrium structures and their welfare implications. We provide a simple example in section 4. Section 5 concludes.

## 3.2 THE STAGE GAME: A WAR OF ATTRITION

Consider a two-agent version of the war of attrition which will arise as a subgame of the dynamic game. Let two typical agents be $i$ and $j$. We assume each agent possesses some information which the other can use via the formation of costly link. Let $b_i \in B, \forall\ i$, denote the utility flow of agent $i$'s information to himself and the other agent. We assume information (or benefit) flows both ways. Suppose that until the link is formed, agent $i$ gets utility $b_i$, and that once the link is formed to agent $j$, his utility flow from that time forward is $b_i + b_j$. We let $C_i > 0$ be the present discounted net cost to agent $i$ if he forms the link. Agent $i$ discounts the future utility at rate $\rho_i$. We assume without loss of generality that $C_i < b_i \ / \ \rho_i$ for all $i$. Assume also that every agent's cost, benefit, and the discount rate are common knowledge. We add a dynamic element by allowing individuals the possibility of waiting for some time before volunteering. The point of waiting is to let someone else volunteer first, but waiting can be costly because until someone performs the service, no one can enjoy its benefits.

In the reduced strategic form of this kind of waiting game, a pure strategy is a time $t \in [0, \infty)$ at which to volunteer if no one else has. Once anyone volunteers, the link is formed and the game ends. In the extensive form, a strategy must specify whether an agent would volunteer at any $t$ in the eventuality that no one had volunteered up to that point. In particular, strategies must specify what an agent would do even at times that would be reached with probability zero given the other's strategy. For example, if someone volunteers with probability one at $t = 0$, strategies must specify what the other would do if any $t > 0$ is reached. Now, if the link is formed by the other agent at time $t$, agent $i$'s payoff would be

$$F_i(t) = \frac{b_i}{\rho_i} + \frac{b_j}{\rho_i} e^{-\rho_i t}.$$

Suppose that the cost of forming a link includes two separate components: a one-time utility cost $f_i$, and a net utility flow of cost $c_i$. The total cost of volunteering at time $t$ is therefore

$$C_i(t) = f_i e^{-\rho_i t} + \frac{c_i}{\rho_i} e^{-\rho_i t}.$$

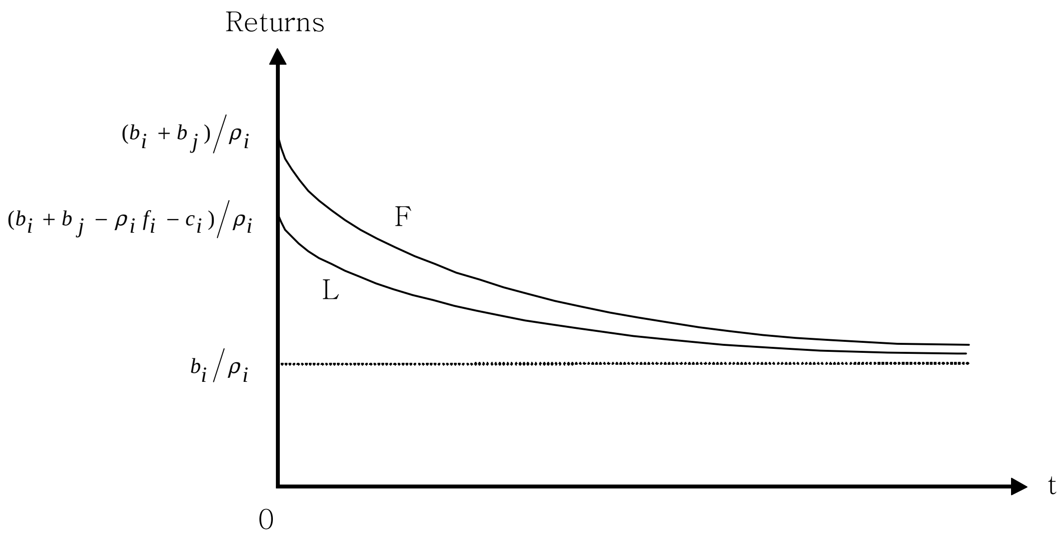


Figure 5: The return functions

Readers might be uncomfortable with the assumption that the initiator pays all link costs and that both agents benefit from the link, since many examples of two-way flow networks have two-sided link costs (such as social or friendship networks) whereas one-sided link costs are often associated with one-way flow networks (such as visiting a website and downloading a paper). However, the existence of the net utility flow of cost, $c_i$, is not essential. What is important is that there must exist at least a one-time utility cost, $f_i$, to initiate the relationship among strangers. Inviting newly arrived neighbors to dinner can be an example. By adding the continuous maintenance cost, we try to capture the idea that the initiator usually pays more attention to maintenance of the relationship or has slightly higher start-up costs.

Assume that $b_j - c_i - \rho_i f_i > 0$ for all $i$, so everyone would still rather form a link than do completely without it. Let $L_i(t) = F_i(t) - C_i(t)$ denote the payoff to $i$ if he is the one to volunteer to form a link at time $t$, and let $S_i(t)$ be $i$'s payoff if he volunteers simultaneously with the other, where $L_i(t) \leq S_i(t) < F_i(t)$. Then $S_i = \lim L_i(t) = \lim F_i(t) = b_i \ / \ \rho_i$ is his payoff if no one ever volunteers.

Hendricks et al. (1988) provide a complete characterization of all the subgame-perfect equilibria for this class of games. Using their results, we can construct the following proposition.

**Proposition 10.** *There is a subgame-perfect equilibrium outcome in which anyone of the two agents establishes a link immediately.*

*Proof.* $L_i(0) > S_i$ for all $t$ and $L_i(0) \geq F_i(t)$ for some $t \in [0, \infty)$ by assumption, so condition (b) (i) and (ii) of Theorem 1 of Hendricks et al. (1988) are satisfied. Therefore there is an equilibrium where someone forms a link immediately with probability one and the other waits with probability one. And there is no $t$ for which $L_i(t) < S_i$ for some $i$, and $L_i(0) > S_i$ for all $i$, so condition (b) of Theorem 4 of Hendricks et al. (1988) is satisfied. Therefore there exists a subgame-perfect equilibrium in which any of the two players establishes a link immediately with probability one. $\square$

Intuitively speaking, since $L_i(0) > S_i$, an agent is always better off forming a link immediately than doing completely without it if no one else volunteers. So suppose no one else ever volunteers, then it is optimal for $i$ to volunteer immediately, and since $i$ volunteers immediately, no one else can do better by changing their strategy. Verifying subgame-perfection requires defining strategies for the extensive form game. If agent $j$ adopts the strategy of waiting forever, then the best agent $i$ can do is to form a link with probability one if any $t$ is ever reached. Given the strategy of agent $i$, agent $j$ has no incentive to change his strategy. So this constitutes a subgame-perfect equilibrium. Therefore, in our setting the empty network does not emerge as a subgame-perfect equilibrium outcome.

The classic war of attrition exhibits multiple equilibria. Many papers show that the "perturbed" version of a war of attrition exhibits a unique equilibrium. Exit failures, hybrid all-pay auctions and time limits all represent perturbations to the war of attrition. In the present paper, in order to avoid the presence of multiple equilibria, we impose a common time limit $T$ and this yields a finite-horizon war of attrition.[2] Allowing $T \to \infty$ eliminates the perturbation. Let's consider the game in which agents have a finite time horizon $T$.

[2] We find that in a very general complete information game, multiplicity of equilibria depends crucially on the assumption that agents have an infinite horizon. If we assume instead that agents have a finite horizon, the game has a unique subgame-perfect equilibrium outcome. This remains true even when the time horizon tends to infinity. Since the infinite horizon game is the limit of the set of finite horizon games, this outcome stands out even in infinite horizon games.

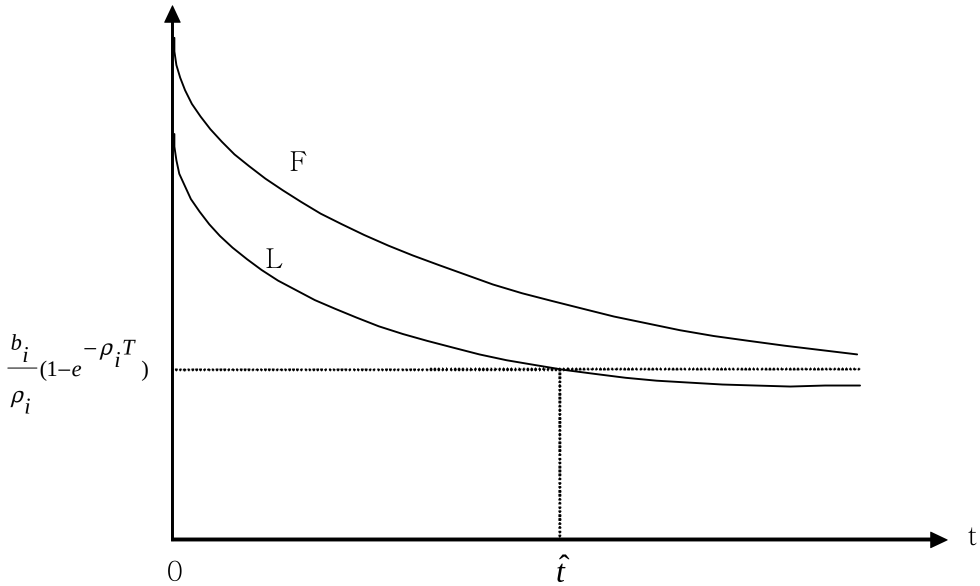


Figure 6: The modified return functions

Then the relevant payoffs become

$$
\begin{aligned}
F_i(t) &= \frac{b_i}{\rho_i}(1-e^{-\rho_i T}) + \frac{b_j}{\rho_i}(e^{-\rho_i t} - e^{-\rho_i T}) && (3.1)\\
C_i(t) &= f_i e^{-\rho_i t} + \frac{c_i}{\rho_i}(e^{-\rho_i t} - e^{-\rho_i T}) && (3.2)\\
L_i(t) &= F_i(t) - f_i e^{-\rho_i t} - \frac{c_i}{\rho_i}(e^{-\rho_i t} - e^{-\rho_i T}) && (3.3)\\
\text{and, } S_i &= \frac{b_i}{\rho_i}(1-e^{-\rho_i T}) && (3.4)
\end{aligned}
$$

Then there exists a $\hat{t_i} < T$ such that $L_i(t) \geq S_i$ for all $t \leq \hat{t_i}$ and $L_i(t) < S_i$ for all $t > \hat{t_i}$.

When agents have a finite horizon, there exists a point in time, $\hat{t_i}$ for each $i$, when it is no longer worth forming a link because the benefits will be felt for only a short time after. Beyond that point, it is a dominant strategy for agent $i$ to never volunteer. Solving for the $\hat{t_i}$ for which $L_i(t) = S_i$ we find

$$\hat{t_i} = T - \frac{1}{\rho_i}\ln\left(\frac{b_j - c_i}{b_j - c_i - \rho_i f_i}\right). \tag{3.5}$$

**Proposition 11.** *Generically, when both the agents have a finite horizon, the unique subgame-perfect equilibrium outcome is where the agent with the larger value of $\hat{t_i}$ establishes a link immediately.*

*Proof.* Without loss of generality, suppose $\hat{t_i} > \hat{t_j}$. No one will form a link at any time $t > \hat{t_i}$. At any $t \in (\hat{t_j}, \hat{t_i}]$ player $i$ can deduce that $j$ will never volunteer. But since $L_i(t) \geq S_i(T)$, $\forall\ t \in (\hat{t_j}, \hat{t_i})$, $i$'s subgame-perfect strategy must be to form a link if any such $t$ is reached. Now consider any time $t \in (\hat{t_j} - \varepsilon, \hat{t_j}]$. Both $i$ and $j$ might volunteer, but $j$ will prefer to wait for $i$ to volunteer at $\hat{t_j}$. Then in any subgame-perfect equilibrium, $i$ must volunteer at any $t \in (\hat{t_j} - \varepsilon, \hat{t_j})$. We, then, analyze at any $t \in (\hat{t_j} - 2\varepsilon, \hat{t_j} - \varepsilon]$, $t \in (\hat{t_j} - 3\varepsilon, \hat{t_j} - 2\varepsilon]$, and so on. By backward induction, the subgame-perfect equilibria are where player $i$ forms a link at any $t \in [0, \hat{t_i})$, randomizes between voluntary link formation and waiting at $t = \hat{t_i}$, and never forms a link at any $t > \hat{t_i}$, and player $j$ always waits. $\square$

Therefore, we observe that the subgame-perfect outcome is where the agent with larger benefit/cost ratio forms the link immediately. This implies that, if all other parameters were identical across the agents, the agent with the lowest cost, the agent with the highest gain, or the most impatient agent volunteers immediately.

### 3.3 THE SEQUENTIAL NETWORK FORMATION GAME

Now we embed the stage game of the two-agent version of a war of attrition into the dynamic network formation game. Let $\mathrm{N} = \{1, 2, \ldots, N\}$ be the finite set of risk-neutral agents and let $i$ and $j$ be typical members of the set. At each stage, two agents are randomly drawn from the set N, and each of these paired agents plays a war of attrition to free ride by waiting as in the previous section. If the same couple appears repeatedly, the game ends immediately and two new agents are drawn.[3] Time is continuous and each stage game has a common time limit $T$. If nobody volunteers to form a link within time limit $T$, the stage game ends and the newly matched agents play the defined game in the subsequent stage. After each

[3]The number of the stage games in the whole game is $\binom{N}{2}$. In this sense, the matching protocol is random permutation. For example, when $N = 3$, the number of the stage games in the whole game is three and a sequence ((1,2)-(2,3)-(1,3)) is one possible realization.

stage game, the resulting relationships among agents are represented by graphs whose nodes are identified with agents and whose links capture the pairwise relations.

Suppose agents $i$ and $j$ are matched at stage $\tau$. We denote this by $(i,j)^\tau$. Let $g^\tau = (g^\tau_{ij}, g^\tau_{ji})$ where $g^\tau_{ij} \in \{0,1\}$. We say agent $i$ ($j$) initiates a link (also pays the cost for it) to agent $j$ ($i$) at stage $\tau$ if $g^\tau_{ij} = 1$ ($g^\tau_{ji} = 1$). With two-way flow assumption, the link $g^\tau_{ij} = 1$ enables both $i$ and $j$ to access each other's information. Let $t^\tau = (t^\tau_i, t^\tau_j)$ where $t^\tau_i \in [0,T]$ denotes the waiting time before agent $i$ initiates a link to $j$ at stage $\tau$. We say agent $i$ forms a link immediately at stage $\tau$ if $t^\tau_i = 0$, and agent $i$ never forms a link or never volunteers to form a link at stage $\tau$ if $t^\tau_i = T$. A history $h^\tau = (h^0, ((j,k)^1, g^1, t^1), \ldots, ((j,k)^{\tau-1}, g^{\tau-1}, t^{\tau-1}))$ begins with the null history and further records for each of the first $\tau - 1$ stages. Denote by $H^\tau$ the set of all such histories and define $H = \bigcup_{\tau=1} H^\tau$. We call $g(h^\tau) = \bigcup_{\tau=1} g^\tau$ the existing graph at stage $\tau$. We will denote $g(h^\tau)$ as $g$ whenever convenient. The set of all possible graphs on N is denoted by G.

Consider newly matched agents $i$ and $j$ at stage $\tau$. If there is a linked path between them and if there is no decay in information (benefits) transmission, they consent to end the stage game immediately since this causes cost only. Otherwise, given the graph structure at the beginning of stage $\tau$, they calculate the benefit and cost of the link formation. Let $b_i(g)$ denote the utility flow from that time forward by constructing a link to $i$ at period $\tau$. For instance, if agents 1, 2, and 3 are linked anyhow and if there is no decay then agent $i$ can enjoy utility flow from that time forward by forming a link to agent 1 by $b_1(g) = b_1 + b_2 + b_3$. The costs of link formation are incurred only by the agent who initiates the link. Let $C_i > 0$ be the present discounted net cost to agent $i$ if he forms the link. This consists of two parts: a one-time utility cost $f_i$ and a net utility flow of cost $c_i$. Each agent can sever the existing link unilaterally at any time without any cost. Recall that the game is played noncooperatively. The game is a complete information game where everyone's informational value, cost of link formation, and discount rate are common knowledge. Then the relevant payoff structure facing agent $i$ at period $\tau$ becomes

$$\begin{aligned}
F_i^\tau(t,g) &= \frac{b_i(g)}{\rho_i}(1-e^{-\rho_i T}) + \frac{b_j(g)}{\rho_i}(e^{-\rho_i t_j^\tau} - e^{-\rho_i T}) && (3.6)\\
C_i^\tau(t) &= f_i e^{-\rho_i t_i^\tau} + \frac{c_i}{\rho_i}(e^{-\rho_i t_i^\tau} - e^{-\rho_i T}) && (3.7)\\
L_i^\tau(t,g) &= F_i(t,g) - f_i e^{-\rho_i t_i^\tau} - \frac{c_i}{\rho_i}(e^{-\rho_i t_i^\tau} - e^{-\rho_i T}) && (3.8)\\
\text{and, } S_i^\tau &= \frac{b_i(g)}{\rho_i}(1-e^{-\rho_i T}) && (3.9)
\end{aligned}$$

When agents have a finite horizon, there exists a point in time, $\hat{t_i}(g)$ for each $i$, when it is no longer worth forming a link, as $T$ is large enough so that the benefits will be felt for only a short time after. Beyond that point, it is a dominant strategy for the player $i$ to never volunteer. Solving for the $\hat{t_i}(g)$ for which $L_i^\tau(t,g) = S_i^\tau$ we find

$$\hat{t_i}(g) = T - \frac{1}{\rho_i}\ln\left(\frac{b_j(g)-c_i}{b_j(g)-c_i-\rho_i f_i}\right). \tag{3.10}$$

Note that link formation is immediate among agents on the equilibrium path. Therefore agents' payoffs and total welfare depend only on the resulting graphs in equilibrium. Recall that the empty network does not emerge as a subgame-perfect equilibrium outcome. Suppose there has been no link before the last stage game starts. As is studied in the previous section, it is a subgame-perfect outcome that one of the two players forms a link immediately in last stage game. Therefore, the empty network does not appear as the subgame-perfect structure.

**Proposition 12.** *In each subgame, when both the agents have a finite horizon, the agent with the larger value of $\hat{t_i}(g)$ establishes a link immediately.*

We characterize the structures of the equilibrium outcomes and their properties.

**Corollary 13.** *1. When $T$ is large enough, minimally connected spanning trees emerge as subgame-perfect equilibrium outcomes if there is no decay in information (benefit) transmission, and the complete network emerges as subgame-perfect equilibrium outcome if there is no spillovers.*
*2. Star networks, which crucially depend on the sequence of the matching, are just a subset of the possible subgame-perfect equilibrium outcomes. Furthermore, cost payment does not depend on whether agent is the center or not even if star networks emerge.*

*Proof.* Recall that the empty network does not emerge as a subgame-perfect equilibrium outcome. And note that there are at most $(N-1)$ links in the resulting graph if $g$ is connected. If $g$ is not minimally connected, there exists an agent who can delete a link and still have a path with every other agent, so $g$ is not subgame-perfect Nash equilibrium outcome. Suppose that $g$ is a nonempty subgame-perfect Nash network, and suppose it is not connected. Since $g$ is nonempty there exists a component $g(C)$ such that $|g(C)| \geq 2$. Choose $i \in g(C)$. As $g$ is not connected, there exists $j \in \mathrm{N}$ such that $j \notin g(C)$. Case (1): If $j$ is an isolated agent, then the payoff to agent $j$ from a link with $i$ is bigger than present payoff, which violates the hypothesis that agent $j$ is choosing a best response at every stage game. Suppose, at stage $\tau$, agent $j$ meets agent $i \in g(C)$. Suppose that $\hat{t_i} < \hat{t_j}$. (This is true since, if $j$ forms a link to $i$, he can use all benefits from the component $g(C) \ni i$. If $\hat{t_i} > \hat{t_j}$, then $i$'s problem belongs to Case (2) to be analyzed later). The only reason, player $i$ must expect, for not forming a link is that there appears an agent $k \notin g(C)$ who will form a link to $j$ first and then to one member of $g(C)$. (If $k \notin g(C)$ forms a link to an existing component $g(C)$ first, then problem returns to original situation). Even if the expectation that $k$ will form a link to $j$ is correct (now $j$ is not an isolated agent), we can show that $j$'s waiting is not best response. Case (2): Suppose instead that $j$ lies in another component $g(D)$ where $|g(D)| \geq 2$. Suppose, at stage $\tau$, agents $i$ and $j$ are drawn. Suppose, without loss of generality, that $|g(C)| \geq |g(D)|$ and $\hat{t_i} < \hat{t_j}$. If the matching is the last ordered one, the player with bigger $\hat{t}$, $j$, will form a link immediately. Now suppose that there is a positive probability that agents $i \in g(C)$ and $k \in g(D)$ will meet each other at stage $\tau + 1$. For the strong result, assume that agents $i$ and $k$ meet with probability one. In this situation, player $j$ has a strong incentive to wait even if $\hat{t_i} < \hat{t_j}$, not because he expects $i$ to initiate the link, but because he waits for player $k$ to form a link to $i$ at the stage $\tau + 1$ at the cost of non-accessibility to benefits of component $g(C)$ during time $T$. But we can find that, when $T$ is large enough, $j$'s immediate link formation incurring the cost is beneficial to both of them. This contradiction implies that $g$ is connected. Therefore, $g$ is minimally connected. □

Note that $\hat{t}$ is determined endogenously through the link formation process. Informally we can describe the network formation process as follows. When the first two players meet each other in the first stage, they play a war of attrition to free ride, and one of the two players with the higher $\hat{t_i}$ forms a link immediately. Then there are two subcases in the next stage. One is the matching of two isolated players. This is the same as the first matching of the game. The other subcase is the matching where one of the two players has been linked from the previous stage and the other one is an isolated player. In this case it is very probable that the isolated player has the higher $\hat{t_i}$ and, therefore, forms a link immediately to the other player because he can enjoy all informational value of the existing component where the opponent player is included. This process continues until the minimally spanning tree completes.

Usually these equilibrium outcomes are inefficient. The efficient network structure in our setting that gives the highest welfare is the center-sponsored star network in which only the agent with the lowest link formation cost plays the central node. But this is just a special case of the minimum spanning tree. The example below illustrates how the players behave, how and which network structure arise in equilibrium. Also this example shows how hard it is for the efficient outcome to appear in our model.

### 3.4 AN ILLUSTRATIVE EXAMPLE

To illustrate ideas presented in the previous sectoins, consider the following example. Assume that N $= \{1, 2, 3\}$, $T = 10$,

$$\begin{aligned} b_1 &= 4,\ c_1 = 1,\ f_1 = 2,\ \rho_1 = 0.5, \\ b_2 &= 4,\ c_2 = 1.5,\ f_2 = 2,\ \rho_2 = 0.5, \\ b_3 &= 5,\ c_3 = 1.5,\ f_3 = 2,\ \rho_3 = 1. \end{aligned}$$

Assume that information (benefit) flows both ways and there is no decay in information transmission (i.e., complete spillovers). In the example the sequence of matching is a random permutation among $\{(1, 2), (2, 3), (1, 3)\}$. Suppose that the first matched pair is $(1, 2)$. Then the order of the matching is $((1, 2), (1, 3), (2, 3))$ with probability .5 (say, Game A) and

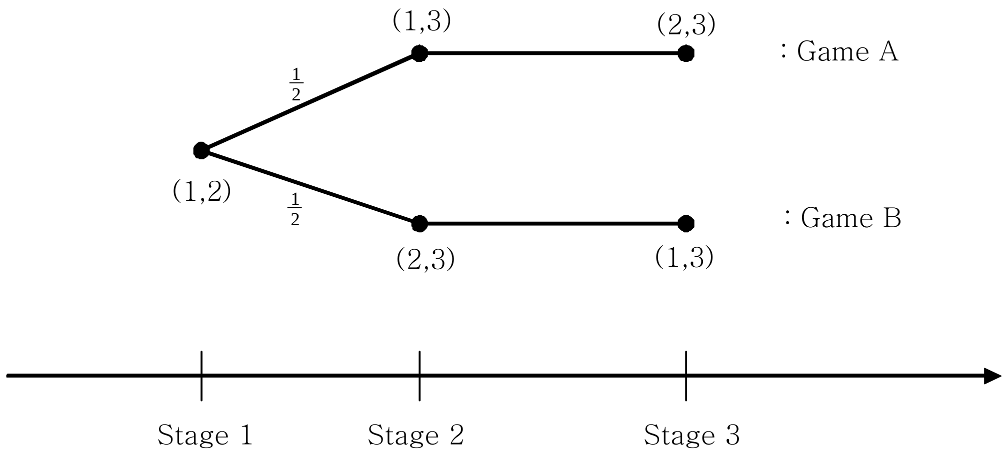


Figure 7: The matching sequence of the game

$((1,2),(2,3),(1,3))$ with probability .5 (say, Game B). Figure 7 shows the sequence of the game. Here each node represents the stage game (the two-agent version of a war of attrition) between the two corresponding agents. In each stage game both the agents try to free ride by letting the other initiate a link and start waiting.

Note that if two links were already formed in the previous stages, the last matched pair has no incentive to form a link, because it causes only cost without additional benefit. So they do not volunteer at all in their meeting, and the game ends. If one link was formed in the previous stages, one of the two agents in the last stage game will form a link immediately. Therefore, in case of three agent network formation game, the outcome structure has at most two links under our assumptions. Note also that if the best choice for an agent is to volunteer, then he will volunteer as soon as possible, because waiting is also costly. For an easy explanation, we use another tree. Figure 8 shows only Game A which has the sequence $((1,2),(1,3),(2,3))$, and each node represents the stage game between the two randomly chosen agents. Figure 8 shows subgame-perfect equilibrium outcome in Game A in which the number indicates the agent who (immediately) initiates a link and pays for the link in each stage game, and x means nobody volunteers to form a link. For instance, consider the first stage of the game played between agents 1 and 2. There are only three possible

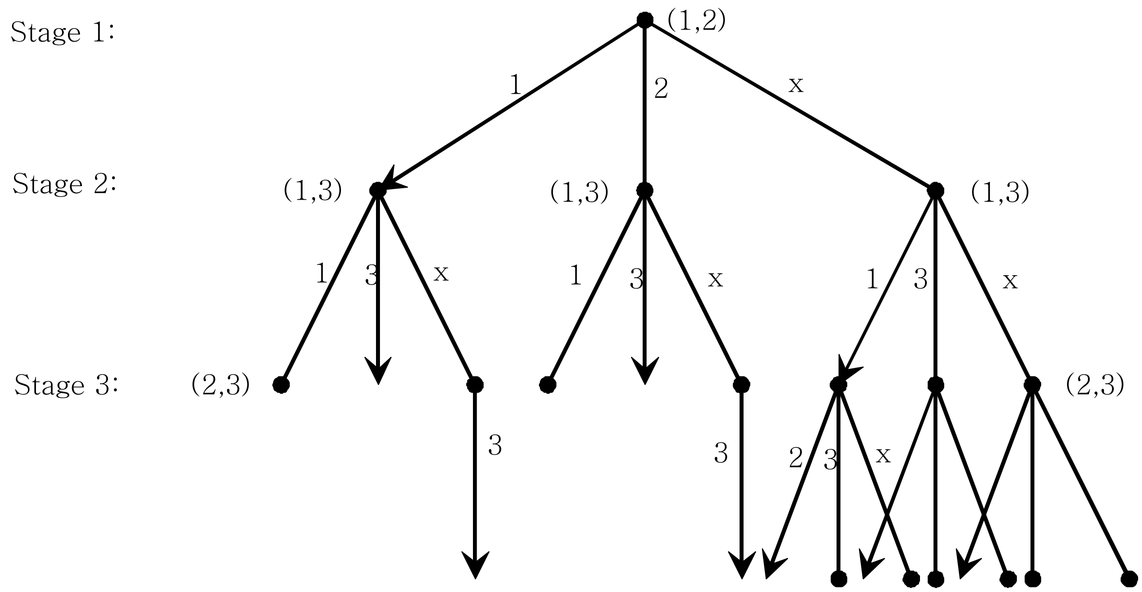


Figure 8: Game A

outcomes. One possible subgame-perfect equilibrium outcome is where agent 1 volunteers immediately, the other is where agent 2 does, and the last case is where nobody volunteers. This applies to every other stage game (node). It is important to remember that each stage game has only three possible subgame-equilibrium outcomes where one of the two agents immediately form a link when they are not linked, or nobody forms a link. Recall that we consider the no decay case only.

Since agents observe the actions of the other agents and the resulting graph among them in each stage, they can calculate their expected payoffs from voluntary link formation and their $\hat{t_i}(g)$ before choosing their actions. We know that in each stage game the agent with higher $\hat{t_i}(g)$ forms a link immediately. We apply backward induction to find out the subgame-perfect equilibrium outcome. Arrows in Figure 8 indicate the final equilibrium outcome. Consider the third stage of the game played between agents 2 and 3. If two links are formed at the previous stages, they don't move. Given that there is a link formed at the first stage and no link at the second stage, for agent 3, forming a link to agent 2 causes additional utility flow $b_1 + b_2$ at the cost of $C_3$, while, for agent 2, establishing a link gives additional utility flow $b_3$ and the cost $C_2$. This consideration gives $\hat{t}(g) = (\hat{t_2}(g), \hat{t_3}(g)) = (9.3271, 9.6323)$. Therefore agent 3 forms a link immediately at the third stage. We can examine the other nodes (stage games) using the same reasoning. Let's consider agent 3 in the second stage. Agent 3 will again form a link immediately if there is a link between agents 1 and 2 at the first stage, since agent 3 knows that he should form a link at the third stage otherwise. If there is no link at the first stage, agents 1 and 2 will form a link at the second and third stage respectively in Game A. Now consider the first stage of the game. Agents 1 and 2 know that either Game A or Game B will be played with even probability after their play. Then they estimate the expected benefits and costs by applying backward induction and they get $\hat{t}(g(h_0)) = (\hat{t_1}(g), \hat{t_2}(g)) = (9.5987, 9.5538)$. Therefore agent 1 forms a link to agent 2 immediately. Now we need to check whether agents 1 and 3 at the second stage have any incentive to change their actions after they observe the first stage game between agents 1 and 2. As soon as they observe that agent 1 formed the link to agent 2 at the first stage, agent 3 will immediately establish a link to agent 1 in order to enjoy increased benefits of agents 1 and 2. This constitutes subgame-perfection. The resulting architecture of the subgame-perfect equilibrium outcome is minimally connected spanning tree.[4] Figure 9 shows the resulting graphs. X indicates that the agent closer to X is the one who initiates the link and pays for it.

We have analyzed only a part of the whole game where the first matching is $(1, 2)$. This means that we will have the resulting graphs shown above with probability $1/3$. Although this is very simple example, we can find important implications. First, we can characterize the equilibrium network structures. In this example, since we have only 3 agents, the structure can be a star or a line. We can, in general, say that the equilibrium outcome is minimally connected spanning tree. We easily check that two resulting networks are inefficient. The only efficient graph in our example is the star network in which agent 1 initiates all links. Second, we show explicitly who initiates the link, who pays for the link formation, and who should be the central agent if the star network emerges as an equilibrium outcome. We find

[4]If there is no spillover in information transmission, i.e., if agents can enjoy the benefit of link formation from the directly linked partners only, the resulting graphs would be the complete networs as long as there is the higher benefit than the cost from the link formation. However, the payment structures are different from each other.

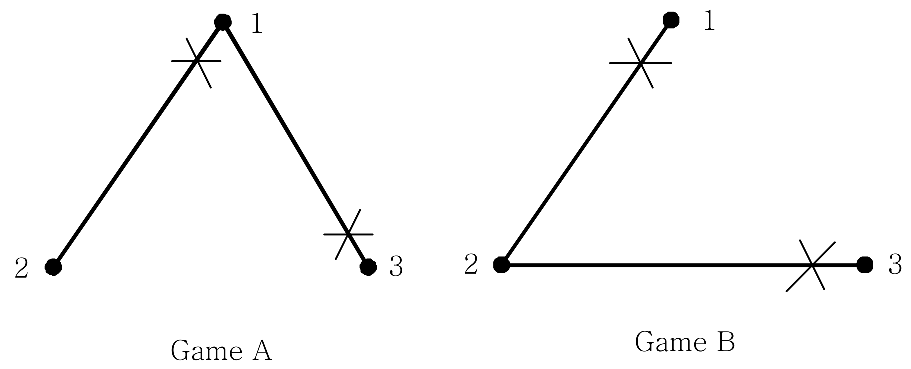


Figure 9: The resulting graphs

there is no relationship between being a central agent and paying the cost. In Game B two peripheral agents form the links and pay for them, while in Game A one link is initiated by the central agent and the other is by peripheral agent. Furthermore, we illustrate that the sequence of the matching plays a critical role in our dynamic model. As shown above, the unique resulting graph critically depends on the matching sequence. Watts (2002) also mention this observation. Watts (2002) says that, even though their approach is cooperative one, the only way for the star network to form is if the agents meet in a particular pattern. For $N = 4$, a star will form if agents meet in the order $\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$, but not if the agents meet in the order $\{(1,2),(3,4),(1,3),(1,4),(2,3),(2,4)\}$. We also observe the same phenomenon. If the benefits and costs are very similar each other and if the sequence of matching is given as $\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$, then periphery-sponsored star network with agent 1 being the center can appear as a resulting graph but not with the sequence $\{(1,2),(3,4),(1,3),(1,4),(2,3),(2,4)\}$. The agent associated with many links (i.e., the agent with high degree) is the one who moves many times in earlier stages than others. When $N = 3$, the central agent is always the one who moves first two stages in a row. When $N = 4$, it is necessary that the central one must be drawn first three times in a row. Here we conclude that the important role of the matching sequence is a common property of sequential network formation.

## 3.5 CONCLUSION AND DISCUSSION

This paper has tried to understand how economic agents behave and which network structures emerge depending on the behaviors of agents when the links among agents have the properties of the public goods. It is well studied that network structures play a significant role in determining the outcome of many important economic relationships. Therefore it is crucial to know how network configurations arise and which network structures emerge. There, however, still remain important questions to be addressed: Who will volunteer to initiate relationships that everyone would rather let someone else do? What will be the resulting network structure in this situation? To answer to these questions, we have modelled the search for a volunteer as a war of attrition in which everyone is tempted to just wait for someone else to do it.

The unique subgame-perfect equilibrium outcome of the dynamic network formation game is that in each stage, ceteris paribus, the agent with the higher benefit/cost ratio of performing the link, or the agent with higer impatience establishes a link immediately. Unlike the results of Bala and Goyal (2000) in which either the center-sponsored star or the empty network is the strict Nash equilibrium outcome, we show that as the set of the players grows bigger, a variety of the equilibrium structures which are minimally connected spanning trees can emerge as subgame-perfect equilibrium structures. And we demonstrate that these equilibrium outcomes are usually inefficient. Also we observe that cost payment for link formation does not depend on whether the agent is the central one or not, even if star networks happen.

In this paper we have applied the war of attrition as a specific behavior among agents in the network formation. We can think of alternative network formation processes such as a bargaining among agents. It would be an interesting question to ask what happens if agents use bargaining process in the network formation. And it would not be a trivial topic to compare the resulting equilibrium outcomes and their welfare implications. One natural extension would be to introduce uncertainty to the model. This paper assumes complete information, which is perhaps unrealistic. If agents have private information on link formation cost or discount factor, we might not have the result of immediate network construction result. Finally, it would be interesting to incorporate the endogenous matching process into the model. We leave these extensions to future research. to the question of whether Folk Theorem is applicable in this setting.

# 4.0 SOCIAL NORMS AND TRUST AMONG STRANGERS

## 4.1 MOTIVATION

One of the most important issues in economics is how to sustain cooperative equilibrium when players have an incentive to deviate from cooperation. The trust game illustrates this situation well. An investor has the option of choosing a costly trusting action by giving money to a trustee. The trustee is then informed of the investor's transfer and can honor the investor's trust by sharing the monetary payoff generated by the investor's transfer. If the investor gives money to the trustee and the latter shares the proceeds of the investment, both players end up with a higher monetary payoff. However, the trustee also has the option of violating the investor's trust. As sharing the proceeds is costly for the trustee, a selfish trustee will never honor the investor's trust because the investor and the trustee interact only once. The investor is therefore caught in a dilemma: if she trusts and the trustee shares, the investor increases her payoff, but she is also subject to the risk that the trustee will abuse this trust. The unique subgame perfect equilibrium prediction is not to trust and not to reciprocate. This game well describes a credit market and the typical transactions between buyers and sellers based on trust in cyberspace.

Economists have long recognized "reputation" as an effective means of enforcing cooperative behavior when there exists an institution to track and disseminate such information, or within a small group where people are intimately familiar with one another's history. These personal enforcement mechanisms are effective only if quick and substantial retaliations are available. The Folk Theorem in the repeated game literature (Fudenberg and Maskin, 1986) provides a formal model of personal enforcement, showing that any mutually beneficial outcome can be sustained as a subgame perfect equilibrium if the same set of players plays the same stage game ad infinitum.

However, many important transactions in reality are infrequent in nature and many transactions happen among essentially anonymous players. eBay claims more than 100 million registered users who buy and sell millions of item each day on the internet. These transactions are done between strangers who have no contact except through cyberspace. In this case only partial information about a stranger's reputation is available at best, and, therefore, the effectiveness of reputation is far less certain. This observation raises important question about economic behavior. What factors drive the emergence of trust and reciprocity in economic transactions? In order to analyze this question, previous papers consider a random matching model under the most extreme information restriction.

Kandori (1992) analyzes the special economic environment where a deviating agent cannot be directly punished since each agent is anonymous and interacts only at randomly determined times with any other particular agent in the population. Nevertheless, a deviator could be indirectly punished if the deviation were to trigger a contagious reaction that destroyed the social norm of cooperation. If the consequences of the eventual destruction of the norm were sufficiently severe, and if the threat of such a contagious reaction were credible, then the threat might sustain a social norm of cooperation. Kandori (1992) shows that there exist conditions that make such a threat credible in the infinitely repeated prisoner's dilemma with a random matching. Specifically, for any fixed population size, Kandori provides an example of a game in which cooperative repeated game equilibria exist, showing that we can define payoffs for the prisoner's dilemma which allow cooperation in a sequential equilibrium. However, when the population is large the argument applies only to games with extreme payoffs.

Ellison (1994), based on Kandori's model, analyzes two main problems. First, he asks whether cooperation is possible in a sequential equilibrium for general payoffs in the prisoner's dilemma. He reaches the positive answer by introducing a publicly observable random variable. Public randomization devices adjust the expected duration, and hence the severity of the punishment. The second problem is a study of the stability and efficiency of the equilibrium in a world with noise. Kandori observed that a single deviation leads to a complete destruction of cooperation due to the property of contagious reaction. With public randomizations, global stability[1] described by Kandori (1992) is easily achieved. Another problem happens in terms of efficiency. If we have a globally stable equilibrium where the continuation payoffs return to the cooperative level so slowly so that the equilibrium with noise has an expected payoff near zero, stability is not plausible. However, this problem is very hard to overcome. If one player were a "crazy" type and always played the Defect action, the contagious strategies would not support cooperative equilibrium. In a large population, we may want to allow for heterogeneity among players, and the existence of the crazy type may be more appropriate.

This paper makes three contributions. First, we theoretically extend Kandori's (1992) arguments to the trust game. As Ellison (1994) notes, the results of the previous papers heavily rely on the fact that the prisoner's dilemma has a dominant strategy equilibrium. Therefore it would be interesting to know whether the results extend to the other class of games. Kandori (1992) only shows that, *under local information processing*[2], a simple form of community enforcement supports cooperation in the one-sided incentive problem[3]. In this paper, we consider the one-sided incentive problem under the most extreme information restriction. That is, players are completely anonymous in that they can neither recognize nor communicate the identity of any of their past opponent. Therefore, players do not observe the outcomes of games in which they are not involved. We, first, develop the concept of the contagious strategy following Kandori (1992), and we provide the sufficient conditions that support the social norm of trust and reciprocity as a sequential equilibrium. Then we introduce public randomizations to discuss the problem of stability. Our theoretic results show that the existence of a contagious equilibrium critically depends on the level of the outside option for the investor, since high outside option represents the high level of threat in our model.

Second, the results of this paper rationalize many experimental results which show the significant level of trust and reciprocity in the laboratory. Many experimental studies have

[1] Public randomizations are used as a coordination device so that all players can simultaneously return to cooperation. Therefore, after any finite history, the continuation payoffs of the players eventually return to the cooperative level with probability one.

[2] Each player carries a label and the necessary information is transmitted by the players' labels. After actions, their labels are updated depending on their original labels and actions by a given rule. (Kandori, 1992)

[3] A canonical example of the one-side incentive problem is the trust game.

used the trust game to identify trust and reciprocity in an investment setting, and found that trust and reciprocity prevail in the laboratory even under the conditions of complete anonymity and one-shot interaction. They regard these behaviors irrational and explain them with some psychological factors such as fairness, altruism and inequality aversion etc. Another contribution of this study is that we identify trusting and reciprocating behavior as an equilibrium phenomenon without changing the fundamental assumption in economics that individuals act in their own self interest. That is, we show that trust and reciprocity are consistent with a rational choice of selfish players, and investigate how such a cooperative social norm is sustained by self-interested group members in the community.

Finally, we expect that this study will help explain the emergence and prevalence of e-Commerce which is important as the internet develops. The success of e-Commerce can be seen as an equilibrium outcome (i.e., the establishment of the social norm of trust and reciprocity) in the anonymous random matching model based on the social norms and community enforcement. More generally, an understanding of the factors that determine the development of the social norm among strangers is of obvious importance to the design and operation of economic transactions and economic institutions.

The paper is organized as follows. Section 2 describes the formal model of repeated random matching game and exhibits a sequential equilibrium which sustains social norms of trust and reciprocity. Section 3 introduces public randomizations and discuss the problem of stability, and also shows that we can expand the set of equilibria which sustain trust and reciprocity. Section 4 concludes and discusses the contributions.

## 4.2 THE MODEL WITHOUT PUBLIC RANDOMIZATION

In this section we describe the structure of the repeated matching game, define the concept of "contagious equilibrium" based on Kandori (1992) in the infinitely repeated trust game with random matching, and then we present the conditions for the equilibrium to exist.

The set of players $N = \{1, 2, \ldots, 2n\}$ is partitioned into two sets of equal size, the set of investors $N_I = \{1, 2, \ldots, n\}$ and the set of trustees $N_T = \{n+1, n+2, \ldots, 2n\}$. In each period, each investor is matched with a trustee according to the uniform random matching

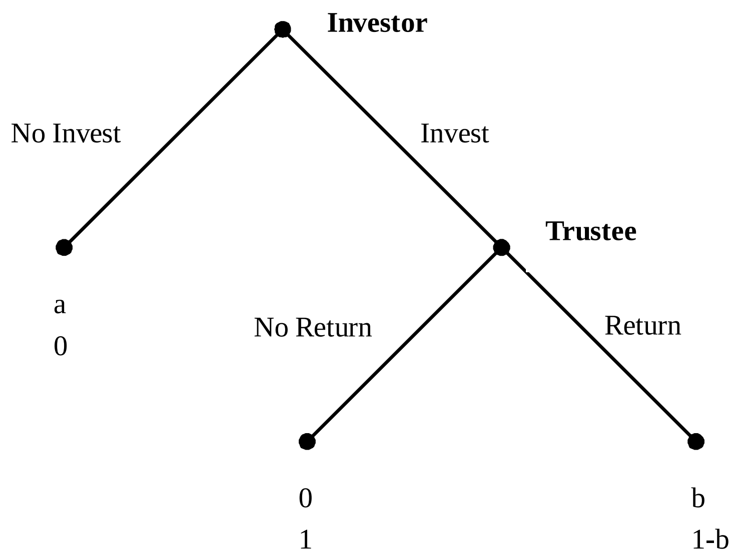


Figure 10: The trust game

rule, and they play the trust game as a stage game. This procedure is repeated infinitely and each player's total payoff is the expected sum of his stage payoffs discounted by $\delta \in (0, 1)$.

In each period, every pair of investor and trustee play the trust game. At the beginning of the game the investor is endowed with one unit of capital. In the first stage, the investor decides whether to invest the capital in the trustee's business or not.[4] If the investor decides not to invest, the game ends and she gains $a < 1$ from the outside option and the trustee gets nothing. If the investor chooses to invest, the capital grows into 1, the return from the trustee's business. In the second stage, the trustee decides whether to return nothing or to return the amount $b$, where $a < b < 1$, to the investor. If the investor chooses to invest in the first stage and the trustee chooses to return in the second stage, the payoff is $b$ for the investor and $1-b$ for the trustee. If the investor chooses to invest but the trustee chooses not to return, then the Investor gets nothing and the Trustee gets 1. We assume $0 < a < b < 1$. The trust game and its payoff structure is described in Figure 10.

If the game is played once, the unique subgame perfect equilibrium is for the investor not to invest in the first stage and for the trustee not to return in the second stage. However,

[4] We will denote Investor as she and Trustee as he for clarity.

since the return from capital in the trustee's business is bigger than the outside option $a$ of the investor, the efficient outcome is for the investor to invest and for the trustee to return. Although the efficient outcome can not be achieved in the one-shot trust game, we will show below that it can be achieved in the "contagious equilibrium" when the trust game is infinitely repeated, even if the opponents are randomly rematched after each period.

We define *No Invest* as the defection of an investor, and *No Return* as the defection of a trustee. Define $d$-type investors or trustees as those whose history include defection of themselves or their partner, otherwise the players are $c$-type.

**Definition 14.** *The "contagious strategy" is defined as follows: An investor invests if she is c-type and does not invest if she is d-type. A trustee returns if he is c-type and does not return if he is d-type.*

The idea of the contagious strategy is that trust is applied to the community as a whole, not to each individual player, since the players are anonymous. Therefore, a single defection by a member means the end of the whole community trust, and a player who experiences dishonest behavior starts defecting all of his or her opponents (Kandori, 1992). Now, we provide an example in which cooperative game equilibria exist, showing that we can define payoffs for the trust game which allow trust and reciprocity in a sequential equilibrium for any fixed number of population.

**Theorem 15.** *Consider the random matching model described above where $2n \geq 4$ players play a trust game. Then for any $\delta$ and $n$, there exist $a$ and $b$ such that (i) $0 < a < b < 1$; and (ii) the contagious strategy constitutes a sequential equilibrium in which (Invest, Return) is the outcome in every period along the equilibrium path under uniformly random matching.*

To prove Theorem 15, we first introduce more notation. Let $X_t$ be the total number of $d$-type investors and $Y_t$ be the total number of $d$-type trustees at the beginning of period $t$. Let $Z_t$ denote the state of period $t$. In particular, $Z_t$ is a one-to-one and onto function from $(X_t, Y_t)$ to the set of natural numbers $\{1, 2, \cdots, n(n+2)\}$:

$$Z_t = (n+1)X_t + Y_t \quad \text{for} \quad X_t + Y_t > 0.$$

Let $A$ be an $n(n+2) \times n(n+2)$ transition matrix with elements $a_{ij} = \Pr\{Z_{t+1} = j \mid Z_t = i\}$. For example, $a_{12} = \Pr\{Z_{t+1} = 2 \mid Z_t = 1\} = \Pr\{(X_{t+1}, Y_{t+1}) = (0, 2) \mid (X_t, Y_t) = (0, 1)\}$.

Similarly, let $B$ be an $n(n+2) \times n(n+2)$ transition matrix with elements $b_{ij} = \Pr\{Z_{t+1} = j \mid Z_t = i$ and one $d$-type trustee deviates to Return$\}$. Define $\rho$ as an $n(n+2) \times 1$ column vector with the $i$th element equal to the conditional probability for the trustee to meet a $c$-type investor when the state is $i$ in period $t$.[5] Finally, let $e_i$ be a $1 \times n(n+2)$ row vector with the $i$th element equal to 1 and all other elements equal to 0.

Following Ellison (1994), we first prove the following lemma, which is useful in proving Theorem 15 and the results with public randomization in the next section.

**Lemma 16.** *Define* $f(\delta) = \delta e_1(B-A)(I-\delta A)^{-1}\rho$ *and* $g(\delta) = \delta e_{n+2}(B-A)(I-\delta A)^{-1}\rho$. *Then*

*(i)* $f(\delta)$ *is continuous and increasing over* $\delta \in (0,1)$, $\lim_{\delta\to 0} f(\delta) = 0$ *and* $\lim_{\delta\to 1} f(\delta) = 1$;

*(ii)* $g(\delta)$ *is continuous and increasing over* $\delta \in (0,1)$, $\lim_{\delta\to 0} g(\delta) = 0$ *and* $\lim_{\delta\to 1} g(\delta) < 1$;

*(iii)* $f(\delta) > g(\delta)$, $\forall \delta \in (0,1)$.

The lemma simply states that the benefit for a $d$-type trustee to deviate on off-the-equilibrium-path is larger when the state is $k = 1$ (i.e., there is no $d$-type investor and one $d$-type trustee) than the benefit when the state is $k = n+2$ (i.e., there is one $d$-type investor and one $d$-type trustee). The left hand side of the inequality can be written into $\delta \sum_{t=0}^{\infty} \delta^t e_1(B-A)A^t\rho$, which is the increase of payoffs for the $d$-type trustee when he chooses to deviate, given that there is no $d$-type investor and one $d$-type trustee in the current period. Similarly, the right hand side can be written into $\delta \sum_{t=0}^{\infty} \delta^t e_{n+2}(B-A)A^t\rho$, which is the increase of payoffs for the $d$-type trustee when he chooses to deviate, given that there is one $d$-type investor and one $d$-type trustee in the current period.

**Lemma 17.** *The contagious strategy constitutes a sequential equilibrium if*

$$a \geq \frac{n-1}{n} b, \tag{4.1}$$

*and*

$$g(\delta) \leq b \leq f(\delta). \tag{4.2}$$

[5] The formulas for $a_{ij}$, $b_{ij}$, and $\rho$ can be found in Appendix A.

*Proof.* As in Kandori (1992), we have only to check that one-shot deviations from the strategy are unprofitable after any history for both investors and trustees.

First, a one-shot deviation from the equilibrium path is unprofitable for a trustee, if

$$\frac{(1-b)}{1-\delta} \geq \sum_{t=0}^{\infty} \delta^t e_1 A^t \rho.$$

The left-hand side is the expected payoff from Return to the investor forever and the right-hand side is the expected payoff from No Return forever. The expression $e_1 A^t$ indicates the distribution over all the possible states after $t$ period and the term $e_1 A^t \rho$ is the probability of meeting a $c$-type investor at time $t$ given that the trustee was the first to defect at $t = 0$. Since the contagious equilibrium requires that the trustee should defect after he has defected, he receives payoff $c$ if he is matched with a $c$-type investor, and gets zero otherwise. This inequality can be simplified to the inequality:

$$1 - b \geq (1-\delta) e_1 (I - \delta A)^{-1} \rho \qquad (\star)$$

where $I$ denotes the identity matrix with size $n(n+2) \times n(n+2)$. Since

$$\begin{aligned}
& (1-\delta) e_1 (I - \delta A)^{-1} \rho + \delta e_1 (B - A)(I - \delta A)^{-1} \rho \\
= \ & e_1 (I - \delta A)^{-1} \rho - \delta e_1 A (I - \delta A)^{-1} \rho \\
= \ & \sum_{t=0}^{\infty} \delta^t e_1 A^t \rho - \delta \sum_{t=0}^{\infty} \delta^t (e_1 A) A^t \rho \\
= \ & e_1 I \rho = 1,
\end{aligned}$$

we can rewrite $(\star)$ as

$$b \leq \delta e_1 (B - A)(I - \delta A)^{-1} \rho.$$

Second, a one-shot deviation from the equilibrium path is unprofitable for the investor if

$$\frac{b}{1-\delta} \geq \frac{a}{1-\delta}.$$

The left-hand side is the expected payoff for the investor from Invest forever and the right-hand side is the expected payoff from No Invest forever. This condition is always satisfied given that $b > a$.

Next, we provide a sufficient condition for a one-shot deviation from an off-the-equilibrium-path to be unprofitable for the investor under any consistent belief, which supports the contagious equilibrium as a sequential equilibrium. A $d$-type investor finds a one-shot deviation from No Invest forever to be unprofitable for any number of $d$-type trustees if

$$\frac{a}{1-\delta} \geq \frac{n-l}{n}b + \delta\frac{a}{1-\delta}, \ \forall\, l = 1, 2, \ldots, n$$

The left-hand side is the expected payoff from No Invest forever and the right-hand side is the expected payoff from Invest in the current period and No Invest forever from the next period. With probability $(n-l)/n$ she meets a $c$-type trustee and gets $b$, or with probability $l/n$ she meets a $d$-type trustee and receive nothing in the current period. Since a $d$-type investor has larger incentives to deviate (i.e., Invest) when the number of $d$-type trustees is smaller in the society, the condition is binding when $n = 1$ and it can be simplified into inequality (4.1).

Finally, a one-shot deviation from an off-the-equilibrium-path is unprofitable for the trustee under any consistent belief. Since there is at least one $d$-type investor and one $d$-type trustee when the $d$-type trustee is currently on off-the-equilibrium-path, any consistent belief for the state $k < n+2$ must be zero. A $d$-type trustee finds a one-shot deviation from No Return forever to be unprofitable given $Z_t = k$, for all $k = n+2, \ldots, n(n+2)$, if

$$1 + \sum_{t=1}^{\infty} \delta^t e_k A^t \rho \geq (1-b) + \delta \sum_{t=0}^{\infty} \delta^t e_k B A^t \rho.$$

The left-hand side is the expected payoff from No Return forever and the right-hand side is the expected payoff from Return in the current period and No Return forever from the next period. The condition is based on the assumption that the trustee meets a $c$-type investor in the current period, otherwise the stage game is over and the trustee needn't make any decision. Recall that the term $e_k A^t \rho$ is the probability of meeting a $c$-type investor after $t$ periods given that the current state is $Z_t = k$. The inequality can be simplified into

$$b \geq \delta e_k (B - A)(I - \delta A)^{-1} \rho \ \text{ for } k = n+2, \ldots, n(n+2).$$

This condition is binding when $k = n + 2$. Therefore it suffices to show the following inequality is satisfied

$$b \geq \delta e_{n+2}(B - A)(I - \delta A)^{-1}\rho.$$

By Lemma 16, we have inequality (4.2). This completes the proof of Lemma 2.

The proof above implies that $0 < \delta e_1(B - A)(I - \delta A)^{-1}\rho < 1$, therefore we can always choose proper values for $a$ and $b$ where $0 < a < b < 1$ and conditions (4.1) and (4.2) are satisfied. □

Condition (4.1) means that the loss of deviation (to Invest) in an off-the-equilibrium-path for an investor is greater than the benefit of deviation, even if there is only one $d$-type trustee. Therefore, the investor has no incentive to deviate from the contagious strategy given any history.

Condition (4.2) controls the trustee's incentive to deviate from the contagious strategy in both on-the-equilibrium-path and off-the-equilibrium-path. When the trustee is on the equilibrium path, the gain of deviation must be less than the gain from deterring starting a defection. When the trustee is off the equilibrium path, the loss of slowing down the contagious procedure must be greater than the gain of slowing down the procedure when there is already defection in the community.

Since the trustee's gain of deviation on the equilibrium path is same as the loss of deviation off the equilibrium path, which is $b$, and the effect of slowing down the contagious procedure is larger when there are less $d$-type players, we can always find $b$ which satisfies condition (4.2).[6]

There are two points we want to make about this result. First, please note that the existence of the contagious equilibrium critically depends on the existence of the outside option. The reasoning is as follows. The concept of contagious equilibrium is based on community enforcement. Players change their partners over time and dishonest behavior against one partner causes sanctions by other members in the society. For the development of a cooperative social norm, this concept requires harsh punishment scheme. Not only are

[6] The other way to interpret Theorem 15 and Lemma 17 is as follows. For any $n$, and for any $a$ and $b$ which satisfy condition (4.1), we can define the relevant interval of $\delta$ which satisfies condition (4.2). This explanation becomes clear with Figure 11.

deviators from the desired behavior punished, but a player who fails to punish is in turn punished (Kandori (1992)). In the trust game the trustee has strong incentives to defect. Then the corresponding cheated investor (now she is $d$-type) must defect forever even if she meets a $c$-type trustee. To sustain this $d$-type investor's defection in the off-the-equilibrium-path, the outside option $a$ must be high enough. Second, we can observe in (4.2) that the discount factor, $\delta$, has a strong effect on the equilibrium payoff structure. The larger the discount factor is, the higher level of Return ($b$) is required to support the equilibrium. The logic is that in order to make a $d$-type trustee with a higher value in $\delta$ defect forever in an off-the-equilibrium-path, we should make a deviation from defection less attractive by imposing a higher value on $b$ (and a lower value on $1 - b$).

## 4.3 THE MODEL WITH PUBLIC RANDOMIZATION

In the previous section, we have shown that a community can sustain an efficient outcome under anonymous random matching by employing the contagious strategy. However, the equilibrium is fragile in the sense that a small amount of noise (defection) causes the breakdown of trust and reciprocity in the whole community. On the contrary, if the equilibrium always goes back to the original payoff point, it is robust to the mistakes of players. Here we achieve "global stability" (Kandori (1992) by introducing public randomization device. Also public randomizations play a significant role in expanding the range of sufficient conditions that support the social norm of trust and reciprocity by allowing the adjustment of players' continuation payoffs.

We assume that, at the beginning of each period $t$, players observe a public random variable $\gamma_t$ which is drawn independently from a uniform distribution on $[0, 1]$. The public randomizations are used to adjust the severity of the punishment.

Given $\gamma_t$ we redefine the types of the players as follows:

- $t = 1$, all players are $c$-type;
- $t > 1$, player $i$ is $c$-type if $(i)$ player $i$ is $c$-type in $t-1$ and (Invest, Return) is the outcome in $t-1$; or $(ii)$ $\gamma_{t-1} > \gamma$. Otherwise player $i$ is $d$-type.

**Definition 18.** *The "contagious strategy with public randomization" is defined as follows: An investor invests if she is c-type and does not invest if she is d-type. A trustee returns if he is c-type and does not return if he is d-type.*

With this preparation, we can construct a "stable" equilibrium which sustain trust and reciprocity under public randomization.

**Theorem 19.** *Consider the random matching model with public randomization where $2n \geq 4$ players play the trust game with $0 < a < b < 1$ and $a \geq \frac{n-1}{n}b$. Then, $\exists \underline{\delta} < 1$ such that $\forall \delta \in [\underline{\delta}, 1)$, the contagious strategy with public randomization constitutes a sequential equilibrium in which (Invest, Return) is the outcome in every period along the equilibrium path under uniformly random matching.*

The formal equilibrium conditions are given as follows.

**Lemma 20.** *The contagious strategy with public randomization constitutes a sequential equilibrium if*

$$a \geq \frac{n-1}{n}b, \tag{4.3}$$

*and*

$$g(\delta\gamma) \leq b \leq f(\delta\gamma). \tag{4.4}$$

The proof of Lemma 20 is similar to the proof of Lemma 17 and can be found in Appendix. Here we provide the intuitive explanation for Theorem 19 and the proof in Figure 11. Suppose the relevant value of $b$ satisfying (4.3) is given. Since function $f(\cdot)$ defined in Lemma 1 is continuous and increasing, we can always find $\underline{\delta} \in (0, 1)$ where $f(\underline{\delta}) = b$. By Lemma 16, we have $g(\underline{\delta}) < f(\underline{\delta}) = b$. We define $\gamma \equiv \underline{\delta}/\delta$. Then for any $\delta \in [\underline{\delta}, 1)$, $\delta\gamma = \underline{\delta}$ and the condition (4.4) in Lemma 20 is satisfied.

Remark The public randomization device dramatically expands the set of $\delta$ which supports the contagious strategy as a sequential equilibrium. Suppose there is no public randomization, (i.e., $\gamma = 1$) and players play the same trust game. Then only $\delta \in [\underline{\delta}, \overline{\delta}]$ supports the equilibrium[7]. Note that $\delta \geq \overline{\delta}$ cannot support the equilibrium. The intuition

[7]This also contrasts with the result of Kandori (1992), where the "contagious equilibrium" requires $\delta$ to be sufficiently large in the prisoner's dilemma.

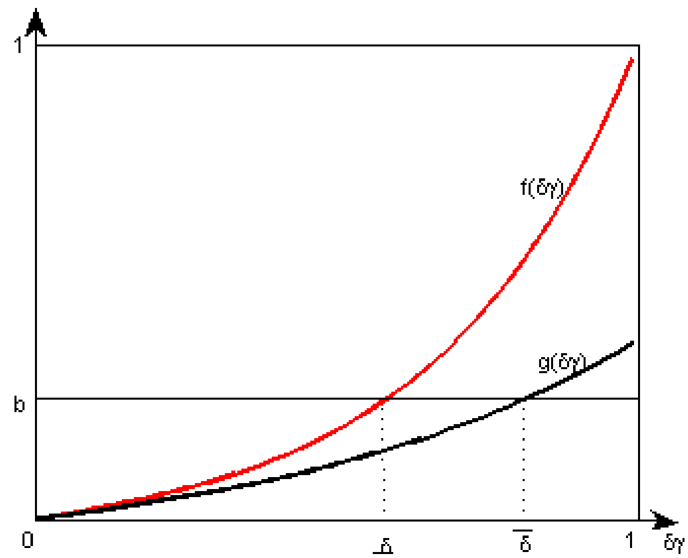


Figure 11: Graphical explanation of Theorem 2 (with $n = 5$ and uniformly random matching: Matlab)

is that if $\delta$ is very high, then a $d$-type trustee has a high incentive to cooperate in an off-the-equilibrium path since $g(\delta) > b$ in this case. This effect is even stronger as the value of $b$ declines. When $b$ is small enough, very short interval of $\delta$ supports the equilibrium. With a small $b$, the interval $[\underline{\delta}, \overline{\delta}]$ supporting the equilibrium gets shorter while the interval $[\underline{\delta}, 1]$ sustaining the equilibrium with randomizatons becomes bigger. Therefore, when the $b$ is small the set of $\delta$ which supports the equilibrium expands dramatically with the introduction of public randomizations.

## 4.4 CONCLUSION

In this paper, we have shown that the social norm of trust and reciprocity is sustained in equilibrium by the use of "contagious" punishments which lead eventually to the breakdown of cooperation after a single deviation. Technically, our model is an extension of the theory of repeated random matching games in which players not only do not observe the outcomes of games in which they are not involved, but also are completely anonymous in that they can neither recognize nor communicate the identity of any of their past opponents. The results illustrate that the contagious punishments are fairy powerful tool for enforcing a social norm of trust and reciprocity.

We observe many important transactions in reality are infrequent in nature and many transactions happen among essentially anonymous players. Electronic transactions are done between strangers who have no contact except through cyberspace. In this case only partial information about a stranger's reputation is available at best, and, therefore, the effectiveness of reputation is far less certain. In all of the results above, we have tried to answer the question of what factors drive the emergence of trust and reciprocity in economic transactions.

The contributions of this paper are as follows. First, we theoretically extend Kandori (1992) and Ellison (1994)'s argument to the trust game. These papers show that community can sustain cooperation without any information processing in the setting of the prisoner's dilemma. And Kandori (1992) only shows that, *under local information processing*, a simple form of community enforcement supports cooperation in the one-sided incentive problem. In this paper, we have considered the one-sided incentive problem under the most extreme information restriction. That is, players know nothing more than his or her personal experience since they are completely anonymous in that they can neither recognize nor communicate the identity of any of their past opponent. This is interesting since, as Ellison (1994) notes, the results of the previous papers heavily rely on the fact that the prisoner's dilemma has a dominant strategy equilibrium. And it is important to know whether the results extend to the other class of games.

Second, our results rationalizes many experimental results which show the significant level of trust and reciprocity in the laboratory. Many experimental researches have been done to identify trust and reciprocity in an investment setting, and found that trust and reciprocity prevail in the laboratory even under the conditions of complete anonymity and one-shot interaction. They regard these behaviors irrational and explain them with some psychological factors such as fairness, altruism and inequality aversion etc. The contribution of this study is that we can identify trusting and reciprocating behavior as an equilibrium phenomenon without changing the fundamental assumption in economics that individuals act in their own self interest. That is, we show that trust and reciprocity are consistent with a rational choice of selfish players, and investigate how such a cooperative social norm is sustained by self-interested group members in the community.

Finally, we expect that this study will help explain the emergence and prevalence of e-Commerce which is important as the internet develops. The results of this study provide the first analysis of the development of trust and reciprocity among strangers based on the concept of social norm. More generally, an understanding of the factors that determine the development of the social norm among strangers is of obvious importance to the design and operation of economic transactions and economic institutions.

We should note that we have left one important question to future research. The contagious strategy is the most extreme punishment scheme, and the contagious equilibrium is just one of a multiplicity of equilibria. Therefore, it would be interesting to characterize all the possible equilibria, or, at least, to know whether other relaxed strategies could support trust and reciprocity. It is also related to the question of whether Folk Theorem is applicable in this setting.

## BIBLIOGRAPHY


[1] Bala, V. and Goyal, S. (2000). “A Noncooperative Model of Network Formation,” *Econometrica*, 68, 1181-1230.

[2] Bloch. F. and Jackson. M. (2005). "Definitions of Equilibrium in Network Formation Games," Working paper. Université d'Aix-Marseille and California Institute of Technology.

[3] Bramoullé. Y. and Kranton. R. (2005). "Risk-Sharing Networks," Working paper. Université Laval and University of Maryland.

[4] Clarke, R. N. (1983). "Collusion and the Incentives for Information Sharing," *Bell Journal of Economics*, 14, 383-394.

[5] Crawford, V. and Sobel, J. (1982). “Strategic Information Transmission,” *Econometrica*, 50, 1431-1451.

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

[7] Dutta, B. and Mutuswami, S. (1997). “Stable Networks,” *Journal of Economic Theory*, 76, 322-344.

[8] Gal-Or, E. (1985). "Information Sharing in Oligopoly," *Econometrica*, 53, 329-343.

[9] Gal-Or, E. (1986). "Information Transmission – Cournot and Bertrand Equilibria," *Review of Economic Studies*, 53, 85-92.

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

[11] Goyal, S., Konovalov, A., and Moraga-Gonzalez, J. L. (2004). "Hybrid R&D," Working paper. University of Essex and Erasmus University Rotterdam.

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

[13] Jackson, M. (2004). “A Survey of Models of Network Formation: Stability and Efficiency,” forthcoming in *Group Formation in Economics: Networks, Clubs, and Coalitions*, edited by G. Demange and M. Wooders, Cambridge University Press: Cambridge U.K., http://www.hss.caltech.edu/~jacksonm/netsurv.pdf.

[14] Jackson, M. and van den Nouweland, A. (2005). "Strongly Stable Networks," *Games and Economic Behavior,* 51, 420-444.

[15] Jackson, M. and Watts, A. (2002). "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, 106, 265-295.

[16] Jackson, M. and Wolinsky, A. (1996). “A Strategic Model of Social and Economic Networks,” *Journal of Economic Theory,* 71, 44-74.

[17] Kirby, A. (1988). "Trade Associations as Information Exchange Mechanism," *Rand Journal of Economics*, 19, 138-146.

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

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

[20] Myerson, R. (1977). "Graphs and Cooperation in Games," *Mathematical Operations Research*, 2, 225-229.

[21] Novshek, W. and Sonnenschein, H. (1982). "Fulfilled Expectations Cournot Duopoly with Information Acquisition and Release," *Bell Journal of Economics*, 13, 214-218.

[22] Radner, R. (1962). "Team Decision Problems," *Annals of Mathematical Statistics*, 33, 857-881.

[23] Roth, A. and Sotomayor, M. (1989). *Two Sided Matchig*, Econometric Society Monographs No.18: Cambridge University Press.

[24] Suzumura, K. (1992). "Cooperative and Noncooperative R&D in an Oligopoly with Spillovers," *American Economic Review*, 82, 1307-1320.

[25] Vives, X. (1984). "Duopoly Information Equilibrium: Cournot and Bertrand," *Journal of Economic Theory,* 34, 71-94.

[26] Watts, A. (2001). "A Dynamic Model of Network Formation," *Games and Economic Behavior*, 34, 331-341.

[27] Bilodeau, M. and Slivinski, A., 1996. Toilet cleaning and department chairing: volunteering a public service. J. Public Econ. 59, 299-308.

[28] Bliss, C., Nalebuff, B., 1984. Dragon-slaying and ballroom dancing: the private supply of a public good. J. Public Econ. 25, 1-12.

[29] Bloch. F. and Jackson. M. (2005). "Definitions of Equilibrium in Network Formation Games," Working paper. Université d'Aix-Marseille and California Institute of Technology.

[30] Bramoulle, Y., Kranton, R., 2003. A network model of public goods: experimentation and social learning. Working paper. Université Laval and University of Maryland.

[31] Ellison, G., 1993. Learning, local interaction and coordination. Econometrica. 61, 1047-1071.

[32] Eshel, I., Samuelson, L., Shaked, A., 1998. Altruists, egoists, and hooligans in a local interaction model. Amer. Econ. Rev. 88, 157-179.

[33] Goyal, S., Konovalov, A., and Moraga-Gonzalez, J. L. (2004). "Hybrid R&D," Working paper. University of Essex and Erasmus University Rotterdam.

[34] Goyal, S., Vega-Redondo, F., 2003. Network formation and social coordination. Working paper. Erasmus University.

[35] Hendricks, K., Weiss, A., Wilson, C., 1988. The war of attrition in continuous time with complete information. Int. Econ. Rev. 29, 663-680.

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

[37] Morris, S., 2000. Contagion. Rev. Econ. Stud. 67, 57-78.

[38] Roth, A. and Sotomayor, M. (1989). *Two Sided Matchig*, Econometric Society Monographs No.18: Cambridge University Press.

[39] Watts, A. (2001). "A Dynamic Model of Network Formation," *Games and Economic Behabior*, 34, 331-341.

[40] Berg, J., Dickhaut, J. and McCabe, K. (1993). "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10, 122-142.

[41] Bolton, G. and Ockenfels, A. (2000). "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review,* 90, 166-193.

[42] Bolton, G., Katok, E. and Ockenfels, A. (2005). "Cooperation among Strangers with Limited Information about Reputation," *Journal of Public Economics,* 89, 1457-1468.

[43] Duffy, J. and Ochs, J. (2005). "Cooperative Behavior and the Frequency of Social Interaction," Working Paper, University of Pittsburgh.

[44] Ellison, G. (1994). "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *Review of Economic Studies,* 61, 567-588.

[45] Engle-Warnick, J. and Slonim, R. (2006). "Learning to Trust in Infinitely Repeated Games," *Games and Economic Behavior,* 54, 95-114.

[46] Ghosh, P. and Ray, D. (1996). "Cooperation in Community Interaction without Information Flows," *Review of Economic Studies,* 63, 491-519.

[47] Greif, A. (1993). "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review,* 83, 525-548.

[48] Kandori, M. (1992). "Social Norms and Community Enforcement," *Review of Economic Studies,* 59, 63-80.

[49] Kranton, R. (1996). "The Formation of Cooperative Relationships," *Journal of Law, Economics, and Organization*, 12, 214-233.

[50] Kreps, D., Milgrom, P., Roberts, J. and Wilson, R. (1982). "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27, 245-252.

[51] Kreps, D. and Wilson, R. (1982). "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

[52] Milgrom, P.and Roberts, J. (1982). "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

[53] Okuno-Fujiwara, M. and Postlewaite, A. (1995). "Social Norms and Random Matching Games," *Games and Economic Behavior*, 9, 79-109.

[54] Tirole, J. (1996). "A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality)," *Review of Economic Studies,* 63, 1-22.

# APPENDIX

## APPENDIX FOR CHAPTER 4

### A: Formulas for Matrix A, Matrix B and Vector $\rho$

Matrix $A$ is an $n(n+2) \times n(n+2)$ transition matrix with elements $a_{ij} = \Pr\{Z_{t+1} = j \mid Z_t = i\}$. Suppose $Z_t^{-1}(i) = (X_t, Y_t) = (p, q)$, and $Z_{t+1}^{-1}(j) = (X_{t+1}, Y_{t+1}) = (r, s)$. Then

$$a_{ij} = \begin{cases} \frac{\binom{q}{r-p}\binom{n-p}{r-p}(r-p)!\binom{p}{p+q-r}(p+q-r)!(n-q)!}{n!}, \\ \text{if } s = r,\ p \leq r,\ q \leq r, \text{ and } r \leq p+q; \\ \\ 0, \text{ otherwise.} \end{cases}$$

Matrix $B$ is an $n(n+2) \times n(n+2)$ transition matrix with elements $b_{ij} = \Pr\{Z_{t+1} = j \mid Z_t = i$ and one $d$-type trustee deviates to Return$\}$. Suppose $Z_t^{-1}(i) = (X_t, Y_t) = (p, q)$, and

$Z_{t+1}^{-1}(j) = (X_{t+1}, Y_{t+1}) = (r, s)$. Then

$$b_{ij} = \begin{cases} a_{ij}, \text{ if } q = 0; \\ \\ \frac{\binom{n-p}{r-p}\binom{q-1}{r-p}(r-p)!\binom{p}{p+q-r}(p+q-r)!(n-q)!}{n!}, \\ \text{if } s = r,\ 1 \le p \le r,\ 1 \le q \le r, \text{ and } r \le p+q-1; \\ \\ \frac{\binom{n-p}{r-p}\binom{q-1}{r-p}(r-p)!\binom{n-r}{1}\binom{p}{p+q-r-1}(p+q-r-1)!(n-q)!}{n!}, \\ \text{if } s = r+1,\ 0 \le p \le r,\ 1 \le q \le r+1, \text{ and } r \le p+q-1; \\ \\ 0, \text{ otherwise.} \end{cases}$$

$\rho$ is an $n(n+2) \times 1$ column vector with the $i$th element equal to the conditional probability for the trustee to meet a $c$-type investor when the state is $i$ in period $t$.

$$\rho = (\underbrace{\frac{n}{n}, \cdots, \frac{n}{n}}_{n \text{ elements}}, \underbrace{\frac{n-1}{n}, \cdots, \frac{n-1}{n}}_{n+1 \text{ elements}}, \cdots, \underbrace{\frac{1}{n}, \cdots, \frac{1}{n}}_{n+1 \text{ elements}}, \underbrace{\frac{0}{n}, \cdots, \frac{0}{n}}_{n+1 \text{ elements}})^T$$

# B: Proofs

We prove Lemma 16 by introducing the following notations similar as in Ellison (1994).

*Proof.* Define $\omega$ as the random variable whose realization is a pairing of all players in each period. Define $o_i(t,\omega)$ as player $i$'s opponent in period $t$ for a given realization of $\omega$.

Define the sets $C_I(t,p,q,\omega)$ and $C_T(t,p,q,\omega)$ by

$$\begin{aligned} C_I(0,p,q,\omega) &= \{p+1, p+2, \cdots, n\}, \\ C_T(0,p,q,\omega) &= \{n+q+1, n+q+2, \cdots, 2n\}, \\ C_I(t+1,p,q,\omega) &= \{i \in C_I(t,p,q,\omega) \mid o_i(t,\omega) \in C_T(t,p,q,\omega)\}, \\ C_T(t+1,p,q,\omega) &= \{i \in C_T(t,p,q,\omega) \mid o_i(t,\omega) \in C_I(t,p,q,\omega)\}. \end{aligned}$$

$C_I(t,p,q,\omega)$ and $C_T(t,p,q,\omega)$ will be the set of $c$-type investors and $c$-type trustees in period $t$ when every player plays contagious strategy and the sets of players $\{1,2,\cdots,p\}$ and $\{n+1, n+2, \cdots, n+q\}$ are the $d$-type investors and $d$-type trustees in period 0.

Define the set $D(t,\omega)$ by

$$\begin{aligned} D(0,\omega) &= \{2n\}, \\ D(t+1,\omega) &= D(t,\omega) \cup \{i \mid o_i(t,\omega) \in D(t,\omega)\}. \end{aligned}$$

$D(t,\omega)$ gives the set of $d$-type players in period $t$ suppose that the trustee $2n$ is the only $d$-type player in period 0.

From equation (1), $Z^{-1}(1) = (0,1)$ and $Z^{-1}(n+2) = (1,1)$, then

$$\begin{aligned} f(\delta) &= \delta e_1 (B-A)(I-\delta A)^{-1}\rho \\ &= \delta \sum_{t=0}^{\infty} \delta^t e_1 (B-A) A^t \rho \\ &= E_\omega[\sum_{t=1}^{\infty} \delta^t \Pr(o_{2n}(t,\omega) \in C_I(t,0,0,\omega) \cap D(t,\omega))], \end{aligned}$$

and

$$\begin{aligned} g(\delta) &= \delta e_{n+2}(B-A)(I-\delta A)^{-1}\rho \\ &= \delta \sum_{t=0}^{\infty} \delta^t e_{n+2}(B-A)A^t \rho \\ &= E_\omega[\sum_{t=1}^{\infty} \delta^t \Pr(o_{2n}(t,\omega) \in C_I(t,1,0,\omega) \cap D(t,\omega))]. \end{aligned}$$

We show that

$$\begin{aligned} E_\omega[\sum_{t=1}^{\infty} \delta^t \Pr(o_{2n}(t,\omega) &\in C_I(t,0,0,\omega) \cap D(t,\omega))] \\ &\geq E_\omega[\sum_{t=1}^{\infty} \delta^t \Pr(o_{2n}(t,\omega) \in C_I(t,1,0,\omega) \cap D(t,\omega))] \end{aligned}$$

by showing that the inequality holds for every realization of $\omega$, i.e.,

$$\begin{aligned} \sum_{t=1}^{\infty} \delta^t I(o_{2n}(t,\omega) &\in C_I(t,0,0,\omega) \cap D(t,\omega)) \\ &\geq \sum_{t=1}^{\infty} \delta^t I(o_{2n}(t,\omega) \in C_I(t,1,0,\omega) \cap D(t,\omega)) \end{aligned} \quad (.1)$$

(The notation $I(E)$ indicates a function equal to one or zero depending on whether the deterministic condition $E$ is true or false.) The definition of $C_I$ implies that

$$C_I(t,1,0,\omega) \subset C_I(t,0,0,\omega),$$

so

$$C_I(t,1,0,\omega) \cap D(t,\omega) \subset C_I(t,0,0,\omega) \cap D(t,\omega).$$

and the inequality (.1) gives the desired result.

Proof of (16):

$$\begin{aligned} f(\delta) &= \delta e_1(B-A)(I-\delta A)^{-1}\rho \\ &= \delta \sum_{t=0}^{\infty} \delta^t e_1(B-A)A^t \rho. \end{aligned}$$

$e_1(B-A)A^t\rho$ is the difference in the probability for a trustee to meet a $c$-type invest after $t$ periods when he chooses to start defection in next period and when he chooses to start defection in the current period. Therefore $e_1(B-A)A^t\rho > 0$ and it goes to 0 with probability 1 when $t$ goes to infinite. $\square$

*Proof.* As in the proof of Lemma 17 and Theorem 15, we have only to prove that one-shot deviations from the strategy are unprofitable after any history for both investors and trustees.

For ease of exposition, let $\Psi_I$ denote the sum of expected payoff for investors when $\gamma_s > \gamma$ for all $s \geq t$, and let $\Psi_T$ denote the sum of expected payoff for trustees when $\gamma_s > \gamma$ for all $s \geq t$, given that all the players follow the contagious strategy with public randomization.

$$\Psi_I = \sum_{t=1}^{\infty} \delta^t \gamma^{t-1}(1-\gamma)\frac{b}{1-\delta},$$

$$\Psi_T = \sum_{t=1}^{\infty} \delta^t \gamma^{t-1}(1-\gamma)\frac{1-b}{1-\delta}.$$

First, a one-shot deviation from the equilibrium path is unprofitable for a trustee, i.e.,

$$\frac{1-b}{1-\delta} \geq \sum_{t=0}^{\infty} \delta^t \gamma^t e_1 A^t \rho + \Psi_T.$$

This inequality can be simplified to yield inequality:

$$1 - b \geq (1-\delta\gamma)e_1(I-\delta\gamma A)^{-1}\rho.$$

By

$$(1-\delta\gamma)e_1(I-\delta\gamma A)^{-1}\rho + \delta\gamma e_1(B-A)(I-\delta\gamma A)^{-1}\rho = 1,$$

we have

$$b \leq \delta\gamma e_1(B-A)(I-\delta\gamma A)^{-1}\rho. \tag{.2}$$

Second, a one-shot deviation from the equilibrium path is unprofitable for the investor if

$$\frac{b}{1-\delta} \geq \frac{a}{1-\delta\gamma} + \Psi_I.$$

This condition is always satisfied given that $b > a$.

Next, we provide a sufficient condition for a one-shot deviation from an off-the-equilibrium-path to be unprofitable for the investor under any consistent belief, which supports the contagious equilibrium as a sequential equilibrium. A $d$-type investor finds a one-shot deviation from No Invest forever to be unprofitable for any number of $d$-type trustees if

$$\frac{a}{1-\delta\gamma} + \Psi_I \geq \frac{n-l}{n}b + \delta\gamma\frac{a}{1-\delta\gamma} + \Psi_I, \ \forall\, l = 1, 2, \ldots, n$$

The condition is binding when $n = 1$ and it can be simplified into inequality

$$a \geq \frac{n-1}{n}b. \tag{.3}$$

Finally, a one-shot deviation from an off-the-equilibrium-path is unprofitable for the trustee under any consistent belief. A $d$-type trustee finds a one-shot deviation from No Return forever to be unprofitable given $Z_t = k$, for all $k = n+2, \ldots, n(n+2)$, if

$$1 + \sum_{t=1}^{\infty} \delta^t \gamma^t e_k A^t \rho + \Psi_T \geq (1-b) + \delta\gamma \sum_{t=0}^{\infty} \delta^t \gamma^t e_k B A^t \rho + \Psi_T.$$

The inequality can be simplified into

$$b \geq \delta\gamma e_k (B-A)(I-\delta\gamma A)^{-1}\rho \ \text{ for } k = n+2, \ldots, n(n+2).$$

This condition is binding when $k = n+2$. Therefore it suffices to show the following inequality is satisfied

$$b \geq \delta\gamma e_{n+2}(B-A)(I-\delta\gamma A)^{-1}\rho. \tag{.4}$$

□

*Proof.* For any $0 < b < 1$, $\exists\ \underline{\delta} \in (0,1)$ such that

$$b = \underline{\delta} e_1 (B-A)(I-\underline{\delta}A)^{-1}\rho$$

by property (i) of Lemma 16. And by property (iii) of Lemma 16,

$$b > \underline{\delta} e_{n+2}(B-A)(I-\underline{\delta}A)^{-1}\rho.$$

Then for any $\delta \in [\underline{\delta}, 1)$, define $\gamma(\delta) = \underline{\delta}/\delta$. So inequality (4.4) is satisfied automatically. □