2013

Strategic interaction and social networks

Quqiong He
Louisiana State University and Agricultural and Mechanical College

STRATEGIC INTERACTION AND SOCIAL NETWORKS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by

Quqiong He
B.S., Wuhan University, 2008
M.S., Louisiana State University, 2010
August 2013

This dissertation is dedicated to my father, Xinrong He and my mother, Heding Ding.

ACKNOWLEDGEMENTS

I would like to thank Professor Sudipta Sarangi for his encouragement and guidance throughout my studies at Louisiana State University. I also would like to thank the other members of my committee, Professors Carter Hill and Robert Newman, for their advice. Moreover, Dr. Ying Pan provided a great amount of assistance in the early stages of my work. I am grateful to all of the wonderful people who are being with me along this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation consists of three essays which investigate individuals' interaction in different contexts using social network analysis. The first essay generalizes the models of link formation of Rogers (2005) by allowing that giving and asking choices can be made separately and simultaneously by each agent. We focus on two specifications of the relationship function: the concave specification and the linear specification. The second essay empirically tests how the pattern of village structure, in terms of lineage network composition, affects people's reciprocal behavior, utilizing data from Chinese Household Income Project Survey 2002. The third essay demonstrates different types of asymmetries and investigates individuals' behavior in a model of friendship networks based on Brueckner (2006).

CHAPTER 1: INTRODUCTION

A social network describes a group of individuals, and the relations between them. Relations, often depicted as links, are means for communication and for allocation of goods and services, such as invitations, information, friendship, opportunities and the like (Jackson and Wolinsky, 1996). The pattern of individuals' interaction, which is embedded in a social network, plays a key role in shaping economic outcomes and thus has broad implications. This promotes both theoretical and empirical analysis of social networks. Existing literature provides extensive research across a wide range of subjects such as hyperlinks between webpages, political alliance, job hunting in labor markets, research collaboration among firms, and provision of public goods (Newman and Girvan, 2004; Granovetter, 1973; Rees, 1966; Baker, Murphy and Gibbons, 2004; Bramoulle and Kranton, 2007).

In a network, establishing and maintaining the relations take time and effort. How to allocate resources across different relations is then a fundamental question, such as how to spend limited effort in obtaining information from others, how to spend limited time in helping each other or how to spend a limited budget in building public goods. Once the relation is created, individuals can exchange information or favor through the relations. A social network analysis aims at investigating in different contexts what kind of relation structure will emerge and what the economic effects of the pattern of relations will be.

The presence of social networks or communities may enforce a set of norms or behaviors, such as altruism, cooperation and trust. Putnam (1999) defined this set of norms or behaviors as social capital — "features of social life, networks, norm, trust that enable participants to act together more effectively to pursue shared objectives." Recent literature relates social capital with community heterogeneity. For example, Alesina and La Ferrara (2002) illustrate a negative relationship between racial fragmentation and trust, a major component of social capital. Leigh (2006) finds the same relationship between ethnic fragmentation and trust. Some studies also argue that ethnic fragmentation is inversely related with public-good provision (Banerjee et al., 2005; Alesina et al., 1999). Thus, social network pattern may affect individuals' interaction and building of social capital.

These important questions stated above motivate my research. This dissertation, presented in the following chapters, investigates individuals' interaction in different contexts using social network analysis. The second chapter of my dissertation studies network formation in a model of asking and giving, where the amount of benefits individuals obtain from their connections depends on those agents' effort in asking as well as their connections' effort in giving. To further explore the relationship between social networks and individuals' behavior, the third chapter empirically tests how the pattern of village structure in China, in terms of lineage network composition, affects people's reciprocal behavior. Examples of reciprocal behavior include helping each other, borrowing and lending, or public-good provision. The last chapter analyzes a special type of networks — friendship network. We introduce different types of asymmetries and investigate individuals' behavior in a model of friendship networks based on Brueckner (2006).

1.1 Network formation in a model of asking and giving

People derive benefits from connecting with each other. These benefits may be pleasure, information, favors and so on. The second chapter of this dissertation studies network formation in a model of asking and giving. This model builds on the model of asking and the model of giving introduced by Rogers (2005). Rogers examines the behavior of asking and giving in separate models and claims that inefficiency comes from the behavior of giving. We generalize

Rogers' models by incorporating two decisions – asking and giving – into one model. More importantly, by considering asking behavior and giving behavior at the same time, we provide a new perspective into some aspects that cannot be obtained in separate models, i.e., the relationship between asking behavior and giving behavior.

We assume in the network each individual is endowed with an intrinsic value and this value is publicly observed. Each individual has a budget constraint which implies a limited amount of resources spent on obtaining information from others. Apart from one's own intrinsic value, an individual also wants to get some information from others. So people have to decide whom to connect with and how much effort to spend in establishing the relationships. Once the relationship is established between two individuals, information naturally flows from one to the other. The amount of information the other agent gets depends on the nature of the relationship, which is represented by a relationship quality function. The relationship quality function has two arguments: the effort of asking information and the effort of giving information. For example, the share of information flowing from agent j to agent i depends on agent i's effort in asking as well as agent j's effort in giving. This chapter aims at studying under different assumptions how agents behave when they face the decisions of both asking and giving and how the socially optimal network structure would respond.

This model contributes to the literature by assuming the amount of information one confers to other agents is endogenously determined by the network structure. This reasonable assumption leads to several important characteristics. First, all paths between two agents generate benefits. Second, we take "feedback effects" into account, whereby the benefits associated with a relationship are counted many times. Feedback effects appear frequently in daily life and have important implications.

1.2 Lineage-based fragmentation and cooperative behavior in rural China

In less developed areas, formal institutions are often missing or weak. Hence, in situations when information asymmetries are crucial, informal institutions instead play an important role. In rural China, many villages are still structured by a number of traditional lineage organizations, which results in lineage-based fragmentation.

Using data from the Chinese Household Income Project Survey (CHIPS) 2002, we define three types of villages: types 1-3, which go from the most homogenous to the most heterogeneous villages. We then measure intra-lineage cooperation by the frequency of mutual help within a lineage. Two kinds of mutual help are considered: monetary help and non-monetary help that is time-consuming. Inter-lineage cooperation is measured by villagers' physical contribution to public goods and the share of village budget spent on public goods. This chapter aims at examining how lineage-based fragmentation affects cooperative behavior in rural China.

This study is novel for three reasons. First, to the best of our knowledge, this is the first study that presents a full picture of cooperative behavior by examining both intra-group and inter-group cooperation. Second, since both the provision of public goods in rural China and fragmentation are measured at the village level, our study presents a more convincing relationship between the two variables than some existing literature does. Third, China serves as an excellent case for studying fractionalization because the lineage composition within a village is exogenously determined.

1.3 Asymmetries in friendship networks

The fourth chapter of this dissertation considers a model of friendship networks based on Brueckner (2006) where costly links with an uncertain success probability yield direct and indirect benefits. We first study cost asymmetries by allowing for an agent with lowest linking costs called the *cost-magnetic* agent. Next we focus on network asymmetries by allowing for a *knows-everyone* agent. We characterize the equilibrium effort levels for both cases for the class of regular networks. We also show that this cannot be done for arbitrary networks.

This work extends the work of Brueckner (2006) by introducing different types of asymmetries in the model of friendship networks. Given the model setup, asymmetries can occur either in values or costs, or in the network structure itself. Brueckner himself proposes value based asymmetry and network asymmetry but considers only specific examples for both types. In the real world, there is yet another type of asymmetry. With the same level of effort, the cost of forming friendships is less for certain individuals but high for many others. This idea motivates the analysis of cost based asymmetry. Since Roy and Sarangi (2009) have already examined value based asymmetry, in this chapter we first introduce cost based asymmetry and then focus on network asymmetry. Unlike Brueckner (2006) where asymmetries are examined only for very small sets of agents, for both instances we consider the general case with n agents.

The examination of asymmetries is important for two reasons. First, most economic environments are not characterized by homogeneous agents. Second, they act as a robustness check for results obtained in the homogenous model. Thus, the present extensions could benefit both decisions makers and researchers in important areas when they face different occasions.

CHAPTER 2: NETWORK FORMATION IN A MODEL OF ASKING AND GIVING

2.1 Introduction

The process of strategic network formation has broad implications ranging from interactions among different individuals, firms and also websites. People derive benefits from connecting with each other. These benefits may be pleasure, information, favors and so on. The goal of this paper is to study a model of network formation in which the amount of benefits agents obtain from their connections depends on those agents' effort in asking as well as their connections' effort in giving.

In an effort to fix ideas, we interpret benefits as information throughout this paper. We consider the following problem in which each individual in the network is endowed with an intrinsic value and this value is publicly observed. For example, we know doctors know medicine and engineers know engineering. Each individual has a budget constraint which implies a limited amount of resources spent on obtaining information from others. Apart from one's own intrinsic value, an individual also wants to get some information from others. So people have to decide whom to connect with and how much effort to spend in establishing the relationships. Once the relationship is established between two individuals, information naturally flows from one to the other. The amount of information the other agent gets depends on the nature of the relationship, which is represented by a relationship quality function. The relationship quality function has two arguments: the effort of asking information and the effort of giving information. For example, the share of information flowing from agent j to agent i depends on agent i's effort in asking as well as agent j's effort in giving.

The utility of each individual is regarded as the total amount of information this individual has. We describe it as the sum of one's intrinsic value and the amount of information one gets from the connections to other agents. The amount of information one obtains from a connection is quantified by the product of the relationship quality function and the total amount of information the connection has. Thus, the information agents have also depends on who their connections are. Utilizing this setup, this paper aims at studying how agents behave when they face the decisions of both asking and giving and how the socially optimal network structure would respond.

The model of asking and giving investigated in this paper benefits from Rogers (2005). Rogers examines the behavior of asking and giving in separate models and claims that inefficiency comes from the behavior of giving. In the model of asking, agents receive information through the relationships they establish while in the model of giving, agents confer information through the relationships. In the real world, however, interaction between two agents often exhibits not only their effort spent in asking but also their effort in giving. It is not necessary to separate asking behavior from giving behavior. Thus, this work generalizes Rogers' models by incorporating two decisions – asking and giving – into one model. More importantly, by considering asking behavior and giving behavior at the same time, this paper provides a new perspective into some aspects that cannot be obtained in separate models, i.e., the relationship between asking behavior and giving behavior.

Much of the literature is based on the assumption that the information one obtains from the connections is exogenously determined (Block and Dutta, 2005; Brueckner, 2006; Jackson and Rogers, 2005; Jackson and Wolinsky, 1996). However, Rogers' models contribute to the literature by assuming the amount of information one confers to other agents is endogenously determined by the network structure. In other words, what a person can offer depends on who his/her acquaintances are. For example, what teachers can teach depends on their own knowledge as well as the knowledge they learn from their friends, colleagues and so on. This reasonable assumption leads to several important characteristics. First, all paths between two agents generate benefits. There is limited work in the literature considering redundancy. For instance, in the connection model of Jackson and Wolinsky (1996), the benefit one agent gets from another only depends on the number of links in the shortest path between them. However, other paths may also generate added value. Second, we take "feedback effects" into account, whereby the benefits associated with a relationship are counted many times. Feedback effects appear frequently in daily life and have important implications. A simple example is two students working on a problem. While they discuss with each other and exchange information, the effort one student devotes in conveying information to the other benefits both of them. They can keep exchanging their new thoughts until finally the problem is solved. On the contrary, the friendship network introduced by Brueckner (2006) considers only the benefits from all direct and indirect friends but ignores feedback effects.

The analysis of network formation in this paper is based on the approach proposed by Bala and Goyal (2000) – using the concept of Nash network. In their study, the costs and benefits of links are exogenously given. Bloch and Dutta (2009) then study a network in which the quality of links is endogenously chosen by the agents. Our study differs fundamentally from their work in two ways. First, the network they analyze is a two-way flow network where both parties at two sides of the link share the same link quality. However, the model of asking and giving is a one-way flow network and the flow of information is directed. Second, the utility in their work is modeled as the link strength of the shortest path between two agents. In this paper, we use a different utility structure: one's utility is the sum of intrinsic value and the information obtained from all connections.

Different assumptions can be made on the relationship quality function. Following Rogers' study, we model the relationship quality function under two specifications: the concave specification and the linear specification. Under the concave specification, the relationship quality function is concave. We find that people spend more effort in asking for help from those with more information, and spend more effort in offering information to those from whom they can receive more information. A social planner would want people to spend more effort in giving if they have better relationships with others. Another finding is related to the relationship between asking behavior and giving behavior. If an agent benefits less from receiving information than his/her connection does, then this agent's effort in asking information from this connection is increasing with the effort in giving information to this connection.

Under the linear specification, the relationship quality function is linear. With this assumption, people only spend effort in asking information. To make the model tractable, we consider a simple network with only three agents. Following Brueckner (2006), we introduce asymmetries into the network by considering an *endowment-attractive* case and a *budget-attractive* case. In both cases, there is an attractive agent while the other two are identical. The attractive agent may either have higher intrinsic quality (*endowment-attractive*) or more budget to spend (*budget-attractive*). Although this setup may appear simple, it has some interesting implications. We find in both cases, non-attractive agents spend all their resources connecting with the attractive agents. This is consistent with the finding of Breuckner (2006). Moreover, in both cases, efficient networks coincide with Nash network.

The rest of the paper is organized as follows: In Section 2, we set up the model. Section 3 discusses the results. The last section summarizes and concludes our findings.

2.2 Model setup

Since this analysis is based on the two models initiated by Rogers (2005), we adopt the notation used in his paper. There is a finite set of agents $\boldsymbol{N} = \{1, \dots, n\}$, which are identified with nodes of a *network* g. A network is a collection of nodes and links which represent the network relations among agents. Suppose agents are endowed with intrinsic values $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \boldsymbol{R}_{+}^{\boldsymbol{n}}$ and budgets $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n) \in \boldsymbol{R}_{+}^{\boldsymbol{n}}$. This setting introduces heterogeneity and can be interpreted as each individual in the society holds a certain amount of information and his resources such as time and money are limited. People know how much and what kind of information they can get from those with different professions. Moreover, resources like wages for different jobs are available on the internet. As a result, it is reasonable to assume further that agents' qualities and budgets are publicly observed. Individuals obtain information from others at the expenses of their resources. Thus, agents have to decide how to allocate their limited resources $\boldsymbol{\beta}$ in establishing relationships with others. We assume agents have two ways to spend their effort: asking for information from others and giving information to others. Formally, ϕ_{ij}^{a} is the amount of resources agent i spends on asking information from j and ϕ_{ij}^{g} is the resources i spends on giving information to j. Each agent i has strategies $\boldsymbol{\phi_i^a} = (\phi_{i1}^{a}, \dots, \phi_{in}^{a})$ and $\boldsymbol{\phi_i^g} = (\phi_{i1}^{g}, \dots, \phi_{in}^{g})$ that satisfy $\phi_{ii} = \phi_{ii}^{a} + \phi_{ii}^{g} = 0, \phi_{ij} = \phi_{ij}^{a} + \phi_{ij}^{g} \geq 0$ for all $i, j \in \boldsymbol{N}$. The budget constraint $\sum_j \phi_{ij} \leq \beta_i$ for all $i \in \boldsymbol{N}$. $\boldsymbol{\phi_i^a}$ represents the allocation of resources spent on asking information from others while $\boldsymbol{\phi_i^g}$ represents the allocation of resources used for giving information to others. $\phi_{ii} = 0$, implying that agents know their own quality and they do not need to ask information from themselves.

Once the relationship is established between two individuals, information flows from one to the other. The amount of information the other agent gets depends on a relationship quality function. The relationship quality function has two arguments: one agent's effort in giving information as well as the other agent's effort in asking information. Formally, the share of information flowing from agent j to i is modeled as the relationship quality function, $f(\phi_{ij}^a$, $\phi_{ji}^g)$, which satisfies $0 \leq f < 1$, and is strictly increasing in both arguments.

We define agent j to be *directly connected* to agent i if there is a link directed from j to i in a network g, or $f\left(\phi_{ij}^a,\ \phi_{ji}^g\right) > 0$. In such case, the link is denoted ji. Agent j is agent i's *direct neighbor* if agent j is directly connected to agent i. Figure 2.1 below describes the flows of information between two agents, i and j, if they are directly connected to each other. The upper arrow pointing to agent j from agent i represents the link ij, indicating that agent i is directly connected to agent j. The lower arrow, reversely, represents link ji, showing that agent j is directly connected to agent i. The share of information flowing from agent j to agent i is $f\left(\phi_{ij}^a,\ \phi_{ji}^g\right)$. Thus, there may exist two links between two agents as shown in the figure below. If $f\left(\phi_{ij}^a,\ \phi_{ji}^g\right) = 0$, then there is no link or no flow of information from j to i.

The network g is a *complete network* if every agent $i \in N$, is directly connected to every other agent $j \in \boldsymbol{N}\backslash\{i\}$. A *path* in g connecting agent i_1 and i_n is a sequence of distinct nodes $\{i_1, i_2, \dots, i_n\}$ and directed links $\{i_1 i_2,\ i_2 i_3, \dots, i_{n-1} i_n\}$ in the network g. The *distance* from i_1 to i_n, denoted $d(i_1, i_n)$, is then the minimum number of links among all the possible paths existing between agent i_1 and i_n. So the distance between direct neighbors is 1. If the distance from agent i to j is greater than 1, then we define agent i as agent j's *indirect neighbor* which means agent i is indirectly connected to agent j .

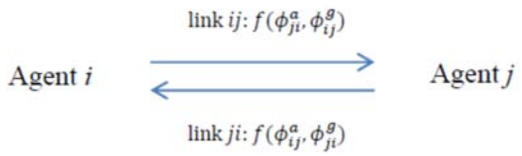


Figure 2.1: The flows of information

Different assumptions can be made on the relationship quality function $f(\phi_{ij}^{a},\ \phi_{ji}^{g})$. Throughout this paper, relationship quality is assumed to be an additively separable function of investments, i.e., $f(\phi_{ij}^{a},\ \phi_{ji}^{g}) = h_1(\phi_{ij}^{a}) + h_2(\phi_{ji}^{g})$. So the effort in asking and the effort in giving are assumed to be substitutes. With endogenous relationship quality, agents are able to adjust their decisions for resources allocation. We focus on two specifications of relationship quality. In the first case, each separable function is concave, which indicates diminishing returns to investment in establishing a link. This assumption is reasonable since, for example, at a low level of effort, it is easy for an agent to get some basic information from partners but after that it becomes increasingly harder to gain additional information. In the second case, the relationship quality function is linear. The effort in asking and the effort in giving are assumed to be perfect substitutes. A formal description of the two specifications is stated as follows:

Assumption 2.1 (Concave specification): $h_1(\cdot)$ *and* $h_2(\cdot)$ *are continuously differentiable, strictly increasing, and strictly concave,* $h_1(0) = h_2(0) = 0$ and $h_1(\cdot) + h_2(\cdot) < 1$so that $f \in [0,1)$. *Also,* $\lim_{x\to 0} h_1'(x) = \infty$, *and* $\lim_{x\to 0} h_2'(x) = \infty$.

Assumption 2.2 (Linear specification): *The functions* $h_1(\cdot)$ *and* $h_2(\cdot)$ *are both the identity mapping. In this case,* $\beta_i \in \left(0, \frac{1}{2}\right)$ *so that* $f(\phi_{ij}^{a},\ \phi_{ji}^{g}) = \phi_{ij}^{a} + \phi_{ji}^{g} \in [0,1)$.

Total utility of agent i is defined as the sum of i's intrinsic value and the information derived through i's direct neighbors via the network structure. Let u_i denote the total utility of agent i, v_i the information derived from direct neighbors so that $u_i = \alpha_i + v_i$. If agent j is directly connected to agent i, then $f(\phi_{ij}^a, \phi_{ji}^g)$ of agent j's information flows to agent i through the link ji. In other words, we get $u_j f(\phi_{ij}^a, \phi_{ji}^g)$ as the amount of information agent i obtains from agent j. Since $u_j = \alpha_j + v_j$, we rewrite $u_j f(\phi_{ij}^a, \phi_{ji}^g)$ as $(\alpha_j + v_j) f(\phi_{ij}^a, \phi_{ji}^g)$, and then sum up all information from direct neighbors to obtain $v_i = \sum_j (\alpha_j + v_j) f(\phi_{ij}^a, \phi_{ji}^g)$. Therefore, the total information of agent i from a network g is given by,

$$u_i = \alpha_i + v_i = \alpha_i + \sum_j (\alpha_j + v_j) f(\phi_{ij}^a, \phi_{ji}^g)$$

where the first term is agent *i*'s intrinsic value, and the second term is the information *i* gets from direct neighbors.

Collecting the above equations in matrix notation we obtain $\boldsymbol{u} = \boldsymbol{\alpha} + f(\boldsymbol{\Phi})\boldsymbol{u}$, where $\boldsymbol{u} = (u_1, \dots, u_n)'$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)'$ are column vectors of utilities and intrinsic values respectively. $f(\boldsymbol{\Phi})$ denotes the matrix with elements $f(\phi_{ij}^a,\ \phi_{ji}^g)$, i.e., $f(\boldsymbol{\Phi})$ is the network structure generated by strategy profile $\boldsymbol{\Phi} = [(\phi_{ij}^a,\ \phi_{ji}^g)]$. Solving for $\boldsymbol{u}$ yields $\boldsymbol{u} = (1 - f(\boldsymbol{\Phi}))^{-1}\boldsymbol{\alpha}$. Letting $\boldsymbol{A} = (1 - f(\boldsymbol{\Phi}))^{-1}$ with elements a_{ij}, we have $\boldsymbol{u} = \boldsymbol{A}\ \boldsymbol{\alpha}$. The matrix $\boldsymbol{A}$ can be rewritten as $\sum_{p=0}^{\infty} f(\boldsymbol{\Phi})^p$ since $I - f(\boldsymbol{\Phi})$ satisfies the well-known dominant diagonal condition. Thus, $(f(\boldsymbol{\Phi})^p)_{ij}$ depicts the total weight of directed paths from *j* to i that have length p. $\boldsymbol{A}_{ij}$ represents the total weight of all paths from *j* to i. Also, since $f \in [0,1)$, the matrix $\boldsymbol{A}$ is convergent.

Another object in the study of networks is to examine the *total value* of a network. The total value in this paper is defined as the sum of individual utilities, $U(g) = \sum_i u_i(g)$. Next we proceed to define Nash networks and efficient networks.

Definition 2.1 *A strategy profile* $\boldsymbol{\Phi}_i^*(g)$ *is said to be a best response of agent i against the strategy* $\boldsymbol{\Phi}_{-i}(g)$ *if* $u_i\big(\boldsymbol{\Phi}_i^*(g),\ \boldsymbol{\Phi}_{-i}(g)\big) \geq u_i\big(\boldsymbol{\Phi}_i'(G),\ \boldsymbol{\Phi}_{-i}(G)\big)$*, for all* $\boldsymbol{\Phi}_i'(g)$.

Definition 2.2 *A network* g *is a Nash network if the strategy for each agent in the network is a best response.*

So in a Nash network, agents maximize their own utilities, given other agents' strategies. Agents have no incentive to deviate from their equilibrium behavior.

Definition 2.3 *A network* g *is efficient if* $U(g) \geq U(g')$ *for all possible structure* g'.

An efficient network has the highest total value among all possible network structures.

2.3 Results

In this section we state our main results. Subsection 3.1 investigates the concave specification. We first characterize the equilibrium and efficient network under the concave specification. Then we study the relationship between agents' asking behavior and giving behavior. In Subsection 3.2, we look at the linear case. To make the model tractable, we follow Brueckner (2006) by introducing asymmetries into a network with only three agents.

2.3.1 Concave specification

In the model of giving from Rogers (2005), there are many Nash networks that are not complete. For example, if there is no link pointing to agent j from i, then j has no incentive to give any information to i for reason that agent j cannot obtain any benefits from doing so. In other words, agent j and agent i stay in different partitions. But this is not true when we incorporate the behavior of asking into the model. Now we characterize the set of Nash networks and the efficient networks and provide some intuition for the results. The formal proofs are given in the Appendix. The first proposition describes the Nash networks under the concave specification.

Proposition 2.1 *Under the concave specification, the Nash network is complete, which satisfies the conditions* $\sum_j \phi_{ij} = \beta_i$ $for\ all\ i \in N, and$ *for all* $i, j, j' \in N$

$$h_1'(\phi_{ij}^a)u_j = h_1'\left(\phi_{ij'}^a\right)u_{j'}, \qquad (2.1)$$

$$a_{ij}h_2'(\phi_{ij}^g) = a_{ij'}h_2'\left(\phi_{ij'}^g\right) \qquad (2.2)$$

Proof See Appendix. ■

The assumptions of $\lim_{x\to 0} h_1'(x) = \infty$ and $\lim_{x\to 0} h_2'(x) = \infty$ ensure that the Nash network is complete. The marginal return is relatively high at a low level of investment. So it is beneficial for every agent to spare some effort interacting with every other agent. Then the solution of the utility maximization problem for every agent is interior. First order conditions lead to equation (2.1) and (2.2). Marginal utilities are equal across different links. Otherwise, agents would have incentive to spend more effort on the relationships with higher marginal utilities. Equation (2.1) indicates that agents spend more resources in asking information from those with more information. Equation (2.2) implies that agents spend more resources in giving information to those who have stronger paths back. In the real world, people usually spend more effort in asking for help from those with more information, and spend more effort in offering help to those who are more likely to help back.

The next proposition describes the efficient networks under the concave specification.

Proposition 2.2 *Under the concave specification, the socially efficient network is complete, and satisfies the conditions* $\sum_j \phi_{ij} = \beta_i$ $for\ all\ i \in N, and$

$$h_1'(\phi_{ij}^a)u_j = h_1'\left(\phi_{ij'}^a\right)u_{j'} \qquad (2.3)$$

$$h_2'(\phi_{ij}^g)\sum_k a_{kj} = h_2'\left(\phi_{ij'}^g\right)\sum_k a_{kj'} \quad (2.4)$$

Proof. See Appendix. ■

The socially efficient network is also complete. Condition (2.3) is the same as condition (2.1). Equation (2.4) is different from equation (2.2) in the sense that a social planner would suggest agents spend to more effort giving information to those who have better relationships with others, which results in benefiting the entire society as a whole. Equation (2.2) and (2.4) indicate the possible differences between Nash networks and efficient networks.

The first order conditions in Nash networks also shed light on the relationship between agents' asking behavior and giving behavior. We start with a perfectly symmetric case as shown in Example 1.

Example 2.1 *Assume that $\beta_i = \beta$ and $\alpha_i = \alpha$ for all $i \in N$. If $h_1(\cdot) = h_2(\cdot)$, then a complete network with $\phi_{ij}^a = \phi_{ij'}^a$ and $\phi_{ij}^g = \phi_{ij'}^g$ for all $i, j, j' \in N$ is a possible Nash network. Moreover, $\phi_{ij}^a > \phi_{ij}^g$.*

Proof See Appendix. ■

This result is intuitive. In a perfectly symmetric environment, agents should spend more effort in asking information from than giving information to the same person. Because asking results in direct benefits while benefits of giving come from the feedback effect. During the transmission of information, the benefits of giving depreciate more than the benefits of asking. Since $h_1(\cdot)$ and $h_2(\cdot)$ have the same shape, agents would rather spend more effort in asking information. If asking behavior and giving behavior weigh differently, i.e., $h_1(\cdot) > h_2(\cdot)$, then agents spend even more effort in asking and even less information in giving. If the reverse is true, i.e., $h_1(\cdot) < h_2(\cdot)$, the result can be ambiguous, depending on the relative shape of the

$h_1(\cdot)$ and $h_2(\cdot)$. Thus, whether there is a relationship between asking behavior and giving behavior is an interesting question which cannot be investigated in separate models of asking and giving as in Rogers (2005). Proposition 2.3 describes the relationship between the agents' effort spent in asking and giving in a general environment.

Proposition 2.3 *Under the concave specification, in any Nash network, if* $a_{ij}^* u_j^* < a_{jj}^* u_i^*$*, then agent i's effort in asking information from agent j is increasing with agent i's effort in giving information to agent j .*

Proof See Appendix. ■

The term $a_{ij}^* u_j^*$ describes i 's utility of obtaining information from j and $a_{jj}^* u_i^*$ is to be interpreted as j 's utility of obtaining information from i. Since $a_{ij}^* u_j^* < a_{jj}^* u_i^*$, j 's benefit of obtaining information from i is greater than i 's benefit of obtaining information from j. In this case, the proposition implies that agent i's effort in asking information from agent j is increasing with i' effort spent in giving. In other words, if agent i decides to spend more resources in asking information from j, then i would also spend more resources in giving information to j. Because j benefits more from receiving information than i does. After j receives those benefits, j has more information to return favors back.

If $a_{ij}^* u_j^* > a_{jj}^* u_i^*$, the relationship between the behavior of asking and giving is uncertain. It depends on the difference between the asking component of the relationship quality function, $h_1(\cdot)$ and the giving component of the relationship quality function, $h_2(\cdot)$.

2.3.2 Linear specification

Under the concave specification, every agent has the incentive to both ask and give. Because at a low level of effort, the marginal return is high. When the relationship quality is a linear function, agents do not necessarily have the incentive to give information to others. The benefit of giving comes from the feedback effect: if your neighbors' information increases by receiving information from you, you would be better off by being able to receive more information from your neighbors. Under the linear specification, the effort in asking is a perfect substitute for the effort in giving. So agents would rather directly ask for information since the benefit of giving depreciates more through paths back to themselves. The following proposition describes this phenomenon.

Proposition 2.4 *Under the linear specification, each agent only spends effort in asking information from other agents. In other words,* $\Phi_{ij}^{g} = 0,\ for\ all\ i\ and\ j.$

Proof. See Appendix. ■

So under the linear specification, every model of asking and giving is reduced to a model of asking. In Nash networks, nobody wants to spend effort in giving information to others. The benefits of giving information depreciate during the transmission process and asking directly reduces depreciation.

Under the concave specification, the analysis is easier because of the assumptions $\lim_{x\to 0} h_1'(x) = \infty$ and $\lim_{x\to 0} h_2'(x) = \infty$. The solution of the utility maximization problem is interior. This is not guaranteed under the linear specification, however. The characterization of Nash networks is difficult. To make the model tractable, we consider two types of asymmetries in a small universe of agents with $n = 3$, following Brueckner (2006). In the real world, network patterns often exhibit heterogeneity. Some agents have higher intrinsic values while others may have higher budget. In the first type of asymmetry, one agent is an "*endowment-attractive* agent", having higher intrinsic value than the other agents, who remain symmetric. In the second type, one individual is a "*budget-attractive*" agent, with more resources to allocate in asking than the other agents, who again remain symmetric. Formal assumptions are given as follows.

Assumption 2.3 (*Endowment-Attractive* Case) *Suppose there are three agents, with agent 1 being the endowment-attractive agent. The intrinsic value of agent 1 is* α_1 *while the resource budget is* β. $\alpha_1 > 0$ *and* $0 < \beta < ½$. *The other two agents, agent 2 and agent 3 are identical, with the same intrinsic value* α *and the same budget* β. $\alpha_1 > \alpha$.

Notice the only difference between the *endowment-attractive* agent and the others is that the *endowment-attractive* agent has higher intrinsic value.

Assumption 2.4 (*Budget-Attractive* Case): *Suppose there are three agents, with agent 1 being the budget-attractive agent. The intrinsic value of agent 1 is* α *while resource budget is* β_1. $\alpha > 0$ *and* $0 < \beta_1 < ½$. *The other two agents, agent 2 and agent 3 are identical, with the same intrinsic value* α *and the same budget* β. $0 < \beta < \beta_1$.

The only difference between the *budget-attractive* agent and the others is that the *budget-attractive* agent has more resources to spend. This discussion proceeds by characterizing Nash networks and efficient networks under assumptions 3 and 4.

Proposition 2.5 *In both endowment-attractive and budget-attractive networks under the linear specification,* (1) *the non-attractive agents spend all effort asking information from agent 1;* (2) *Nash networks and efficient networks coincide.*

Proof. See Appendix. ■

According to Proposition 2.4, agents under linear specification only spend effort in asking. Proposition 2.5 further shows that when asymmetries are introduced, non-attractive agents only connect with attractive agent. To understand this conclusion, consider first the *endowment-attractive* case. Agent 1 has higher intrinsic value, which makes him/her more attractive to the other agents. Other agents are willing to spend all their effort in asking information from agent 1 so that they can get more information. In the *budget-attractive* case, agent 1 is more able to get information from other agents because of the higher budget level, which again makes him/her more attractive. As a result, other agents are willing to spend all their effort in asking information from agent 1. This is consistent with the results from Brueckner (2006): non-attractive agents spend more effort linking with the attractive agent.

2.4 Conclusion

In this paper we study a setting in which agents spend resources in both giving information to and asking information from connections to their neighbors. We generalize the models of link formation of Rogers (2005) by combining the model of asking and the model of giving and allowing that giving and asking choices can be made separately and simultaneously by each agent. We focus on two specifications: the concave specification and the linear specification. Under the concave specification, the results show that people usually spend more effort in asking for help from those with more information, and spend more effort in offering help to those from whom they can receive more information. A social planner wants people to spend more effort in giving if they have better aggregate relationships with others. If an agent's direct neighbor benefits more from receiving information, then this agent's effort in asking information from is increasing with the effort in giving information to this neighbor.

Next, we turn our attention to the linear case. In the linear case, we find people only spend resources in asking because the behavior of giving suffers more depreciation. In an effort to further study the impact of asymmetries and make the model tractable, we follow Brueckner (2006) by considering an *endowment-attractive* case and a *budget-attractive* case in a small universe of agents with $n = 3$. In the *endowment-attractive* case, the attractive agent has higher intrinsic value while in the *budget-attractive* case, the attractive agent has higher budget level. In both cases, non-attractive agents spend all their resources connecting with the attractive agents. This conclusion is consistent with the finding of Breuckner (2006). Moreover, in both cases, efficient networks coincide with Nash networks.

This paper analyzes of network formation under heterogeneous environments. Rogers (2005) takes redundancy and feedback effects into account. The model of asking and giving not only benefits from Rogers' work but also provides an approachable avenue for studying the relationship between asking and giving behavior. This provides many directions for future work. First, other specifications of the link quality function may be assumed. Another extension may be making people's giving behavior interdependent. For example, if one agent refuses to help the other agent, then the other agent's willingness to offer help will be reduced. The results in the linear case are for a simple network with only three agents. Thus another future direction can be based upon examining whether the results still hold in a larger universe of agents.

2.5 Appendix

Proof of Proposition 2.1: The utility maximization problem of i, taking the strategies of others as given, is

$$\max_{\phi_i^a,\phi_i^g} u_i \;\; s.t. \sum\nolimits_j \phi_{ij} \le \beta_i$$

The assumption that $\lim_{x\to 0} h_1'(x) = \infty$, and $\lim_{x\to 0} h_2'(x) = \infty$ ensures that the solution for this problem is interior. Then the first order conditions are $\frac{\partial u_i}{\partial \phi_{ij}^a} = \frac{\partial u_i}{\partial \phi_{ij'}^a}$, and $\frac{\partial u_i}{\partial \phi_{ij}^g} = \frac{\partial u_i}{\partial \phi_{ij'}^g}$, for all $j, j' \neq i$.

Since $\frac{\partial \boldsymbol{u}}{\partial \phi_{ij}^a} = \frac{\partial A}{\partial \phi_{ij}^a}\alpha$, and $\frac{\partial \boldsymbol{u}}{\partial \phi_{ij}^g} = \frac{\partial A}{\partial \phi_{ij}^g}\alpha$, it follows with

$$\frac{\partial \boldsymbol{u}}{\partial \phi_{ij}^{a}} = A\frac{\partial f(\Phi)}{\partial \phi_{ij}^{a}}A\alpha = h_1{}'(\phi_{ij}^{a})u_j\begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

and

$$\frac{\partial \boldsymbol{u}}{\partial \phi_{ij}^{g}} = A\frac{\partial f(\Phi)}{\partial \phi_{ij}^{g}}A\alpha = h_2{}'(\phi_{ij}^{g})u_i\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

The i^{th} components of the derivatives are simply

$$h_1'(\phi_{ij}^{a})u_j = h_1'\left(\phi_{ij'}^{a}\right)u_{j'}, \text{and } a_{ij}h_2'(\phi_{ij}^{g}) = a_{ij'}h_2'\left(\phi_{ij'}^{g}\right) \quad \blacksquare$$

Proof of Proposition 2.2: The utility maximization problem is

$$\max_{\phi_i^a,\phi_i^g}\sum_i u_i \;\; s.t. \sum_j \phi_{ij} \leq \beta_i$$

An efficient network exists since the choice sets are compact and $\sum_i u_i$ is continuous. The assumption that $\lim_{x\to 0} h_1'(x) = \infty$, and $\lim_{x\to 0} h_2'(x) = \infty$ ensures that the solution for this problem is interior. Then the first order conditions are $\frac{\partial \sum_i u_i}{\partial \phi_{ij}^{a}} = \frac{\partial \sum_i u_i}{\partial \phi_{ij'}^{a}}$, and $\frac{\partial \sum_i u_i}{\partial \phi_{ij}^{g}} = \frac{\partial \sum_i u_i}{\partial \phi_{ij'}^{g}}$, for all $j, j' \neq i$.

From the proof of proposition 2.1, we have

$$h_1'(\phi_{ij}^{a})u_j = h_1'\left(\phi_{ij'}^{a}\right)u_{j'}, \qquad h_2'(\phi_{ij}^{g})\sum_k a_{kj} = h_2'\left(\phi_{ij'}^{g}\right)\sum_k a_{kj'}. \quad \blacksquare$$

Lemma 1 $\frac{\partial a_{ij}}{\partial \Phi_{kl}^{a}} = h_1'(\phi_{kl}^{a})a_{ik}a_{lj}$; $\frac{\partial a_{ij}}{\partial \Phi_{kl}^{g}} = h_2'(\phi_{kl}^{g})a_{il}a_{ik}$.

Proof. Following Rogers (2005) and differentiating $\boldsymbol{AA^{-1}} = \boldsymbol{I}$, we get

$$\frac{\partial \boldsymbol{A}}{\partial \Phi_{kl}^{a}}\boldsymbol{A}^{-1} + \boldsymbol{A}\frac{\partial \boldsymbol{A}^{-1}}{\partial \Phi_{kl}^{a}} = 0 \text{ and } \frac{\partial \boldsymbol{A}}{\partial \Phi_{kl}^{g}}\boldsymbol{A}^{-1} + \boldsymbol{A}\frac{\partial \boldsymbol{A}^{-1}}{\partial \Phi_{kl}^{g}} = 0.$$

Right-multiplying by $\boldsymbol{A}$ and rearranging produces

$$\frac{\partial \boldsymbol{A}}{\partial \Phi_{kl}^{a}} = \boldsymbol{A}\frac{\partial f(\boldsymbol{\Phi})}{\partial \Phi_{kl}^{a}}\boldsymbol{A} \text{ and } \frac{\partial \boldsymbol{A}}{\partial \Phi_{kl}^{g}} = \boldsymbol{A}\frac{\partial f(\boldsymbol{\Phi})}{\partial \Phi_{kl}^{g}}\boldsymbol{A},$$

since $\boldsymbol{A} = (I - f(\boldsymbol{\Phi}))^{-1}$.

The results are just the scalar forms of the above two equations. ■

Proof of Example 2.1: The first order conditions indicate

$$h_1'(\phi_{ij}^{a})u_j a_{ii} = h_2'(\phi_{ij}^{g})u_i a_{ij}$$

In a perfectly symmetric case, $u_j = u_i$. Since $a_{ii} > a_{ij}$, $h_1'(\phi_{ij}^{a}) < h_2'(\phi_{ij}^{g})$. If $h_1(\cdot) = h_2(\cdot)$, $\phi_{ij}^{a} > \phi_{ij}^{g}$. Because $h_1(\cdot)$ and $h_2(\cdot)$ are concave.

Proof of Proposition 2.3: The first order conditions are

$$\frac{\partial u_i}{\partial \phi_{ij}^{a}} = \frac{\partial u_i}{\partial \phi_{ij'}^{a}} = \frac{\partial u_i}{\partial \phi_{ij}^{g}} = \frac{\partial u_i}{\partial \phi_{ij'}^{g}}, \qquad \forall\, i, j, \text{and } j'.$$

The marginal utilities agents *i* obtains from each link should be the same. Otherwise, agent *i* has the incentive to change his/her behavior. From the proof of Proposition 2.1,

$$h_1'(\phi_{ij}^{a})u_j a_{ii} = h_1'(\phi_{ij'}^{a})u_{j'} a_{ii} = h_2'(\phi_{ij}^{g})u_i a_{ij} = h_2{}'(\phi_{ij'}^{g})u_i a_{ij'}.$$

Differentiating $h_1'(\phi_{ij}^{a})u_j a_{ii} = h_2'(\phi_{ij}^{g})u_i a_{ij}$ yields

$$h_1''(\phi_{ij}^{a})\frac{\partial \Phi_{ij}^{a}}{\partial \Phi_{ij}^{g}}u_j a_{ii} + h_1'(\phi_{ij}^{a})\frac{\partial u_j}{\partial \Phi_{ij}^{g}}a_{ii} + h_1'(\phi_{ij}^{a})u_j\frac{\partial a_{ii}}{\partial \Phi_{ij}^{g}} =$$

$$h_2''(\phi_{ij}^{g})u_i a_{ij} + h_2'(\phi_{ij}^{g})\frac{\partial u_i}{\partial \Phi_{ij}^{g}}a_{ij} + h_2'(\phi_{ij}^{g})u_i\frac{\partial a_{ij}}{\partial \Phi_{ij}^{g}}.$$

Rearrange the above equation so that we obtain the following expression describing the relationship between the effort spent on asking information from another agent and the effort spent on giving information to the same person:

$$\frac{\partial \Phi_{ij}^{a}}{\partial \Phi_{ij}^{g}} = \left[h_2''(\phi_{ij}^{g}) u_i a_{ij} + h_2'(\phi_{ij}^{g}) \frac{\partial u_i}{\partial \Phi_{ij}^{g}} a_{ij} + h_2'(\phi_{ij}^{g}) u_i \frac{\partial a_{ij}}{\partial \Phi_{ij}^{g}} - h_1'(\phi_{ij}^{a}) \frac{\partial u_j}{\partial \Phi_{ij}^{g}} a_{ii} \right.$$

$$\left. - h_1'(\phi_{ij}^{a}) u_j \frac{\partial a_{ii}}{\partial \Phi_{ij}^{g}} \right] / h_1''(\phi_{ij}^{a}) u_j a_{ii}.$$

Substituting $\frac{\partial u_i}{\partial \Phi_{ij}^{g}} = h_2'(\phi_{ij}^{g}) u_i a_{ij}, \frac{\partial u_j}{\partial \Phi_{ij}^{g}} = h_2'(\phi_{ij}^{g}) u_i a_{jj}$ from the proof of Proposition 2.1 and $\frac{\partial a_{ij}}{\partial \Phi_{ij}^{g}} = h_2'(\phi_{ij}^{g}) a_{ij} a_{ij}, \frac{\partial a_{ii}}{\partial \Phi_{ij}^{g}} = h_2'(\phi_{ij}^{g}) a_{ij} a_{ii}$ from Lemma 1,

$$\frac{\partial \Phi_{ij}^{a}}{\partial \Phi_{ij}^{g}} == \left[h_2''(\phi_{ij}^{g}) u_i a_{ij} + h_2'(\phi_{ij}^{g}) h_2'(\phi_{ij}^{g}) u_i a_{ij} a_{ij} + h_2'(\phi_{ij}^{g}) u_i h_2'(\phi_{ij}^{g}) a_{ij} a_{ij} \right.$$

$$\left. - h_1'(\phi_{ij}^{a}) h_2'(\phi_{ij}^{g}) u_i a_{jj} a_{ii} - h_1'(\phi_{ij}^{a}) u_j h_2'(\phi_{ij}^{g}) a_{ij} a_{ii} \right] / h_1''(\phi_{ij}^{a}) u_j a_{ii}.$$

Simplifying the above equation with $h_1'(\phi_{ij}^{a}) u_j a_{ii} = h_2'(\phi_{ij}^{g}) u_i a_{ij}$ produces

$$\frac{\partial \Phi_{ij}^{a}}{\partial \Phi_{ij}^{g}} = \left[h_2''(\phi_{ij}^{g}) u_i a_{ij} + h_2'(\phi_{ij}^{g}) h_1'(\phi_{ij}^{a}) u_j a_{ii} a_{ij} + h_2'(\phi_{ij}^{g}) h_1'(\phi_{ij}^{a}) u_j a_{ii} a_{ij} \right.$$

$$\left. - h_1'(\phi_{ij}^{a}) h_2'(\phi_{ij}^{g}) u_i a_{jj} a_{ii} - h_1'(\phi_{ij}^{a}) u_j h_2'(\phi_{ij}^{g}) a_{ij} a_{ii} \right] / h_1''(\phi_{ij}^{a}) u_j a_{ii}$$

$$= \left[h_2''(\phi_{ij}^{g}) u_i a_{ij} + h_2'(\phi_{ij}^{g}) h_1'(\phi_{ij}^{a}) u_j a_{ii} a_{ij} - h_1'(\phi_{ij}^{a}) h_2'(\phi_{ij}^{g}) u_i a_{jj} a_{ii} \right]$$

$$/ h_1''(\phi_{ij}^{a}) u_j a_{ii}$$

$$= \left[h_2''(\phi_{ij}^{g}) u_i a_{ij} + h_2'(\phi_{ij}^{g}) h_1'(\phi_{ij}^{a}) a_{ii} (u_j a_{ij} - u_i a_{jj}) \right] / h_1''(\phi_{ij}^{a}) u_j a_{ii}$$

Note that $h_1''(\phi_{ij}^{a}) < 0, h_2''(\phi_{ij}^{g}) < 0, h_1'(\phi_{ij}^{a}) > 0, h_2'(\phi_{ij}^{g}) > 0, u_j > 0$ and $a_{ij} > 0$

If $a_{ij} u_j < a_{jj} u_i$, this indicates $\frac{\partial \Phi_{ij}^{a}}{\partial \Phi_{ij}^{g}} > 0$.

Therefore, then agent *i*'s effort in asking information from agent *j* is increasing with agent *i*'s effort in giving information to agent *j*. ■

Proof of Proposition 4: Under linear specification, $f\left(\phi_{ij}^{a},\ \phi_{ji}^{g}\right)=\phi_{ij}^{a}+\phi_{ji}^{g}$. This implies,

$$f(\boldsymbol{\Phi})=\begin{bmatrix} 0 & \phi_{12}^{a}+\phi_{21}^{g} & \cdots & \phi_{1n}^{a}+\phi_{n1}^{g} \\ \phi_{21}^{a}+\phi_{12}^{g} & 0 & \cdots & \phi_{2n}^{a}+\phi_{n2}^{g} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{n1}^{a}+\phi_{1n}^{g} & \phi_{n2}^{a}+\phi_{2n}^{g} & \cdots & 0 \end{bmatrix}.$$

Since $\boldsymbol{u}=(1-f(\boldsymbol{\Phi}))^{-1}\boldsymbol{\alpha}$, we first construct $1-f(\boldsymbol{\Phi})$,

$$\boldsymbol{A}^{-1}=\mathbf{1}-f(\boldsymbol{\Phi})=\begin{bmatrix} 1 & -(\phi_{12}^{a}+\phi_{21}^{g}) & \cdots & -(\phi_{1n}^{a}+\phi_{n1}^{g}) \\ -(\phi_{21}^{a}+\phi_{12}^{g}) & 1 & \cdots & -(\phi_{2n}^{a}+\phi_{n2}^{g}) \\ \vdots & \vdots & \ddots & \vdots \\ -(\phi_{n1}^{a}+\phi_{1n}^{g}) & -(\phi_{n2}^{a}+\phi_{2n}^{g}) & \cdots & 1 \end{bmatrix}.$$

Recall $\boldsymbol{u}=\boldsymbol{A}\,\boldsymbol{\alpha}$. To get each agent's utility u_i, first we need to know the matrix $\boldsymbol{A}$:

$$\boldsymbol{A}=\frac{adj(\mathbf{1}-f(\boldsymbol{\Phi}))}{|\mathbf{1}-f(\boldsymbol{\Phi})|}.$$

Then $u_i=[\boldsymbol{A}]_{i\text{th row}}\cdot\ \boldsymbol{\alpha}=\frac{[adj(\mathbf{1}-f(\boldsymbol{\Phi}))]_{i\text{th row}}}{|\mathbf{1}-f(\boldsymbol{\Phi})|}\cdot\ \boldsymbol{\alpha}$. Notice ϕ_{ij}^{g} does not show up in $[adj(\mathbf{1}-f(\boldsymbol{\Phi}))]_{ith\text{ row}}$ but show up in $|\mathbf{1}-f(\boldsymbol{\Phi})|$ as negative elements. This indicates less ϕ_{ij}^{g} results in a higher utility. If each agent aims at maximizing his/her own benefits, he/she will not spend effort in giving, i.e., $\Phi_{ij}^{g}=0$. ■

Proof of Proposition 5: First, look at the *endowment-attractive* case. There is no interior solution for Nash equilibrium. To prove this, suppose the contrary is true: there is an interior solution. The three agents' utility should satisfy the following two relations:

$$u_1=\alpha_1+\beta u_x$$

$$u_x=\alpha+e_{x1}u_1+(\beta-e_{x1})u_x$$

Solving for u_x yields

$$u_x=\frac{\alpha_1 e_{x1}+\alpha}{1-\beta+e_{x1}-e_{x1}\beta}.$$

Taking derivative with respect to e_{x1} produces

$$\frac{\partial u_x}{\partial e_{x1}} = \frac{(1-\beta)(\alpha_1 - \alpha)}{(1-\beta+e_{x1}-e_{x1}\beta)^2} > 0$$

So agent 2 or 3 can get more benefit as they spend more effort in asking information from the *endowment-attractive* agent. A contradiction arises, ruling out the initial assumption. In a Nash network, both agent 2 and agent 3 spend all their effort asking information from agent 1.

In an efficient network, the total utility of all the three agents is maximized. That is,

$$\max_{e_{x1}} \; u_1 + 2u_x.$$

Taking derivative of the total utility with respect to e_{x1} produces

$$\frac{\partial\,(u_1 + 2u_{x)}}{\partial e_{x1}} = \frac{\beta(1-\beta)(\alpha_1-\alpha) + 2(1-\beta)(\alpha_1-\alpha)}{(1-\beta+e_{x1}-e_{x1}\beta)^2} > 0$$

So total utility increases as agent 2 and agent 3 spend more effort in asking information from the *endowment-attractive* agent. Hence, in an efficient network, both agent 2 and agent 3 spend all their effort asking information from agent 1. Efficient networks coincide with Nash networks.

Now, we look at the *budget-attractive* case. Likewise, there is no interior solution for Nash equilibrium. To prove this, suppose the contrary is true: there is an interior solution. The three agents' utility should satisfy the following two relations:

$$u_1 = \alpha + \beta_1 u_x$$

$$u_x = \alpha + e_{x1}u_1 + (\beta - e_{x1})u_x$$

Solving for u_x yields

$$u_x = \frac{\alpha e_{x1} + \alpha}{1-\beta+e_{x1}-e_{x1}\beta_1}\,.$$

Taking derivative with respect to e_{x1} produces

$$\frac{\partial u_x}{\partial e_{x1}} = \frac{(\beta_1 - \beta)\alpha}{(1 - \beta + e_{x1} - e_{x1}\beta_1)^2} > 0$$

So agent 2 or 3 can get more benefit as they spend more effort in asking information from the *budget-attractive* agent. A contradiction arises, ruling out the initial assumption. In a Nash network, both agent 2 and agent 3 spend all their effort asking information from agent 1.

In an efficient network, the total utility of all the three agents is maximized. That is,

$$\max_{e_{x1}} \ u_1 + 2u_x.$$

Taking derivative of the total utility with respect to e_{x1} produces

$$\frac{\partial\,(u_1 + 2u_{x)}}{\partial e_{x1}} = \frac{\beta_1(\beta_1 - \beta)\alpha + 2(\beta_1 - \beta)\alpha}{(1 - \beta + e_{x1} - e_{x1}\beta)^2} > 0$$

So the total utility increases as agent 2 and agent 3 spend more effort in asking information from the *budget-attractive* agent. Hence, in an efficient network, both agent 2 and agent 3 spend all their effort asking information from agent 1. Efficient networks coincide with Nash networks. ■

CHAPTER 3: LINEAGE-BASED FRAGMENTATION AND COOPERATIVE BEHAVIOR IN RURAL CHINA[1]

3.1 Introduction

Developing countries, including China, make tremendous efforts to promote rural development and reduce poverty. Since the success of many economic endeavors, such as exchanges of goods and services and public-good provision, depends on cooperation, the study of cooperative behavior in rural areas is of great importance. In the presence of imperfect contract enforcement, informal institutions then play an important role in rural development. Existing studies have associated fractionalization, measured by ethnic, linguistic, and religious heterogeneity, with trust, economic growth and the quality of governance (Easterly and Levine, 1997; Alesina et al., 1999, 2002). In rural China, though the economy has made great strides towards modernization, many villages are still structured by a number of traditional lineage organizations, which results in lineage-based fragmentation. The goal of this paper is to examine how lineage-based fragmentation affects cooperative behavior in rural China.

To the best of our knowledge, this is the first paper that presents a full picture of cooperative behavior by examining both intra-group and inter-group cooperation. Bowles and Gintis (2008) argue that cooperation can take the form of mutually beneficial transactions that may fail to materialize without trust and reciprocity (intra-group); it can also take the form of public-good provision that requires agreement and collective action (inter-group). In this paper, we measure intra-group cooperation by the frequency of mutual help that occurs between lineage members, and inter-group cooperation by individuals' contribution to build village infrastructures and the share of village budget that is spent on village public goods.

[1] We are grateful to Sudipta Sarangi and R. Carter Hill for suggestions. We also thank Matthew Jackson and Francis Bloch for valuable comments at the Networks and Development Conference 2012.

There is a growing body of literature studying the impact of heterogeneity on provision of public goods. The findings generally indicate that heterogeneity in ethnicity, religion and social class undermines inter-group cooperation and public-good provision (Alesina et al., 1999; Banerjee, Iyer and Somanathan, 2005; Bandiera et al., 2005). Our findings are consistent with the literature in this perspective. We find that villages that are more diversified in terms of lineage composition spend a lower share of the village budget on village public goods and people in more diversified villages contribute less labor to build village infrastructures. Since both the provision of public goods in rural China and fragmentation are measured at the village level, our study presents a more convincing relationship between the two variables than some existing literature does. For example, Alesina (1999) studies the relationship between ethnic composition in U.S. metropolitan areas and the share of metropolitan government expenditure on public goods such as education. However, spending on education is mostly determined at a much more local level, school-districts. Under the mismatching scenario, it would be difficult to determine the causal relationship between fragmentation and provision of public goods.

The empirical studies on intra-group cooperation are rare. Conflict theory in sociology suggests that diversity fosters in-group solidarity as well as out-group distrust (Blalock, 1967). In other words, with growing diversity of the population, people stick to their own group more and trust others less. According to this theory, one would expect that the people who live in more diversified villages, in terms of lineage compositions, should be more willing to cooperate with the same lineage members than those who live in homogenous villages. However, our findings imply the opposite case. Putnam (2007) claims that the fundamental assumption behind conflict theory – in-group trust and out-group trust are negatively correlated – is unwarranted. In other words, bonding with own-group members is not necessarily at the cost of bridging with other groups. Our paper provides important empirical evidence for Putnam's arguments.

China serves as an excellent case for studying fractionalization because the lineage composition within a village is exogenously determined. The lineage culture in rural China can be dated back to hundreds of years ago. Extended families related in men's line live in one settlement and form a lineage. The size of a lineage ranges from a few to a few hundred households. All men in one lineage are descendants of a common ancestor, and, consequently bear the same surname. Over generations, the common surname becomes the lineage identity and promotes solidarity among lineage members (Peng, 2004). Shortly after communist China was founded in 1949, the central government set up administrative villages in order to strengthen the Party's rule and to build up the Commune system. The administrative villages, the lowest level of administrative agency in China, also serve as the lowest level of collective farming unit in the Commune system. To meet the needs of collective farming, administrative villages arbitrarily included one or more adjacent lineages (Wang, 2006).[2] Therefore, the lineage composition within a village is exogenously determined by the shock of China's administrative re-organization. In addition, in 1958, China enacted the household registration system, which inhibits free migrations and essentially ties rural people to the land where they were born. Thus, the lineage composition in rural villages has remained static since 1958.[3] The identification in this paper arises from the exogenous and predetermined fragmentation. By contrast, the measurements of fragmentation in the existing literature are usually endogenous. For instance,

[2] A very large lineage can be broken into several single-lineage villages.

[3] The household registration system has been partially relaxed since the 1980s. The surplus rural laborers pour to cities seeking non-agricultural jobs. However, rural workers do not have the same access as urban citizens to medication, pension, housing and children's schooling in cities, which makes permanent rural-to-urban migrations still extremely difficult. Most rural workers have to commute between cities and their original villages several times a year.

Easterly and Levine (1997) and Alesina et al. (2002) use ethnic divisions within a county; Miguel and Gugerty (2005) use ethno-linguistic diversity within a district in Kenya; Egel (201l) uses the number of tribes within a subdistrict in Yemen. The fragmentation within a country or a region is probably associated with other characteristics of the country or region that can directly affect the outcomes.

In addition, lineage-specific culture traditionally is more predominant in the South than in the North of China (Freedman, 1965).[4] The South-North divide enables us to apply the Difference-in-Difference (D-in-D) method to further refine the identification. First, we examine the difference in cooperative behavior between the people from lineage-heterogeneous villages and those from lineage-homogenous villages. Then we further investigate whether the difference is stronger in the South than in the North. If the answer is affirmative, this indicates that the lineage-based heterogeneity affects people's cooperative behavior. We also apply the D-in-D models to exclude the possibility that there may be other unobserved differences between homogeneous villages and heterogeneous villages that have impacts on people's cooperative behavior. Notice the D-in-D method does not assume that the lineage-heterogeneous villages and the lineage-homogenous villages are the same in all other aspects. Instead, the identification assumption is that the two kinds of villages can be different in other aspects but those differences, if there are any, do not vary from the South to the North. In the paper, we present evidence that the assumption holds.

[4] Freedman (1965) proposes three reasons to explain the South-North difference. First, the political center of China is usually established in the North. Hence, the South is far from formal government control. Second, rice-cultivation in the South demands extensive irrigation. Inter-household cooperation in irrigation could be the base from which the lineage organizations emerged. Third, the population in the South has many immigrants from the North. The exigency of frontier life could stimulate the development of lineages.

Using data from the Chinese Household Income Project Survey (CHIPS) 2002, we measure intra-lineage cooperation by the frequency of mutual help within a lineage. Two kinds of mutual help are considered: monetary help and non-monetary help that is time-consuming. We find that lineage-based fragmentation has a negative effect on the frequency of both monetary and non-monetary mutual help. It turns out villagers do not treat them differently when it comes to lineage obligations and enforcement. Inter-lineage cooperation is measured by villagers' physical contribution to public goods and the share of village budget spent on public goods. Our results show that lineage-based fragmentation has a negative effect on inter-lineage cooperative behavior as well. In other words, people in lineage-homogenous villages are more likely to engage in reciprocal behavior with their lineage members, as well as contribute to the provision of public goods that are jointly shared across lineages. We also find that the association between the lineage-based homogeneity and the cooperative behavior is stronger in the South.

The rest of the paper is organized as follows. Section 2 describes the data. We provide some background information about lineages and public-good provision in rural China. Section 3 presents the empirical models. Section 4 discusses the results. The last section concludes.

3.2 Background and data

We use data from the rural section of the CHIPS 2002 survey. In this portion of the survey, 9200randomly-selected households were interviewed from 961 villages in 22 provinces.[5] Figure 3.1 presents a map of the provinces in China and the surveyed provinces have been shaded. To measure the cultural differences between the South and the North, we separate 22 provinces into

[5] Although there are in total 34 province-level administrative units in China, the 22 provinces in CHIPS 2002 provide a nationally representative sample. The 22 provinces were selected from four distinct regions in China---metropolitan region, eastern region, central region, and western region --- to reflect variations in economic development and geography (Li et al, 2008).

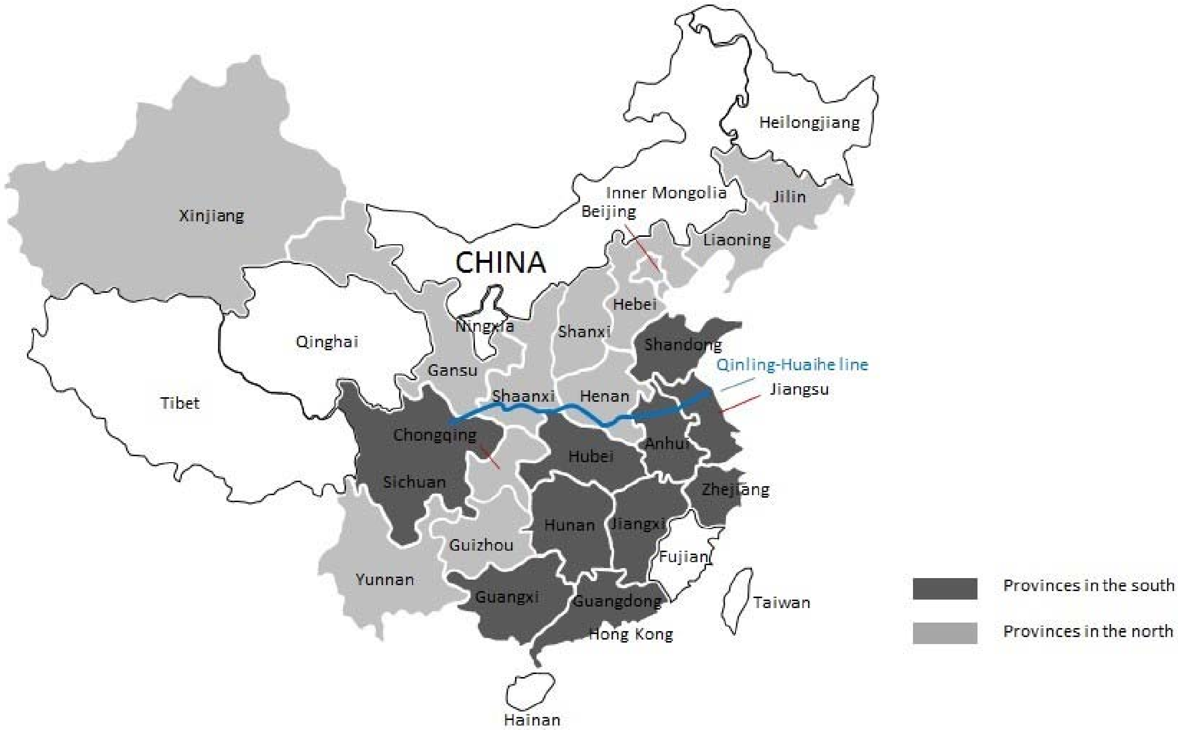


Figure 3.1: The surveyed provinces in China

two groups: Southern provinces and Northern provinces. As shown in Figure 3.1, darker shade denotes Southern provinces. The geographic border between Northern and Southern China is defined by the line of Huaihe River and Qinling mountains.[6] The South/North of China in this analysis includes the provinces located in the South/North of the line. There are four exceptions: Shandong, Chongqing, Yunnan and Guizhou provinces. We group Shandong province into the South, though it locates in the North of the line, because lineage culture in Shandong province traditionally is strong (Wang, 2007). Likewise, Chongqing, Yunnan and Guizhou are grouped as the north, despite that their geographic locations are in the South. This is because these provinces have large minority populations. Unlike Han ethnicity, minorities usually do not have

[6] The Qinling-Huaihe line is an important agro-climatic demarcation line in China. On the two sides of this line, the climate, flora and fauna, and agricultural products are very different.

the lineage culture. In total, ten provinces are defined as in the South and twelve provinces are defined as in the North.[7]

In our sample, individual-level questions were answered by heads of households. For each village, the village-level questions were answered by a village representative who was familiar with geographic, demographic and economic characteristics of the village. A village representative could be the party branch secretary, the head of village committee, or the village accountant, whoever was available during the survey. Table 3.1 presents summary statistics for all the variables that are at the core of this analysis. In the following subsection, we describe our variables of interest.

3.2.1 Three types of villages

Two village-level questions in the survey enable us to categorize the 961 villages into three types in terms of fragmentation. The two questions are as follows：

- Q1. "Is the percentage of households belonging to the largest lineage more than 50%?"
- Q2. "Is the percentage of households belonging to the top five largest lineages more than 50%?"

In the sample, villages that answered "yes" to Q1 are defined as type 1 villages. Villages that answered "no" to Q1 and "yes" to Q2 are type 2 villages. Villages who answered "no" to both Q1 and Q2 are type 3 villages. Thus type 1 villages are the most homogenous villages as the majority of households in a type 1 village are from the largest lineage. Type 3 villages are the

[7] The provinces in "the South" include Jiangsu, Zhejiang, Anhui, Jiangxi, Shandong, Hubei, Hunan, Guangdong, Guangxi, and Sichuan. The provinces in "the North" include Beijing, Hebei, Shanxi, Liaoning, Jilin, Henan, Shannxi, Gansu, Xinjiang, Chongqing, Guizhou, and Yunnan.

Table 3.1: Summary statistics

Dependent Variables	Definition	Mean (Std. Dev.)		
		All	South	North
Borrow	=0 if none/few or sometimes; =1 if often or very frequently	0.383 (0.486)	0.375 (0.484)	0.391 (0.488)
Help	=0, 1, 2, or 3. The smallest value of help is 0 which means the respondent answers “non/few” or “sometimes” to all the three categories (help farming, help house building and help taking care of others); the largest value of help is 3 which means the respondent answers “often” or “very frequently” to all the three categories	1.645 (1.026)	1.581 (1.050)	1.717 (0.994)
Fulfill	=1 if the respondent physically fulfills the assigned collective working requirement without paying any penalty; =(actually completed unpaid working days/ required unpaid working days) if not	0.891 (0.285)	0.855 (0.330)	0.932 (0.219)
Share	the share of village budget spent on education, medical system, and other public goods	0.149 (1.767)	0.143 (0.164)	0.156 (0.192)
Sgrowth	Change of *share* from 1998 to 2002	-0.005 (0.166)	-0.007 (0.159)	-0.002 (0.174)

(Table 3.1 continued)

Village Characteristics	Definition	All	South	North
Type1	=1 if one largest lineage dominates in the village; =0 if not	0.299 (0.458)	0.357 (0.480)	0.238 (0.427)
Type2	=1 if five largest lineages dominate in the village; =0 if not	0.368 (0.483)	0.388 (0.488)	0.348 (0.477)
Mountain	=1 if in mountainous area; =0 if in hilly or plain area	0.505 (0.500)	0.536 (0.499)	0.473 (0.500)
Suburb	=1 if is the suburb of a city/middle city; =0 if not	0.080 (0.272)	0.076 (0.265)	0.084 (0.278)
Distance to stations	Distance from the closest transportation terminals in Kilometer	5.449 (8.236)	4.05 (5.255)	6.855 (10.219)
Distance to county	Distance from the nearest county in Kilometer	24.128 (21.054)	22.346 (17.328)	26.185 (24.135)
Poverty	=1 if the village is in a county designated as province or national level poverty county or in a town designated as the province level poverty town; =0 if not	0.317 (0.466)	0.283 (0.451)	0.352 (0.478)
Poptotal	Population of the village in 2002	1811.825 (1185.602)	1928.862 (1307.258)	1691.578 (1033.645)
Vincome	Net income per capita of 2002 of the village in Yuan	2453.4 (1497.472)	2860.242 (1711.333)	2035.351 (1094.09)
Plantarea	Total planting area of 2002 of the village in Mu	3553.328 (2928.316)	3011.117 (2503.255)	4109.295 (3216.957)
Canal98	=1 if using canal as a major irrigating method in 1998; =0 if using well	0.692 (0.462)	0.850 (0.358)	0.515 (0.500)

(Table 3.1 continued)

Village Characteristics	Definition	All	South	North
Village head				
Years in office	= the number of years the village head being in office	5.197 (0.499)	5.356 (4.958)	5.032 (5.011)
Age	=1 if the age of the village head is 29 or below; =2 if between30-34; =3 if between 35-39; =4 if between 40-44; =5 if between 45-49; =6 if between 50-54; =7 if 55 or above	4.518 (1.449)	4.551 (1.419)	4.484 (1.480)
Education	=1 if the educational level of the village head is primary school or less; =2 if junior middle school; =3 if senior middle school; =4 if technical secondary school; = 5 if college or above	2.463 (0.914)	2.512 (0.887)	2.411 (0.939)
Management Experience	=1 if the village head has the experience of enterprise management; =0 if otherwise	0.381 (0.486)	0.374 (0.484)	0.389 (0.488)
Non-agri business Experience	=1 if the village head has the experience of operating nonagricultural family businesses; =0 if otherwise	0.409 (0.492)	0.422 (0.494)	0.395 (0.489)

(Table 3.1 continued)

Village Characteristics	Definition	All	South	North
The Party Secretary				
Years in office	= the number of years the party secretary be in office	7.123 (6.787)	7 (6.320)	7.25 (2.239)
Age	=1 if the age of the party secretary is 29 or below; =2 if between30-34; =3 if between 35-39; =4 if between 40-44; =5 if between 45-49; =6 if between 50-54; =7 if 55 or above	4.860 (1.409)	4.807 (1.417)	4.913 (1.400)
Education	=1 if the educational level of the party secretary is primary school or less; =2 if junior middle school; =3 if senior middle school; =4 if technical secondary school; = 5 if college or above	2.583 (1.004)	2.632 (0.948)	2.532 (1.057)
Management Experience	=1 if the party secretary has the experience of enterprise management; =0 if otherwise	0.387 (0.487)	0.386 (0.487)	0.388 (0.488)
Non-agri business Experience	=1 if the party secretary has the experience of operating nonagricultural family businesses; =0 if otherwise	0.459 (0.496)	0.466 (0.499)	0.451 (0.498)

(Table 3.1 continued)

Personal Characteristics	Definition	All	South	North
Female	=1 if female ; =0 if male	0.255	0.237	0.275
		(0.436)	(0.425)	(0.446)
Age	Age of the respondent	45.354	45.802	44.851
		(10.692)	(10.610)	(10.762)
Marriage	=1 if married; =0 otherwise	0.951	0.951	0.952
		(0.215)	(0.216)	(0.214)
Cadre[a]	=1 if is a cadre; =0 if not	0.160	0.168	0.150
		(0.366)	(0.374)	(0.357)
Education[b]	Years of schooling	7.010	7.057	6.958
		(2.716)	(2.672)	(2.763)
Hhincome	Total net household income of 2002 in Yuan	10704.25	12308.3	8903.212
		(8594.038)	(10037.08)	(6128.024)
Hhsize	Total number of residents living in the household for 6 months or more	4.100	4.062	4.122
		(1.306)	(1.267)	(1.347)
Surname	=1 if the respondent belongs to the largest lineage; =0 if not	0.412	0.446	0.375
		(0.492)	(0.497)	(0.484)
Past disaster	Number of natural disasters suffered in the past five years (1998-2002)=1 if none; =2 if one; =3 if two; =4 if three or more	1.990	1.752	2.259
		(1.108)	(1.000)	(1.160)

Note:

a. In 2002, 1 USD= 8.2770 Yuan , according to *China Statistical Yearbook 2011*

b. Cadre means administrators in China. In both Russia's and China's revolutionary eras, the word refers to a group of leaders active in promoting the revolution of the communist party. It no longer has any revolutionary implications in today's China.

c. If there is a missing value, replace it with a value estimated from education level. For example, if the education level is college or above, I use 17 years of education; if professional school, I use 14 years of education; if middle level professional, technical or vocational school, I use 12 years; if senior middle school, use 12; if junior middle school, use 9; if 4 or more years of elementary school, use 5; if 1-3 years of elementary school, use 2; if illiterate or semi-illiterate, use 0 most heterogeneous villages as each village consists of a number of small lineages. Type 2 is a medium type. Figure 3.2 provides a visual comparison of lineage composition in the three types of villages. In our sample, 30 percent of the villages are type 1 and 37 percent are type 2.

There is another household-level question that also provides lineage information:

- "Does your family belong to the largest lineage in the village?"

We use the surveyed households to calculate the percentage of the largest-lineage households in each village. Our calculations show that, on average, type 1 villages have 71%, type 2 villages have 37% and type 3 villages have 15% respectively of households belonging to the largest local lineage. Thus type 3 villages are the most heterogeneous villages while type 1 villages are almost twice as homogenous as type 2 villages.

We use type 3 villages as the reference group and examine whether people in type 1 and 2 villages are more cooperative. Therefore, we define two binary variables: *TYPE1* and *TYPE2*. *TYPE1* is 1 if the respondent belongs to a type 1 village and 0 otherwise. Similarly *TYPE2* is 1 if the respondent belongs to a type 2 village and 0 otherwise.

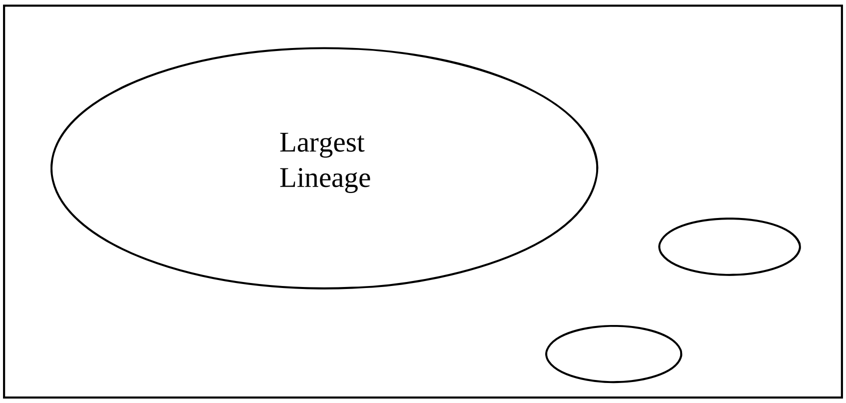
3.2.2 Intra-lineage reciprocity variables

Next we describe the construction of the variables capturing intra-lineage interaction. We use the following question from CHIPS 2002:

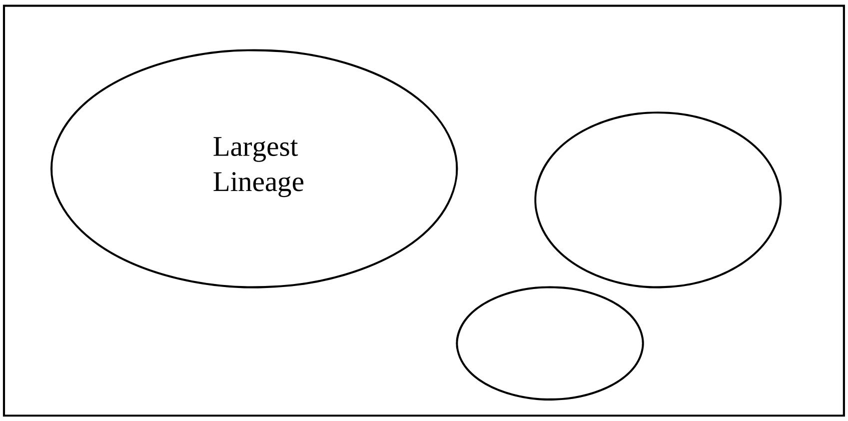
- "How often do you offer the following types of mutual help to your relatives and neighbors?"

The types of mutual help include: (*i*) borrowing and lending money; (*ii*) helping in farming during the busy season; (*iii*) helping in house building; and (*iv*) taking care of the elderly, the sick, and babies. Since extended families of one lineage cluster together in the same area of a

Type 1: Most dominating



Type 2: Moderately dominating



Type 3: Least dominating

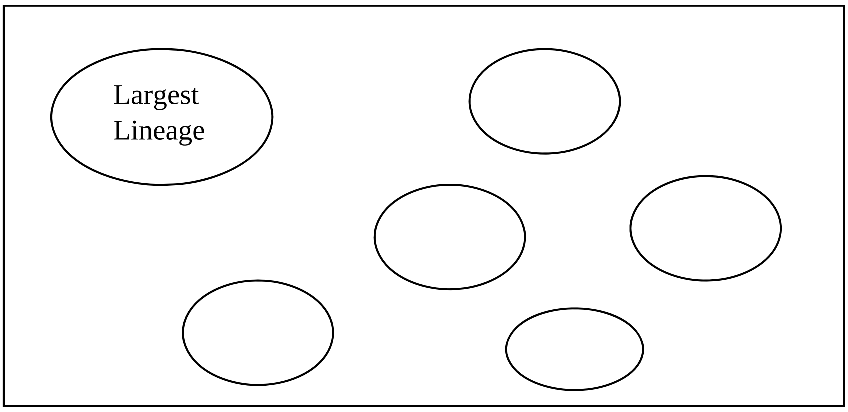


Figure 3.2: Three types of villages by lineage compositions

village, we assume that the respondent's relatives and neighbors are mostly from his/her own lineage. Therefore, we use the answers to this question to proximate intra-lineage reciprocity. Among the four types of help listed above, the first one reflects monetary reciprocity while the other three capture the non-pecuniary favors, especially those favors that require an investment of time. Therefore, we define two dependent variables: *borrow* and *help* to separate monetary from non-monetary reciprocities. The binary variable *borrow* takes the value 1 if the respondent answered that borrowing or lending money happened often or very frequently, and zero if the respondent answered that this mutual help happened rarely or never. With regard to the other three types of mutual help, (*ii*), (*iii*) and (*iv*), we first construct a binary variable for each type in the same way as we construct *borrow*. Then we define *help* by adding up the three indicators. Hence, *help* takes values 0, 1, 2 or 3 where a larger number implies more mutual help regardless of the type of non-monetary help. For example, 0 means that all the three binary help variables are 0s: the respondent answered mutual help in all the three types (*ii*), (*iii*) and (*iv*) happened rarely or never; 3, the greatest possible number of *help,* indicates all the three binary variables take the value of 1: the respondent answered mutual help in type (*ii*), (*iii*) and (*iv*) happened often or very frequently.

3.2.3 Inter-lineage cooperation variables

A commonly used measurement of inter-group cooperation in literature is the provision of public goods (Alesina et al., 1999; Banerjee et al., 2005). Following the literature, we examine villagers' physical contribution to villages' public goods and the share of villages' budgets spent on public goods. Village public goods, such as irrigation facilities, roads, and schools, are jointly consumed by all villagers, regardless of their lineage membership. Therefore, people's willingness to invest in public goods reflects the inter-lineage reciprocities in any given village.

Before 2002, all villagers (between the ages of 18 and 65) in China were required by law to provide unpaid labor to build local public goods, such as irrigation systems, dams, roads, school buildings.[8] The number of regulated days of unpaid work varied from place to place, but was usually around 7-21 days per year. If villagers could not physically participate in the unpaid work, they were charged fines for each day they missed. The villages could use the collected fines to hire other people to replace the missing laborers. Since 2002, China has gradually reduced and eventually waived this unpaid-labor duty. However, this reform did not take effect at the same pace across the nation. When CHIPS 2002 was conducted, this reform had still not been implemented in 140 out of the 961 surveyed villages which provide us with data on the fulfillment of the unpaid-labor requirement. In these 140 villages, each surveyed household reported the number of days that they were required to work for free and the number of days they actually completed in 2002. Based on this information, we construct the variable *fulfill,* which is the ratio of the number of completed days to the number of required days, to measure the households' physical contribution to village public goods.

To further investigate monetary contribution to public goods, we construct another two village level variables for this: (*i*) *share*, which measures the share of village budget spent on education, the medical system, and other commonweal expenditure (i.e., expenditure on environment protection and public safety); (*ii*) *sgrowth*, which measures the change of *share* from 1998 to 2002.

[8] According to the *Regulations on Peasants' Fees and Services* (1992) announced by the State Council of the People's Republic of China. Before the tax-for-fee reform around 2002, households were required to supply labor for free to local authorities mostly for the construction of local infrastructure. The number of regulated days varied with local needs. Local authorities were responsible for enforcing this regulation. The unpaid labor requirement should take place during off-season in farming.

Alesina et al. (1999) find a negative relationship between public good provision in U.S. cities and ethnic diversity from the median voter theorem. However, given the nature of political institutions in China, it is not entirely clear how the provision of public goods was determined in rural villages. In fact, due to a series of political and financial reforms in rural China around 1990s, villagers had been granted increasing power to determine who the village leaders were, and how to spend the village budget (Zhang et al., 2004). Before 1988, villages were governed by village Party branches, whose members were appointed by county-level Party committee. In 1988, the *Organic Law of Village Committees* was implemented. This law let villagers elect village councils, who share the administrative power with the Party branches. The village council members are chosen from villagers. Though the specific way of sharing the power between village councils and the Party branches varies from village to village, the right to vote for their own leaders increases villagers' awareness of participating in public affairs and their desire to communicate their demands (Coniff, 2004). In 1998, a revised version of the *Organic Law of Village Committees* was passed, aiming at further improving the democracy of rural governance. The new law clarifies the regulations on how the elections should be held (for example, open primaries should be hold to nominate the candidates). In her testimony to the U.S. congress, in 2002, Anne Thurston stated that "*there is some evidence, though we certainly need more research, that governance in such villages has improved, finances have become more transparent, and corruption has declined.*" Despite the fact that rural China is far from being fully democratic, there are still some channels though which villagers can participate in public affairs. This is all we need for our analysis. In addition, as rural China is making progress in switching to democratic electoral process, one would expect that the provision of public good should become increasingly aligned with the median voter's opinion. If the median voters in homogenous and heterogeneous villages have different opinions in the supply of public good, one should observe that the provision of public goods would have distinctive paths of growth over time across different types of villages. Therefore, we investigate *sgrowth*, the change in the share of public goods in village budget from 1998 to 2000.

3.2.4 Other control variables

The survey process collected detailed data about individual and household information as reported in the last panel of Table 3.1. We now discuss some of these variables which are specific to our data set. *CADRE* is a binary variable indicating whether or not the respondent is a village cadre and 15.8 percent of respondents were village cadres.[9] *EDUCATION* measures the respondent's years of schooling. *SURNAME* is a binary variable which takes value 1 if the respondent belongs to the largest lineage in the village. Following the study by Alesina and La Ferrara (2002), we construct a *PAST DISASTER* variable which takes value of 1, 2, 3 or 4. A larger number indicates that the respondent suffered more natural disasters in the last five years (1998 – 2002). Alesina and La Ferrara (2002) includes a similar indicator, "recent traumas" in their model, which is equal to 1 if the respondent suffered a negative experience in the past year such as divorce, diseases, accidents, financial misfortune. Their study shows "recent traumas" has a negative impact on trust. Due to data limitations, we do not have all the details on villagers' past experiences. Our analysis uses *PAST DISASTER* instead to check if the number of disasters suffered by the respondent affects his/her cooperative behavior.

Village characteristics are reported in the second panel of Table 3.1. The mean of *MOUNTAIN* is 0.505, implying that approximately half of the villages are located in the

[9] Cadre means administrators in China. In both Russia's and China's revolutionary eras, the word refers to a group of leaders active in promoting the revolution of the communist party. It no longer has any revolutionary implications in today's China (Pan, 2012).

mountainous area. Since the village location plays an important role in the prevalence of lineage culture (Freedman, 1965), we also include other geographic controls, such as whether the village is the suburb of a city, the distance to the nearest transportation terminals and the distance to the nearest county. Irrigation is the most important form of long-term inter-lineage cooperation in rural China (Freedman, 1965). Therefore we control for the variable, denoted as *CANAL98*, which measures whether the village used the canal as the major irrigating method in 1998. In rural China, the village is led by a village head (the chairman of the village committee) and a party secretary. The village economy depends heavily on village leadership (Oi, 1999). Thus, the characteristics of village leaders are controlled for our analysis. A total of five measures are used: the number of years the village leader have been in office, age of the village leader, education level of the village leader, enterprise management experience and the experience of operating non-agriculture business family business.

The outcome variables are listed in the first panel of Table 3.1, followed by village characteristics in the second panel and individual characteristics in the last panel.

3.3 The models

In this section we focus on describing the models used in the analysis. We use individual level models to examine intra-lineage cooperative behavior and villagers' physical contribution to public goods. These models are presented in subsection 3.1. When we study the impact of lineage-based fragmentation on the share of village budget spent on public goods, we use village level models demonstrated in subsection 3.2.

3.3.1 Intra-lineage

Our basic model for intra-lineage relationships is

$$outcome_{ijp} = \beta_1 TYPE1_{jp} + \beta_2 TYPE2_{jp} + Y_{jp}\delta + X_{ijp}\gamma + \alpha_p + \varepsilon_{ijp} \quad (3.1)$$

where the subscripts indicate individual i in village j of province p. The outcome variable is *help* or *borrow*. *TYPE1* and *TYPE2* are the village-type indicators. Y_{jp} is a vector of other village characteristics. It includes net income per capita, and its squared form so that we are able to test whether the effect of income per capita on cooperative behavior is stronger or weaker as villages get wealthier. Y_{jp} also includes an indicator for mountainous area, a suburb indicator, distance to the closest transportation station, distance to the closest transportation terminals, poverty indicator, total population, total planting area, characteristics of village leaders and a binary variable indicating whether the village uses a canal as the major irrigating method in 1998. X_{ijp} is a vector of individual characteristics. It includes age and a quadratic form of age, so that we can examine whether the effect of age on cooperative behavior is stronger or weaker as people get older. It also includes a sex indicator, years of schooling, marital status, a cadre indicator, an indicator of marital status, household income, family size, a binary variable indicating whether the individual belongs to the largest lineage in the village and the number of natural disasters he/she suffered in the past five years (1998-2002). α_p is a vector of province fixed effects, ruling out systematic differences between provinces.

Since *borrow* is a binary variable, we use a probit model to estimate regression coefficients. *Help* is a discrete ordinal variable. More specifically, we classify the frequency of non-monetary help into 4 categories, with 3 thresholds. Therefore we use an ordered probit model when *help* is the outcome variable.

We expect intra-lineage cooperation is more frequent in homogenous villages than in heterogeneous villages. The identification strategy in this paper arises from the exogenous and predetermined fragmentation. We use model (3.1) to examine whether intra-lineage cooperation is more frequent in homogenous villages than in heterogeneous villages. Since the lineage culture traditionally is more predominant in the South than in the North of China, to further refine the identification, we separate the entire sample into two subsamples – Southern provinces and Northern provinces. We apply model (3.1) to each subsample and investigate whether the difference is stronger in the South than in the North. If the answer is affirmative, this indicates that the lineage-based heterogeneity affects people's cooperative behavior. However, to achieve the necessary identification, the model needs an assumption, i.e., the villages of the three types are not different in unobserved aspects that can directly affect people's cooperative behavior. Otherwise, our estimates may be biased. Hence, we use the following D-in-D model to identify the impact of fragmentation under a relaxed assumption that unobserved differences between types are allowed:

$$outcome_{ijp} = \beta_3 TYPE1_{jp} + \beta_4 TYPE2_{jp} + \beta_5 TYPE1_{jp} * SOUTH_{jp} + \beta_6 TYPE2_{jp} * SOUTH_{jp}$$
$$+Y_{jp}\delta + X_{ijp}\gamma + \alpha_p + \varepsilon_{ijp} \quad (3.2)$$

where *SOUTH* is a dummy variable which is equal to 1 if the respondent is from the South of China. β_3 and β_4 measure the impact of lineage composition in the North. $\beta_3 + \beta_5$ and $\beta_4 + \beta_6$ measure the impact of lineage composition in the South. β_5 and β_6 reflect whether the impact of lineage composition are different between the South and the North. The lineage culture is more prevalent in Southern China than in Northern China. If homogeneity promotes intra-lineage cooperation, the impact should be stronger in the South. Then β_5 and β_6 will be positive. The assumption of the model is that the differences in other characteristics across of the three-type villages, apart from lineage culture, do not change from the South to the North. We empirically test whether this assumption holds in next section.

3.3.2 Inter-lineage

The measurements of inter-lineage cooperation include *fulfill,* which is an individual-level outcome, and *share* and *sgrowth*, which are two village-level outcomes. The models for the *fulfill* variable are the same as in the models (3.1) and (3.2) except that the fine charged for each missed day of unpaid work is also included in Y_{jp}, besides all other village characteristics. The models for the *share* variable are as follows:

$$outcome_{jp} = \beta_7 TYPE1_{jp} + \beta_8 TYPE2_{jp} + Y_{jp}\delta + \alpha_p + \varepsilon_{jp} \qquad (3.3)$$

$$outcome_{jp} = \beta_9 TYPE1_{jp} + \beta_{10} TYPE2_{jp} + \beta_{11} TYPE1_{jp} * SOUTH_{\ jp}$$
$$+ \beta_{12} TYPE2_{jp} * SOUTH_{jp} + Y_{jp}\delta + \alpha_p + \varepsilon_{jp}. \qquad (3.4)$$

The literature (Alesina et al., 1999; Banerjee, Iyer and Somanathan, 2005; Bandiera et al., 2005 and Vigdor, 2004) finds that ethnic diversity discourages people from contributing to public goods. Thus, we hypothesize that $\beta_7,\ \beta_8,\ \beta_{11}\ and\ \beta_{12} > 0$.

Note that when *sgrowth* is the outcome variable, we aim at examining whether the cooperation gap between a homogeneous village and a heterogeneous village is larger in 2002 because of the political reform implemented since 1998. So the general model is:

$$share_{jpt} = \beta_{13} TYPE1_{jp} + \beta_{14} TYPE2_{jp} + \beta_{15} T + \beta_{16} TYPE1_{jp} * T +$$
$$\beta_{17} TYPE2_{jp} * T + Y_{jpt}\delta + \varepsilon_{jpt} \qquad (3.5)$$

where T is an year indicator. $T = 1$ if the year is 2002; $T = 0$ if the year is 1998.

In 1998, the model can be rewritten as:

$$share_{jp,1998} = \beta_{13} TYPE1_{jp} + \beta_{14} TYPE2_{jp} + Y_{jp,1998}\delta + \varepsilon_{jp,1998}. \qquad (3.6)$$

In 2002, rewrite (3.5) as

$$share_{jp,2002} = \beta_{13} TYPE1_{jp} + \beta_{14} TYPE2_{jp} + \beta_{15} + \beta_{16} TYPE1_{jp} +$$
$$\beta_{17} TYPE2_{jp} + Y_{jp,2002}\delta + \varepsilon_{jp,2002}. \qquad (3.7)$$

Thus, we can obtain the following model for the variable *sgrowth* from (3.7) minus (3.6):

$$sgrowth = \beta_{15} + \beta_{16}TYPE1_{jp} + \beta_{17}TYPE2_{jp} + \Delta Y_{jp}\delta + \Delta\varepsilon_{jp}. \quad (3.8)$$

where ΔY_{jp} indicates the change of village characteristics from 1998 to 2002.

Similarly, we derive the following D-in-D model for *sgrowth*:

$$sgrowth = \beta_{18} + \beta_{19}TYPE1_{jp} + \beta_{20}TYPE2_{jp} + \beta_{21}TYPE1_{jp} * SOUTH_{\ jp}$$

$$+ \beta_{22}TYPE2_{jp} * SOUTH_{jp} + \Delta Y_{jp}\delta + \Delta\varepsilon_{jp}. \quad (3.9)$$

Since in a more democratic environment, people have more capacity to participate in collective decisions. Then we can expect public-good provision should increase more in a more homogenous village because people's decision of cooperation could be better realized in 2002. In other words, $\beta_{16},\ \beta_{17},\ \beta_{21}\ and\ \beta_{22} > 0$.

3.4 Results

This section discusses results. We first present evidence that our assumption of D-in-D models is justified. Then we explore how lineage-based fragmentation affects intra-lineage cooperation and inter-lineage cooperation respectively. The regressions control for a large set of individual characteristics and village characteristics as presented in Table 3.1. When *borrow* is the outcome variable, we use a probit model; when *help* is the outcome variable we use an ordered probit model; when the outcome variable is *fulfill*, *share* or *sgrowth*, the results are based on OLS estimates. In addition, we cluster the standard errors by county.

3.4.1 The D-in-D model

As discussed in previous section, we apply the D-in-D model to investigate whether our results are robust under a relaxed assumption – we allow for unobserved differences across the three types of villages that can directly affect people's cooperative behavior. However, we also assume that the differences across the three types of villages, other than lineage culture, do not change from the South to the North. Now we examine the validity of this assumption for the D-in-D model. We use the individual and village characteristics as the left-hand side variables and *TYPE1*, *TYPE2*, and the interaction terms $TYPE1 * SOUTH$ and $TYPE2 * SOUTH$ as the right-hand side variables. We test whether those characteristics are different across the three village types (evaluated by the coefficients of *TYPE1* and *TYPE2*), and more importantly, whether those differences, if there are any, change from the South to the North (evaluated by the coefficients of $TYPE1 * SOUTH$ and $TYPE2 * SOUTH$).

Table 3.2 presents the results. Notice that the coefficients of *TYPE1*, *TYPE2*, $TYPE1 * SOUTH$ and $TYPE2 * SOUTH$ for the dependent variable "*SURNAME*" are all significantly positive. "*SURNAME*" is a dummy variable which is 1 if the individual is from the largest lineage in the village and zero otherwise. The four positive coefficients indicate that people in type 1 and type 2 villages are more likely than type 3 villages to be from the largest local lineage, and this situation is more likely to be the case in the South. This illustrates that type 1 and type 2 villages are more homogenous than type 3 villages and the difference in lineage-based fragmentation across types of villages is greater in the South than in the North. However, we do not see other variables having the same pattern as "Surname". For the other variables, though the coefficients of *TYPE1* and *TYPE2* can be statistically different from zero, the coefficients of $TYPE1 * SOUTH$ and $TYPE2 * SOUTH$ are not. For example, $TYPE1 * SOUTH$ ($TYPE2 * SOUTH$) does not have a significant impact on the respondent's education level, implying that there is no evidence that the differences in education level between people in type 1 (type 2) villages and type3 villages vary from the South to the North. Thus, the results in Table 3.2 indicate that while the three types of villages can be different in aspects other than the lineage-based fragmentation, there is no evidence that these other differences change from the South to the North.

3.4.2 Intra-lineage relationship

Using models (3.1) and (3.2), we examine the impact of lineage-based fragmentation on intra-lineage relationships. Our goal is to investigate whether *borrow* and *help* are more likely to happen in type 1 and type 2 villages than in type 3 villages. Regression results are reported in Table 3.3 (*borrow*) and Table 3.4 (*help*). Table 3.3 presents marginal probit coefficients calculated at the means while Table 3.4 presents ordered probit coefficients.

In Table 3.3, column 1 demonstrates that frequent monetary help among lineage members is more likely to happen in types 1 and 2 villages than in type 3 villages by 6.4 and 5.1 percentage points respectively. The estimated marginal coefficient of *TYPE1* is statistically significant at the 5 percent level of significance while the estimated marginal coefficient of *TYPE2* is statistically significant at the 10 percent level. Then we split the entire sample into two subsamples: the South and the North. In the South of China, as presented in column 2, the possibility of frequent borrowing and lending within a lineage is 11.6 percentage points higher in type 1 villages and 9.3 percentage points higher in type 2 villages than in type 3 villages. The point estimates of *TYPE1* and *TYPE2* in the South sample are both statistically significant and greater in magnitude than those in the entire sample. By contrast, there is no evidence that lineage-based fragmentation has an impact on *borrow* in the North (column 3). To further test whether the impact of *TYPE1* and *TYPE2* are different between the south and the North, we use the D-in-D model, whose results demonstrate that both the coefficients of *TYPE1* and *TYPE2* in the South is significantly different from that in the North at 10 percent level of significance. This confirms the causal impact of lineage-based fragmentation on intra-lineage monetary help.

Table 3.2: Examining the assumption of the D-in-D model
Dependent variables: individual characteristics and village characteristics

	Independent Variables								
Dep. Var.	Type1		Type2		Type1*South		Type2*South		Obs
Individual Characteristics									
Age	1.681**	[0.714]	1.131*	[0.623]	0.027	[1.001]	-0.514	[0.869]	6,840
Marriage	-0.001	[0.011]	-0.015	[0.011]	-0.008	[0.016]	-0.004	[0.016]	6,840
Cadre	0.016	[0.029]	0.009	[0.023]	0.014	[0.037]	0.012	[0.033]	6,816
Education	0.172	[0.189]	-0.148	[0.177]	-0.077	[0.241]	0.064	[0.219]	6,840
Hhincome	0.04	[0.050]	-0.027	[0.058]	-0.074	[0.083]	0.001	[0.076]	6,836
Hhsize	0.13	[0.099]	-0.003	[0.087]	-0.045	[0.136]	0.017	[0.114]	6,840
Surname	0.406***	[0.049]	0.079*	[0.044]	0.145**	[0.060]	0.139***	[0.052]	6,839
Past disaster	-0.026	[0.156]	0.099	[0.129]	0.126	[0.190]	0.086	[0.165]	6,774

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the county level. All specifications include province dummies.

(Table 3.2 continued)

Dep. Var.	Independent variables								
	Type1		Type2		Type1*South		Type2*South		Obs
Village Characteristics									
Suburb	-0.034	[0.026]	-0.019	[0.027]	-0.013	[0.047]	-0.008	[0.040]	733
Mountain	-0.056	[0.076]	0.039	[0.067]	-0.041	[0.106]	0.005	[0.086]	733
Distance to Transportation	-0.188	[1.480]	2.551*	[1.535]	-0.216	[1.642]	-2.559	[1.698]	718
Distance to County	-4.075	[4.404]	-1.821	[3.748]	6.217	[5.328]	2.138	[4.974]	729
Poverty	-0.036	[0.088]	0.064	[0.074]	0.174	[0.123]	0.054	[0.104]	733
Poptotal	-0.121	[0.122]	-0.1	[0.089]	-0.07	[0.162]	0.098	[0.115]	733
Vincome	-0.067	[0.068]	-0.132**	[0.066]	-0.016	[0.103]	0.116	[0.090]	727
Plantarea	-0.188	[0.151]	0.053	[0.105]	0.14	[0.223]	0.048	[0.176]	729
Canal98	-0.09	[0.058]	-0.067	[0.069]	0.014	[0.080]	0.086	[0.076]	694
Penalty	3.484	[2.766]	1.795	[2.368]	-4.276	[3.030]	-0.524	[2.713]	153

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the county level. All specifications include province dummies

(Table 3.2 continued)

Dep. Var.	Independent Variables Type1		Type2		Type1*South		Type2*South		Obs
Village Characteristics									
Village leaders (the party secretary)									
Years in office	0.921	[1.030]	-0.806	[0.806]	-2.204	[1.376]	1.406	[1.247]	733
age	0.115	[0.223]	0.001	[0.219]	-0.187	[0.306]	0.231	[0.288]	733
Education	-0.213	[0.213]	-0.147	[0.158]	0.029	[0.246]	0.119	[0.201]	733
Management experience	-0.013	[0.093]	0.072	[0.089]	-0.139	[0.108]	-0.168	[0.108]	733
Non-agri business experience	0.002	[0.094]	-0.025	[0.079]	-0.148	[0.118]	-0.051	[0.101]	733
Village leaders (village head)									
Years in office	-0.294	[0.644]	-0.852	[0.600]	-0.179	[0.948]	0.331	[0.917]	726
age	0.711***	[0.264]	0.336	[0.207]	-0.822**	[0.346]	-0.519*	[0.285]	729
Education	-0.305*	[0.160]	-0.357**	[0.147]	0.1	[0.193]	0.236	[0.184]	729
Management experience	-0.084	[0.071]	-0.074	[0.071]	-0.035	[0.092]	-0.037	[0.093]	729
Non-agri business experience	-0.029	[0.085]	-0.098	[0.068]	-0.016	[0.112]	0.143	[0.095]	729

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the county level. All specifications include province dummies

With regard to non-monetary help, we present ordered probit model coefficients in Table 3.4. But those do not measure marginal effect. We therefore compute the marginal impact of *TYPE1* and *TYPE2* on the probabilities that respondents lie in each of the four categories of *help* (from 0 to 3), and report these in table 3.5. In Table 3.4, column 1 reports that there is no evidence type 1 and type 2 villages are different from type 3 villages when we use the entire sample. However, when we restrict the sample to the South, the likelihood of frequent non-monetary help among lineage members is higher in type 1 and type 2 villages than in type 3 villages (Table 3.4 column 2). For example, the second panel of Table 3.5 presents that in the south, the possibility of frequent non-pecuniary help in all the three types (*help* is equal to 3) within a lineage is higher in types 1 and 2 villages than type 3 villages by, respectively, 7.3 and 6 percent. The coefficients of *TYPE1* and *TYPE2* even become negative when we restrict the sample to the North (Table 3.4 column 3). The D-in-D model indicates that the impact of *TYPE1* and *TYPE2* are both significantly different between the South and the North (Table 3.4 column 4).

The above evidence indicates that within-lineage reciprocity, both monetary and non-monetary, is more likely to happen in homogeneous villages than heterogeneous villages. We now explain why this is the case. Essentially an individual's lineage is like an organization to which he or she belongs and whose members tend to know each other quite well and have information about each other's social and economic activities. The importance of lineage organizations grows in the presence of asymmetric information or other market imperfections. Such an organization can enforce informal transactions, because it directs both punishment and reciprocity at not only individual but also members of his/her group (La Ferrara, 2003). Moreover, the enforcement can be better as the size of lineages increases (Pan, 2012). The reciprocity among lineage members can be regarded as a form of implicit contacts, in the sense that "I help you today because I expect you to help me tomorrow" (Posner, 1980). In a large lineage where everyone knows everyone else due to clustering over generations, a deviant may be denied future exchanges not only with the victim but also with a lot of other lineage members. In other words, the cost of defection potentially rises as the lineage size increases. Consequently, reciprocity will be more frequent in large lineages than small lineages. Notice that the size of a lineage in a homogeneous village, on average, is greater than a lineage in a heterogeneous village. This explains why we see more frequent reciprocity in homogenous villages than in heterogeneous villages.

To further support this hypothesis, we use the following model to test whether the cooperative behavior is more frequent in the largest local lineage than in the other smaller lineages in the same village:

$$outcome_{ij} = \beta_{23} SURNAME_{ij} + X_{ij}\gamma + \alpha_j + \varepsilon_{ij} \qquad (3.10)$$

where $outcome_{ij}$ is *help* or *borrow* for individual *i* in the village *j*. $SURNAME_{ij}$ is the binary variable which is 1 if the individual i belongs to the largest lineage in the village *j*. X_{ij} is the same vector of individual characteristics as in model (3.1). α_j is a vector of village fixed effects. The variable of interest for this model is $SURNAME_{ij}$. We expect that $\beta_{23} > 0$, particularly in the South. We restrict our sample to type 1 villages because the size difference between the largest lineage and other lineages is the greatest in type 1 villages.

The results are presented in Table 3.6. Columns 1 to 3 report probit coefficients for *borrow* while columns 4 to 6 report ordered probit coefficients for *help*. We use the entire sample in columns 1 and 4. Columns 2 and 5 restrict the sample to the respondents from the South of China. Columns 3 and 5 restrict the sample to those from the North of China. Columns 1 and 4 demonstrate that belonging to the largest lineage has a positive and significant effect on both monetary and non-monetary help. When we use the regional subsamples, the coefficient of *SURNAME* is statistically insignificant for the monetary help (column 2 and 3). However, for the non-monetary help, bearing the largest surname has a positive effect in the South and no effect in the North (column 5 and 6). The above results provide evidence that within-lineage reciprocity could increase with the sizes of lineages.

We examine within-lineage reciprocity in different ways. First we focus on monetary help and then on non-pecuniary but time-consuming help. Both of them are negatively associated with lineage-based fragmentation. This implies that villagers do not treat monetary help and non-monetary help differently when it comes to lineage obligations and enforcement. Table 3.3 and 4 also reports the coefficients of several interesting individual variables. First, *borrow* increases with age while *help* does not. One explanation is that lending is less likely to happen if there is asymmetric information about the riskiness of the borrowers. Older people are more experienced and may have more information about other lineage members. Hence, they are more likely to offer monetary help than younger people. Non-pecuniary help, on the other hand, although consumes time, it is less risky and asymmetric information plays a less important role. Second, the coefficients of *HHINCOME* are not significant for both monetary and non-monetary reciprocity within a lineage. This result implies that it does not seem to be the case that people with more monetary budgets also tend to help other lineage members more. Both *borrow* and *help* are positively associated with family size. Due to economies of scale, larger families may have more information about other lineage members and have more hands to offer help.

3.4.3 Inter-lineage relationship

Next, we use the models (3.1)-(3.4) to test whether homogeneous villages are more willing to contribute to public goods than heterogeneous villages. Table 3.7 presents the results. The outcome variables are *fulfill* for columns 1-4, *share* for columns 5-8 and *sgrowth* for columns 9-12. Recall that *fulfill* is an individual-level outcome variable measuring the respondent's physical contribution to public goods while the other two outcomes are village-level outcomes about the share of village budget spent on public goods. Columns 1-3 use model (3.1); column 4 uses model (3.2); columns 5-7 use model (3.3); column 8 uses model (3.4); columns 9-11 use model (3.8); and column 12 uses model (3.9).

Column 1 reports that type 1 villages completed more required unpaid labor days than type 3 villages by 10 percentage points. When we restrict the sample to the South, the effect of village types is stronger. Column 2 reports that type 1 and type 2 villages completed more of the unpaid-labor-day requirement by, respectively, 16.6 and 15.7 percentage points. In contrast, when the sample is restricted to the North (column 3), there is no evidence that the types of villages have an impact on the *fulfill* outcome. Column 4 uses the D-in-D model and shows that the impact of type 1 and type 2 are statistically significantly larger in the South than in the North.

Column 5 shows that type 1 and type 2 villages spend more of the village budget on public goods than type 3 villages by, respectively, 4.3 and 3.3 percentage points. The estimated coefficient of *TYPE2* is only marginally significant. When we use the sample of the South, the point estimations of *TYPE1* and *TYPE2* in column 6 are very close to those derived from the full sample. The coefficient of *TYPE1* is statistically significant but the coefficient of *TYPE2* is not. When we use the sample of the North, neither *TYPE1* nor *TYPE2* has an impact on village spending on public goods. Column 8 uses the D-in-D model. The results show that the impacts of *TYPE1* and *TYPE2* are statistically insignificantly different between the South and the North.

Table 3.3: Intra-lineage cooperation
Dependent variables: *borrow* (Probit)

	borrow			
	(1)all	(2)south	(3)north	(4)D-in-D
Type1	0.064**	0.116**	-0.003	0.000
	[0.032]	[0.044]	[0.048]	
				[0.047]
Type2	0.051*	0.093**	-0.003	-0.01
	[0.030]	[0.040]	[0.045]	
				[0.048]
Type1*South				0.104*
				[0.056]
Type2*South				0.102*
				[0.059]
Age	0.010**	0.006	0.014*	0.010**
	[0.005]	[0.005]	[0.008]	[0.005]
ln(hhincome)	-0.015	-0.008	-0.023	-0.015
	[0.012]	[0.016]	[0.017]	[0.012]
Hhsize	0.023***	0.023***	0.025***	0.023***
	[0.005]	[0.007]	[0.008]	[0.005]
Observations	8,193	4,483	3,710	8,193

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Coefficients are marginal probabilities calculated at the means from the probit models. Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the county level. All specifications include province dummies. All individual and village characteristics are included in the regression but not all of them are reported here. We also tried another specification that drops the three provinces: Chongqing, Yunnan and Guizhou. The results are similar.

Table 3.4: Intra-lineage cooperation
Dependent variables: *help* (ordered probit)

	help			
	(1)all	(2)south	(3)north	(4) D-in-D
Type1	0.108	0.237*	-0.096	-0.101
	[0.086]	[0.121]	[0.098]	[0.095]
Type2	0.047	0.197*	-0.182*	-0.174*
	[0.079]	[0.111]	[0.094]	[0.092]
Type1*South				0.346**
				[0.154]
Type2*South				0.377**
				[0.149]
Age	0.010	0.003	0.020	0.010
	[0.008]	[0.010]	[0.013]	[0.008]
ln(hhincome)	-0.010	-0.015	-0.010	-0.010
	[0.027]	[0.039]	[0.035]	[0.026]
Hhsize	0.034***	0.021	0.052***	0.035***
	[0.012]	[0.017]	[0.015]	[0.012]
Observations	8,193	4,483	3,710	8,193

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Coefficients are based on ordered probit estimates. Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the county level. All specifications include province dummies. All individual and village characteristics are included in the regression but not all of them are reported here. We also tried another specification that drops the three provinces: Chongqing, Yunnan and Guizhou. The results are similar.

Table 3.5: Magnitude of the effects: *help* and lineage-based fragmentation

	help			
	0	1	2	3
All				
Type1	-0.026	-0.016	0.009	0.033
Type2	-0.011	-0.007	0.004	0.014
South				
Type1	-0.060**	-0.034*	0.021**	0.073*
Type2	-0.050*	-0.028*	0.018*	0.060*
North				
Type1	0.021	0.016	-0.009	-0.028
Type2	0.040*	0.030**	-0.017*	-0.054*
D-in-D				
Type1*South	-0.075**	-0.057**	0.021***	0.112**
Type2*South	-0.082***	-0.062**	0.022***	0.122**

Note:
Figures in the table indicate the change in the probability of a respondent giving this value of *help* associated with the change of *TYPE1* (*TYPE2*) from 0 to 1. The estimation is based on the ordered probit model in Table 4. *Help* takes values 0, 1, 2 and 3 where a larger number implies more mutual help regardless of the types of help. For example, 0 means that all the three binary help variables are 0s: the respondent answered mutual help in all the three types (ii), (iii) and (iv) happened rarely or never; 3, the greatest possible number of help, indicates all the three binary variables take the value of 1: the respondent answered mutual help in type (ii), (iii) and (iv) happened often or very frequently.

Table 3.6: Intra-lineage cooperation in type 1 villages
Dependent variable: *borrow* and *help*

	Borrow			*help*		
	(1)All	(2)South	(3)North	(4)All	(5)South	(6)North
Surname	0.286*	0.248	0.24	0.360**	0.660***	0.118
	[0.167]	[0.244]	[0.235]	[0.169]	[0.225]	[0.265]
Age	0.045	0.004	0.123**	0.06	0.072	0.099
	[0.039]	[0.054]	[0.057]	[0.047]	[0.077]	[0.063]
Marriage	0.1	0.956**	-1.062**	-0.256	0.547	-1.263**
	[0.327]	[0.426]	[0.524]	[0.333]	[0.385]	[0.580]
Cadre	-0.097	-0.430*	0.365	0.288	-0.008	0.656**
	[0.173]	[0.232]	[0.317]	[0.184]	[0.239]	[0.282]
Education	0.000	0.009	-0.021	-0.04	-0.009	-0.065
	[0.029]	[0.039]	[0.047]	[0.027]	[0.034]	[0.044]
ln(hhincome)	0.07	0.211	-0.212	-0.007	0.068	-0.082
	[0.134]	[0.183]	[0.214]	[0.144]	[0.216]	[0.176]
Hhsize	0.082*	0.137**	0.063	0.038	-0.021	0.091
	[0.049]	[0.066]	[0.080]	[0.048]	[0.070]	[0.069]
Past disaster	0.03	0.072	-0.066	-0.023	-0.098	0.011
	[0.107]	[0.136]	[0.178]	[0.093]	[0.114]	[0.157]
Observations	904	542	362	993	604	389
Number of villages	277	168	109	262	159	103

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Coefficients are based on probit estimates (*borrow*) and ordered probit estimates (*help*). Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the village level. All specifications include village dummies. We also tried another specification that drops the three provinces: Chongqing, Yunnan and Guizhou. The results are similar.

The above evidence indicates that inter-lineage cooperative behavior is more likely to happen in homogeneous villages than in heterogeneous villages. Why does lineage composition matter? Alesina et al. (1999) claims ethnic groups can have different preferences even over a seemingly neutral public good. The same argument can be applied to lineage organizations. According to median voter's theorem, if there are many distinct preferences across groups, the chosen type of public goods is not preferred by a large fraction of the population (Alesina et al., 1999). In this case, individuals contribute fewer resources to public goods, because a large fraction of their resources are used to provide public goods shared with other groups. In a type 3 village, there are lots of small lineages. Villagers from a type 3 village decrease their contribution for the reason that most beneficiaries of the public goods do not belong to their own groups. This explains why homogeneous villages contribute more in public good provision both physically and monetarily than heterogeneous villages do.

Column 9-12 examine how lineage-based fragmentation affects the change of *share* from 1998 to 2002. As shown in column 9, the increase in the share of village budget that is spent on public goods is larger in type 1 villages than type 3 villages by 3.3 percentage points. When the sample is restricted to the South, the impact of village type is even larger. The share of public good spending increases more in type 1 and type 2 villages than in type 3 villages by, respectively 5.2 and 3.3 percentage points. When the sample is restricted to the North, type 1 and type 2 villages do not differ from type 3 villages in the growth of the share of public goods spending. Column 12 uses the D-in-D model. The point estimation of $TYPE1 * SOUTH$ and $TYPE2 * SOUTH$ are both positive. However, they are statistically insignificant. The reason perhaps is that the observation is at village level. Compared to Table 3.3 and Table 3.4, the number of observations in columns 5-12 is much smaller, which may make the estimation less precise.

The above evidence suggests that from 1998 to 2002, the share of village budget spent on public goods increases more in a homogeneous village than in a heterogeneous village. In 1998, the median voters in homogenous and heterogeneous villages may have different opinions about the supply of public goods. The median voters in homogeneous villages support more public good provision because lots of beneficiaries are from their own groups. As China is making progress in switching to a democratic electoral process, median voters play more and more important roles in village affairs. So over time, their opinions are better accepted when the village committees make decisions about how much to spend on public goods. Thus, in 2002, we observe there is a greater increase in the share of village budget spent on public goods in more homogeneous villages

3.5 Conclusion

This paper studies the relationship between lineage-based fragmentation and villagers' cooperative behavior. Rural China provides an excellent environment to conduct this study because China's central government arbitrarily grouped adjacent lineages into one administrative village during the communization movement. As a result, some villages are composed of one or a few large lineages while the others are composed of a number of small lineages. The exogenously determined lineage composition within a village presents a pseudo experiment of lineage-based heterogeneity.

Using data from CHIPS 2002, we define three types of villages: types 1-3, which go from the most homogenous to the most heterogeneous villages. We find that people of types 1 and 2

Table 3.7: Inter-lineage cooperation
Dependent variables: *fulfill, share* and *sgrowth*

VARIABLES	***fulfill***				***Share***				***sgrowth***			
	(1)all	(2)south	(3)north	(4)D-in-D	(5)all	(6)south	(7)north	(8) D-in-D	(9)all	(10)south	(11)north	(12)D-in-D
Type1	0.104**	0.166***	0.008	-0.077	0.043**	0.043**	0.06	0.060*	0.033*	0.052**	0.015	0.001
	[0.051]	[0.059]	[0.067]	[0.073]	[0.017]	[0.021]	[0.037]	[0.032]	[0.019]	[0.024]	[0.037]	[0.033]
Type2	0.064	0.157***	0.061	-0.048	0.033*	0.031	0.044	0.047	0.022	0.033*	0.004	-0.001
	[0.051]	[0.054]	[0.071]	[0.088]	[0.018]	[0.020]	[0.039]	[0.035]	[0.017]	[0.017]	[0.036]	[0.035]
Type1*South				0.275***				-0.027				0.049
				[0.098]				[0.036]				[0.039]
Type2*South				0.202*				-0.022				0.036
				[0.105]				[0.040]				[0.039]
Observations	1,453	947	506	1,453	685	394	291	685	671	389	282	671
R-squared	0.348	0.437	0.486	0.359	0.165	0.132	0.256	0.166	0.062	0.075	0.113	0.064

Note:
*** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1. Robust standard errors are in brackets, corrected for heteroskedasticity and clustering of the residuals at the county level. All specifications include province dummies. Individual and village characteristics are not reported. We also tried another specification that drops the three provinces: Chongqing, Yunnan and Guizhou. The results are similar.

villages are more likely than those of type 3 to have both within-lineage and across-lineage cooperation. In terms of within-lineage cooperation, we find that both monetary and non-monetary reciprocity among lineage members are more likely to happen in types 1 and 2 villages than type 3 villages. With regard to across-lineage cooperation, villagers from type 1 and 2 villages are found to fulfill higher percentage of the requirement of pay-free labor than those from type 3 villages. We also find that types 1 and 2 villages spend more share of village budget on public goods than type 3 villages, and the share of public goods in village budget grows faster in type 1 villages than type 3 villages from 1998 to 2002.

In order to present more robust identification, we utilize the differences in the lineage culture between the South and the North of China. Traditionally, the lineage culture is more prevalent in the South than in the North. Consequently, if the relationship between lineage-based heterogeneity and cooperative behavior is causal, we should find the relationship is stronger in the South than in the North. Using a D-in-D model, we find that the impacts of the type 1 and type 2 indicators on villagers' cooperative behavior are indeed significantly greater in the South than in the North of China. [10]

This paper contributes to a large empirical literature on the relationship between cooperation and diversity by examining intra-group and inter-group cooperative behavior simultaneously. Our empirical findings suggest that within-group and across-group reciprocity are not necessarily negatively correlated. In other words, in-group trust does not necessarily happen at the cost of out-group trust. The results presented in this paper also provide some insights into the role that lineage networks play in the success of economic growth in rural China. Actually, the lineage affiliation can serve as a good substitute to promote intra-lineage

[10] except for the outcomes of the share of village budget on public goods and the growth of the share

cooperation when there is lack of formal institutions. Moreover, the existence of large lineages may promote rural development because lineage-based homogeneity supports and facilitates public-good provision.

CHAPTER 4: ASYMMETRIES IN FRIENDSHIP NETWORKS

4.1 Introduction

In a model of friendship networks, costly effort by agents yields expected benefits from direct and indirect links with other agents. This model was introduced by Brueckner (2006) and differs from the rest of the literature. It captures the realism involved in creating friendships: the relationship between two agents is a probabilistic outcome which depends on the effort incurred by both agents. It is also assumed that the probability of link success between any two individuals is independent of the probability of a successful relation between any other pair of individuals. When a friendship link between agents i and j is successful, it provides each of the two players benefits associated with this direct link as well as benefits from all the other direct links of the other player. In other words by having a friendship with j, player i acquires direct link benefits from j as well as indirect benefits of a lesser value from all the other direct links of player j. These indirect benefits capture the notion of friends of friends.

This paper extends the work of Brueckner (2006) by introducing different types of asymmetries in the model of friendship networks. Given the model setup, asymmetries can occur either in values or costs or in the network structure itself. Brueckner himself proposes value based asymmetry and network asymmetry but considers only specific examples for both types. In the real world, there is yet another type of asymmetry. Certain individuals are good at social networking while many others are poor at even hosting a holiday party. This is to say, with the same level of effort, the cost of forming friendships is less for certain individuals but high for many others. This idea motivates the analysis of cost based asymmetry. Since Roy and Sarangi (2009) have already examined value based asymmetry, in this paper we first introduce cost based asymmetry and then focus on network asymmetry. Unlike Brueckner (2006) where asymmetries are examined only for very small sets of agents, for both instances we consider the general case with n agents.

We consider individuals' behavior under different types of asymmetries when determining the allocation of resources and the structure of social relationships. The examination of asymmetries is important for two reasons. First, most economic environments are not characterized by homogeneous agents. Second, they act as a robustness check for results obtained in the homogenous model. Thus, the present extensions could benefit both decisions makers and researchers in important areas when they face different occasions.

Very briefly, the literature in economics on network formation can be divided into two well-known approaches. One approach due to Jackson and Wolinsky (1996) involves costly mutual consent in link formation. This idea is present in the model of friendship networks, since it is necessary for both agents to incur costly effort in order to establish a link. The equilibrium concept used in Jackson and Wolinsky however is a non-strategic link based concept called pairwise stability. Pairwise stability requires that no pair of agents who have a link wish to delete it and pairs of agents with no links do not wish to add one. Friendship networks on the other hand use Nash equilibrium as the stability concept. In this sense it is closer to the second approach introduced by Bala and Goyal (2000a) where a Nash network consists of all agents playing a best response to other linking strategies. In this framework only the agent initiating a link incurs its costs and thus mutual consent is not modeled explicitly. Note that cost and benefits of links in the network in both these approaches are usually exogenously given.[11]

[11] Probabilistic link formation and expected benefits have been analyzed for Nash networks by Bala and Goyal (2000b) and Haller and Sarangi (2005). Bloch and Dutta (2008) also consider stochastic links whose strength depends on the effort of the agents involved in the link. However their setup is different and the focus is not on asymmetries.

Our paper extends Brueckner's results for asymmetric situations with a small number of agents by allowing for an arbitrary number of finite agents. Although, Brueckner states that the general case cannot be solved, we are able to examine the equilibrium effort levels by looking at different partitions of the effort space.

The rest of the paper is organized as follows. In Section 2 we set up the model and discuss the cost asymmetric and structure asymmetric networks. Section 3 offers a summary of our results.

4.2 The model

In this section, we set up the model of friendship networks presented in Brueckner (2006). Then, we introduce different types of asymmetries and provide the results for these cases. Given different types of network structures, we focus on equilibrium effort levels for agents.

4.2.1 Model setup

Let $N = \{1,2,\ldots,n\}$ be the set of agents. We use $a(i)$ to denote the neighborhood set of agent i, i.e., $a(i) \equiv \{j \mid i \text{ and } j \text{ are acquainted}\}$. Let $e_{ij} \in [0,+\infty)$ denote the effort expended by agent i in attempting to establish a friendship link with $j \in a(i)$. The probability of friendship between i and j is a function of the effort of both agents, which is denoted by $P(e_{ij}, e_{ji})$. The function P is a concave function which satisfies $0 \leq P < 1$, and is increasing in both arguments implying $\frac{\partial P}{\partial e_{ij}} > 0$ and $\frac{\partial P}{\partial e_{ji}} > 0$. The second partial derivatives of P are assumed to be negative, implying $\frac{\partial^2 P}{\partial e_{ij}^2} < 0$ and $\frac{\partial^2 P}{\partial e_{ji}^2} < 0$. In addition, P is a symmetric function, i.e., $P\left(e_{ij}, e_{ji}\right) = P(e_{ji}, e_{ij})$.[12] When $e_{ij} = e_{ji} = 0$, $P = 0$ between i and j. In other words, i and j cannot be friends if both of

[12] Hence, $P\left(e_{ij}, e_{ji}\right) = 2[\arctan(e_{ij} + e_{ji})]/\pi$ can be an example of function P that satisfies all the properties stated above.

them do not spend any effort in establishing the relationship between them. If $P \neq 0$ between i and j, then we say i and j are *connected*.

Effort is costly for all agents. The cost of effort exerted by agent i to establish the ij friendship is $C(e_{ij})$. It is assumed C is increasing and strictly convex. In other words, $C' > 0$ and $C'' > 0$. Also, $C(0) = 0$. Indirect benefits do not require benefits, i.e., the benefits i acquires from indirect friends are free. Therefore, agent i's effort cost for all her possible friendship links is given by $\sum_{j \in a(i)} C(e_{ij})$.

Next we describe the benefit from friendship. In this paper, we follow Brueckner (2006) and only consider benefits from direct friend and friend's friend. This is not an uncommon assumption in the literature since benefits often vanish according to distance. To keep the problem tractable and focus on asymmetries, it is convenient to assume that benefits stop at a length of 2.[13] Let $u_{ij} > 0$ and $v_{ik} > 0$ denote agent i's benefit from a direct friend j and j's direct friend k respectively. Then, assume $u_{ij} > v_{ik}$, i.e., the benefit from a direct friendship is always greater than the benefit from an indirect friendship. We also assume that these friendship benefits are cumulative. This is equivalent to saying that agent i can be both a direct and an indirect friend of agent j, and get benefits from both associations. The expected benefits from friendships can be written as the total benefits minus the costs of effort as shown below,

$$B_i = \sum_{j \in a(i)} P\left(e_{ij}, e_{ji}\right)\left[u_{ij} + \sum_{h \in a(j),\ h \neq i} v_{ih} P\left(e_{jh}, e_{hj}\right)\right] - \sum_{j \in a(i)} C\left(e_{ij}\right). \qquad (4.1)$$

The first term $P\left(e_{ij}, e_{ji}\right)u_{ij}$ in (4.1) is the expected benefits of individual i from all her direct friendship links. The second summation combined with $P\left(e_{ij}, e_{ji}\right)$ captures the expected

[13] This allows us to go beyond direct links by considering also indirect links which is important to demonstrate the importance of a network.

benefits from indirect friendships formed with agent j. The last term on the right hand side of (4.1) gives the cost of effort of agent i.

Agent i has to decide how much effort to devote in establishing a friendship link across different agents with whom individual i is acquainted. Denote individual i's decision vector of effort spent across links as strategy S_i. And S_{-i} is the strategy profile played by all agents except agent i.

Definition 4.2.1 *The strategy S_i is said to be a best response of agent i against the strategy profile S_{-i}, if $B_i(S_i, S_{-i}) \geq B_i(S_i', S_{-i})$ for all other feasible S_i'.*

Definition 4.2.2 *The set of all agent i's best responses to S_{-i} is denoted by $BR_i(S_{-i})$. A network is defined as Nash network if $S_i \in BR_i(S_{-i})$ for each agent i.*

Hence in a Nash network, each agent i chooses optimal effort level $e_{ij}, j \in a(i)$ to maximize her own benefits, given other agents' strategies.

The first-order condition for the choice of e_{ij} can be written as

$$\frac{\partial B_i}{\partial e_{ij}} = \frac{\partial P(e_{ij}, e_{ji})}{\partial e_{ij}} \left[u_{ij} + \sum_{h \in a(j),\ h \neq i} v_{ih} P(e_{jh}, e_{hj}) \right] - C'(e_{ij}) = 0. \qquad (4.2)$$

Equation (4.2) balances the marginal gains of both direct and indirect friendships against the marginal cost of effort.

Observe that this is a finite game, according to Nash's theorem, the equilibrium exists. Note that, for the fully symmetric case, where friendship benefits are uniform across all individuals and each person is acquainted with everyone else, we have $u_{ij} \equiv u$ and $v_{ij} \equiv v$, for all i and j, and (4.2) is reduced to

$$P'(e, e)[u + (n-2)vP(e, e)] = C'(e). \qquad (4.3)$$

Next, we compute the socially optimal effort level. The social welfare function is given by $W = \sum_{i=1}^{n} B_i$. According to (1), we have

$$W = \sum_{i \in N} \Big(\sum_{j \in a(i)} P(e_{ij}, e_{ji}) \Big[u_{ij} + \sum_{h \in a(j),\ h \neq i} v_{ih} P(e_{jh}, e_{hj}) \Big] - \sum_{j \in a(i)} C(e_{ij}) \Big) \qquad (4.4)$$

Definition 4.2.3 *The network g^* is said to be efficient if $W(g^*) \geq W(g')W$ for any possible network structure g'.*

So an efficient network has the maximum social welfare. Following Brueckner (2006), we can write the first order conditions for optimal choice as follows, if symmetry is taken into account,

$$2P'(e^*, e^*)[u + 2(n-2)vP(e^*, e^*)] = C'(e^*), \qquad (4.5)$$

where e^* denotes the socially optimal effort level. From Proposition 1 (P. 854, Brueckner (2006)), we know that a Nash network is not efficient. People do not expend enough effort in forming friendship links.

Our first result brings out a general property of Nash networks when agents have symmetric value and costs: in a non-empty equilibrium, the entire society is connected.

Proposition 4.2.1 *A non-empty Nash network is connected.*

Proof. First, we assume the opposite is true: a non-empty Nash network is connected. Let agent i be the agent who has the largest number of direct friends, j be one of i's direct friends and k be any agent in another partition since the network is not connected. Then for agent j, $P'(e_{ji}, e_{ij})[u + v_j] = C'(e_{ji})$ where v_j stands for all of agent j's indirect benefits via i. If agent k also spends some small effort e_{ki} in establishing a relationship with agent i and $e_{ki} < e_{ji}$, then $P'(e_{ji}, e_{ij}) < P'(e_{ki}, 0)$. Moreover, the indirect benefit k can obtain, v_k, is greater than v_j. Because k also gets indirect benefit from j besides all of j's indirect benefits via i.

Hence, $P'(e_{ki}, 0)[u + v_k] > P'(e_{ij}, e_{ji})[u + v_j] = C'(e_{ji}) > C'(e_{ki})$. In other words, for agent k, spending more effort in establishing friendship with agent i is profitable. This indicates that the current network is not an equilibrium, which results in a contradiction. So a Nash network must be connected. ■

We will now proceed to introduce different types of asymmetries into the model of friendship networks.

Given the structure of the model setup there can be three possible types of asymmetries: (i) value based asymmetry, (ii) cost based asymmetry and (iii) network asymmetry. Brueckner (2006) introduces value based asymmetry and calls it the magnetic agent problem. The magnetic agent is simply the agent who offers the highest benefits through a direct or indirect link. Brueckner solves this model for the case of $n = 3$, noting that it is not possible to generalize this case further. Roy and Sarangi (2009) revisit this model and provide a solution for the class of m-regular networks.

Definition 4.2.4 *A network with* n *agents is said to be* m*-regular if every agent has* m*-direct neighbors. Formally,* $|a(i)| = m$ *, where* $m \in [2, n-1]$.

Proposition 4.2.2 [Proposition 2, Page 5, Roy and Sarangi (2009)] *Consider the set of m-regular networks where every agent has access to benefits from the magnetic agent. Let* $\tilde{e} > \hat{e}$*. Then non-magnetic agents expend more effort attempting to link with agent 1 than she expends attempting to link with them. The non-magnetic agents expend an intermediate amount of effort in linking with one another. More precisely,* $e_{x1} > e_{xx} \geq e_{1x}$.

In this paper we introduce another type of magnetic agent we label as the *cost-magnetic* agent. This is the agent who has the lowest connection cost among all agents. Before allowing for n agents, we first discuss the results for the case of $n = 3$ as Brueckner (2006) does. We find that the result is consistent with that in Brueckner (2006). A general solution for n agents remains difficult and m-regular networks are no longer feasible here.[14] So we restrict our analysis to a class of inter-linked stars.

Definition 4.2.5 *A network with* n *nodes is said to be an inter-linked star with* m *centers if each center has exactly* $\mathrm{n}-1$ *links to the other agents. Formally, for each center* i*,* $|a(i)| = n-1$.

Finally, we consider the case of a *knows-everyone* agent. This notion has also been introduced by Brueckner (2006) and is an asymmetric network situation where there exists an agent who is connected to everyone else. Brueckner shows the outcome in this network for the case of $n = 5$ agents. In this paper we introduce the notion of a modified m-regular network and characterize the solution of the universe of n agents.

Definition 4.2.6 *A knows-everyone network with* n *agents is said to be modified* m*-regular if every non-attractive agent has* m *direct neighbors, while the only attractive agent knows everyone. Formally,* $|a(i)| = m$*, where* $m \in [2, n-2]$*;* $|a(1)| = n-1$.

We will start by discussing cost-based asymmetry in the next section.

4.2.2 Cost asymmetry

In this section we introduce cost asymmetry by means of the *cost-magnetic* agent. To be specific, it costs less to connect directly to the *cost-magnetic* agent than any other agent. We start with the simple case with only three agents, all of whom are connected with each other. Without loss of generality, let agent 1 be the *cost-magnetic* agent, and $x = 2$ or 3 be the other two agents. The cost of linking to a magnetic agent is given by $C_1(e_{x1})$ while the cost of linking to a non-magnetic agent is given by the usual $C(e_{xx})$, or $C(e_{1x})$. Here $0 < C_1(e) < C(e)$ for any e. Due to the properties of the cost function it is also the case that $C_1'(e) < C'(e)$ for any e.

[14] We will discuss this later formally.

Assume further that efforts required to establish friendship links are substitutes so that $P(e_{ij}, e_{ji})$ can be written as $P(e_{ij} + e_{ji})$. Next, define the total effort for each link as follows:

$$\tilde{e} = e_{1x} + e_{x1}$$

$$\hat{e} = 2e_{xx}$$

Our question is whether links involving agent 1 are more likely to form and the following results can be established. Consistent with the findings of Brueckner (2006), the answer is affirmative as shown in the following proposition.

Proposition 4.2.3 *Consider the cost-magnetic case with* 3 *agents, non-magnetic agents expend more effort attempting to link with agent 1 than agent 1 expends attempting to link with them. More precisely,* $e_{x1} > e_{xx} > e_{1x}$. *The inequality* $\tilde{e} > \hat{e}$ *holds.*

Proof. The first order conditions are as follows:

$$P'(\tilde{e})[u + vP(\hat{e})] = C'(e_{1x}), \qquad (4.6)$$

$$P'(\tilde{e})[u + vP(\tilde{e})] = {C_1}'(e_{x1}), \qquad (4.7)$$

$$P'(\hat{e})[u + vP(\tilde{e})] = C'(e_{xx}). \qquad (4.8)$$

(1) Consider first why $\tilde{e} \leq \hat{e}$ is not possible. If $\tilde{e} < \hat{e}$, from (4.6) and (4.8) we can get $P'(\tilde{e})[u + vP(\hat{e})] > P'(\hat{e})[u + vP(\tilde{e})]$. This indicates $e_{1x} > e_{xx}$. Since $\tilde{e} < \hat{e}$, $e_{1x} < e_{xx}$. Then if we look at (4.7) and (4.8), $C'(e_{xx}) > C'(e_{x1}) > {C_1}'(e_{x1})$, which is equivalent to $P'(\tilde{e})[u + vP(\tilde{e})] > P'(\tilde{e})[u + vP(\tilde{e})]$ or $P'(\hat{e}) > P'(\tilde{e})$. This is a contradiction since the function P' is decreasing.

Furthermore, in the equilibrium, $\tilde{e} \neq \hat{e}$. To see why, we first assume $\tilde{e} = \hat{e}$. From (4.6) and (4.7), it is easy to see $e_{x1} > e_{1x}$. But from (4.6) and (4.8), $e_{1x} = e_{xx}$. So $\tilde{e} \neq \hat{e}$. This is a contradiction. Thus the inequality $\tilde{e} > \hat{e}$ holds in the equilibrium.

(2) If $\tilde{e} > \hat{e}$, from (4.6) and (4.8) we can get $P'(\tilde{e})[u + vP(\hat{e})] < P'(\hat{e})[u + vP(\tilde{e})]$. This indicates $e_{1x} < e_{xx}$. Since $\tilde{e} > \hat{e}$, $e_{1x} > e_{xx}$ is true. In other words, $e_{x1} > e_{xx} > e_{1x}$.■

In the *cost-magnetic* case with only 3 agents, the inequality $\tilde{e} > \hat{e}$, holds implying that direct friendships involving agent 1 are more likely to form than direct friendships involving non-magnetic agents.

Remark 4.2.1 *It is not possible to generalize Proposition 4.2.3 to arbitrary networks.*

To see why, consider an arbitrary network g with n agents. Let agent i have n_i neighbors, where $n_i \in [1, n-1]$. The first order conditions are as follows:

$$P'(\tilde{e})[u + (n_i - 1)vP(\hat{e})] = C'(e_{1i}), \tag{4.9}$$

$$P'(\tilde{e})[u + (n_1 - 1)vP(\tilde{e})] = C_1{}'(e_{i1}), \tag{4.10}$$

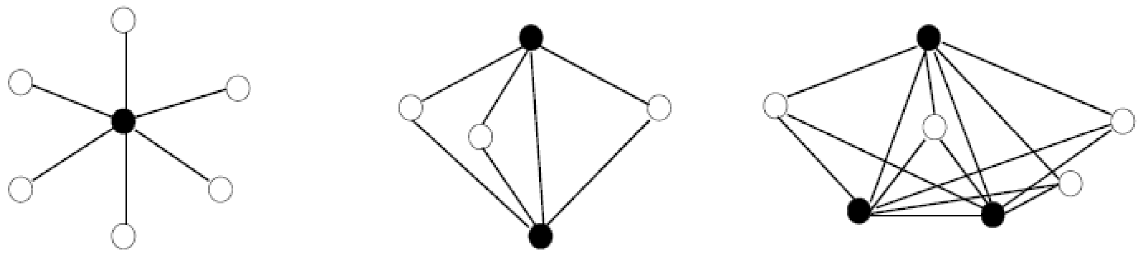
$$P'(\hat{e})\left[u + vP(\tilde{e}) + (n_j - 2)vP(\hat{e})\right] = C'\left(e_{ij}\right), \tag{4.11}$$

$$\text{or } P'(\hat{e})\left[u + (n_j - 1)vP(\hat{e})\right] = C'\left(e_{ij}\right). \tag{4.12}$$

Consider the inequality $e_{i1} < e_{1i}$. Then using equation (4.9) and (4.10), it is easy to verify that the result will depend on the comparison between $n_1 - 1$ and $n_i - 1$. The same problem exists for all the other inequalities.

Since it is not possible to get a general result for arbitrary networks, we restrict our analysis to a special class of networks, inter-linked stars. Inter-linked stars are possible Nash networks structure when the benefit of linking with a non-center agent cannot cover the cost. However, linking with a center is still possible because of the indirect benefits that the center brings. Figure 4.1 gives some basic ideas about how inter-linked stars look like, where black dots indicate centers.

We assume one of the centers is the *cost-magnetic* agent. Again, let agent 1 be the *cost-magnetic* agent and agent 2 be the other centers. Now x are the non-center agents. The cost of linking to a magnetic agent is given by $C_1{}'(e_{x1})$ and $C_1{}'(e_{21})$ while the cost of linking to a non-magnetic agent is given by the usual $C'(e_{x2})$, $C'(e_{2x})$, $C'(e_{12})$ or $C'(e_{1x})$. In an inter-linked star, there are three effort levels. We define the total effort for each link as follows:



(1) One center (2) Two center (3) Three centers

Figure 4.1: Inter-linked stars with 7 nodes

$$\tilde{e} = e_{1x} + e_{x1},$$

$$\bar{e} = e_{12} + e_{21},$$

$$\hat{e} = e_{2x} + e_{x2}.$$

Our goal here is to identify sufficient conditions on the effort level given the probability function under which a given network can be supported as a Nash equilibrium. The following results are based on different partitions of the space of effort.

Proposition 4.2.4 *Let g be an inter-linked star with n agents and m centers. In Nash networks, when $\tilde{e} > \hat{e}$, the agent x expends more effort attempting to link with agent 1 than attempting to link with 2. Agent 1 expends less effort in linking with x than agent 2 expends attempting to link* *with x. More precisely,* $e_{1x} < e_{x2} < e_{x1}$ *while* $e_{1x} < e_{2x} < e_{x1}$. *When* $\tilde{e} < \hat{e}$, *the results are opposite, i.e.,* $e_{2x} < e_{x1} < e_{x2}$ *while* $e_{2x} < e_{1x} < e_{x2}$.

Proof. To maintain the importance of the center, we assume $m < n/2$. The first order conditions are as follows:

$$P'(\tilde{e})[u + (m-1)vP(\hat{e})] = C'(e_{1x}), \qquad (4.13)$$

$$P'(\tilde{e})[u + (n-m-1)vP(\tilde{e}) + mvP(\bar{e})] = C_1{}'(e_{x1}), \qquad (4.14)$$

$$P'(\hat{e})[u + (n-m-1)vP(\hat{e}) + mvP(\bar{e})] = C'(e_{x2}), \qquad (4.15)$$

$$P'(\hat{e})[u + (m-1)vP(\tilde{e})] = C'(e_{2x}). \qquad (4.16)$$

(1) We first examine the case when $\tilde{e} > \hat{e}$. Since $\tilde{e} > \hat{e}$, we can get $P'(\tilde{e}) < P'(\hat{e})$ and $(m-1)vP(\hat{e}) < (m-1)vP(\tilde{e})$. So it is easy to obtain $P'(\tilde{e})[u + (m-1)vP(\hat{e})] < P'(\hat{e})[u + (m-1)vP(\tilde{e})]$, which indicates $e_{2x} > e_{1x}$ from (4.13) and (4.16).

Now we establish $e_{x1} > e_{x2}$. Assume the opposite is true, $e_{x1} \leq e_{x2}$ holds. Then $e_{x1} + e_{1x} \leq e_{x2} + e_{2x}$ or $\tilde{e} \leq \hat{e}$. This is a contradiction.

Since $(n-m-1)vP(\hat{e}) + mvP(\bar{e}) > (m-1)vP(\hat{e})$, from (4.13) and (4.15) it is easy to get $C'(e_{1x}) < C'(e_{x2})$. So we obtain $e_{x2} > e_{1x}$.

from (4.15) and (4.16), $e_{2x} \leq e_{x1}$ is not possible. Otherwise, the assumption $\tilde{e} > \hat{e}$ would be violated. Therefore, $e_{1x} < e_{x2} < e_{x1}$ and $e_{1x} < e_{2x} < e_{x1}$.

(2) If $\tilde{e} < \hat{e}$, then the results are opposite. We first establish $e_{x2} < e_{1x}$. Now $P'(\tilde{e}) > P'(\hat{e})$ and $(m-1)vP(\hat{e}) > (m-1)vP(\tilde{e})$. Thus, $P'(\tilde{e})[u + (m-1)vP(\hat{e})] > P'(\hat{e})[u + (m-1)vP(\tilde{e})]$, which indicates $e_{x2} < e_{1x}$ from (4.13) and (4.16).

Now we establish $e_{x1} < e_{x2}$. Assume $e_{x1} \geq e_{x2}$ holds. Then $e_{x1} + e_{1x} \geq e_{x2} + e_{2x}$ or $\tilde{e} \geq \hat{e}$. This is a contradiction.

Since $(n-m-1)vP(\tilde{e}) + mvP(\bar{e}) > (m-1)vP(\tilde{e})$, from (4.14) and (4.16) it is easy to get $C_1'(e_{x1}) > C'(e_{2x})$. So we obtain $e_{x1} > e_{2x}$.

$e_{x2} \leq e_{1x}$ is not possible. Otherwise, the assumption $\tilde{e} < \hat{e}$ would be violated.

Therefore, when $\tilde{e} < \hat{e}$, $e_{2x} < e_{x1} < e_{x2}$ and $e_{2x} < e_{1x} < e_{x2}$.

(3) in the equilibrium, $\tilde{e} \neq \hat{e}$. To see why, we first assume $\tilde{e} = \hat{e}$. From (4.14) and (4.15), it is easy to see $e_{x1} > e_{x2}$. But from (4.13) and (4.16), $e_{1x} = e_{2x}$. So $\tilde{e} \neq \hat{e}$. This is a contradiction. ■

Remark 4.2.2 *The cost-magnetic agent may lose importance in a general case with n agents.*

When we consider the general case with n agents, the magnetic agent is not as important as in the case with only 3 agents. Other non-magnetic agents may also be attractive if they have a sufficient number of friends. However, $\tilde{e} > \hat{e}$ allows us to maintain the importance of the *cost-magnetic* agent ensuring the specific ranking of the effort levels shown in the proposition. Agent x are establishing friendship with the centers. If they have a stronger relationship with the *cost-magnetic* agent, then they will spend even more effort into this relationship because the *cost-magnetic* agent is easy to get along with. However, if they don't have such a good relationship with the *cost-magnetic* agent, then the lower link cost is no longer attractive for agent x.

4.2.3 Network asymmetry: the *knows-everyone* agent

Brueckner (2006) introduced the *knows-everyone* problem as another type of asymmetric network. An attractive agent who knows every other agent here reflects asymmetry in network structure. In this problem, one attractive agent is acquainted with the entire universe of agents, while other non-attractive agents are each acquainted with only a subset of the non-attractive agents. That means agents have different sets of neighborhoods although the friendship benefits are symmetric across individuals. As before, we will focus only on regular networks and assume that each non attractive agent is linked to the *knows-everyone* agent and $m-1$ other agents. Note that because of the popularity of the attractive agent, the typical regular network does not apply here. Hence below we introduce the notion of a modified m-regular network. Without loss of generality, assume agent 1 is the attractive agent, while x describes other agents. We have two possible effort levels here:

$$\tilde{e} = e_{1x} + e_{x1}$$

$$\hat{e} = 2e_{xx}$$

In a modified m-regular network, when $m = n-1$, we have a complete network. When the number of agents is even, m can only be an odd number. For instance, there exist networks with $n = 7$ and $m = 2, 3, 4,$ or 5; however when $n = 8$, there is no network for $m = 4$; but it does exist for $m = 3, 5$. Obviously we will only consider situations where the modified m-regular network exists.

For example, in Figure 4.2 we illustrate the difference between m-regular networks and modified m-regular networks. Each network has 6 agents, and every agent is acquainted with 3 neighbors, except that in a modified m-regular network, agent 1 is acquainted with all other agents.

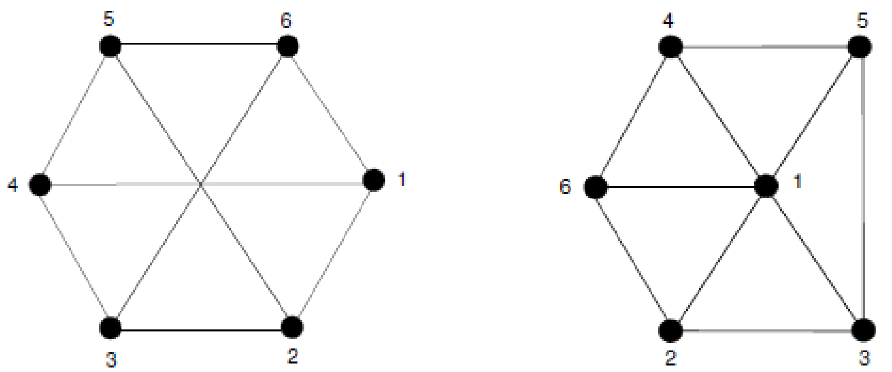


Figure 4.2: M-regular and modified m-regular network with $n = 6, m = 3$

$$P'(\tilde{e})[u + (m-1)vP(\hat{e})] = C'(e_{1x}), \qquad (4.17)$$

$$P'(\tilde{e})[u + (n-2)vP(\tilde{e})] = C'(e_{x1}), \qquad (4.18)$$

$$P'(\hat{e})[u + vP(\tilde{e}) + (m-2)vP(\hat{e})] = C'(e_{xx}). \qquad (4.19)$$

Equations (4.17)-(4.19) are the first order conditions when the number of agents is odd, and $m \leq n-2$. Though it is not possible to set up the model when the number of agents is even, we can still get a general result for the effort investment regarding the different types of links. In practice, when m is small, the importance of the *knows-everyone* agent is maintained. The attractive agent, i.e., the *knows-every* agent provides higher direct and indirect benefits since she knows more people. Intuitively, non-attractive agents have a higher incentive to link with her. On the other hand, attractive agent does not get as much as she gives, so she puts the least effort in forming relationships. However, in a general case where each other agent know a certain number of neighbors, the importance of the attractive agent is diluted and agents' behavior may vary. The proposition below illustrates the results, which are based on different partitions of the space of effort.

Proposition 4.2.5 *Let g be a modified m-regular network. In Nash networks, when $\tilde{e} > \hat{e}$, the non-attractive agents expend more effort attempting to link with agent 1 than agent 1 expends attempting to link with them. The non-attractive agents expend an intermediate amount of effort in attempting to link with one another. That is, $e_{1x} < e_{xx} < e_{x1}$. When $\tilde{e} > \hat{e}$, the results are opposite, i.e., $e_{1x} > e_{xx} > e_{x1}$.*

Proof. Consider the situation when $\tilde{e} > \hat{e}$,

(1) We will first establish $e_{x1} < e_{x1}$. By contradiction, assume $e_{x1} \geq e_{x1}$, then $C'(e_{1x}) \geq C'(e_{x1})$. It follows from (4.17) and (4.19), that $P'(\tilde{e})[u + (m-1)vP(\hat{e})] \geq$

$P'(\tilde{e})[u+(n-2)vP(\tilde{e})]$. Since $\in [2, n-2]$, $m-1 < n-2$, the only way this relationship can hold is $\tilde{e} < \hat{e}$. This is a contradiction. Hence $e_{x1} < e_{x1}$.

(2) Now we establish $e_{xx} > e_{1x}$. Again we prove by contradiction. So assume that $e_{xx} \leq e_{1x}$ holds. Then, $C'(e_{1x}) \geq C'(e_{xx})$. Then from (4.17) and (4.19), it follows that $P'(\tilde{e})[u+(m-1)vP(\hat{e})] \geq P'(\hat{e})[u+vP(\tilde{e})+(m-2)vP(\hat{e})]$. However, given our assumptions the reverse of this inequality holds with a strict sign. This is a contradiction. Hence $e_{xx} > e_{1x}$.

(3) Since $\tilde{e} > \hat{e}$ and $e_{x1} < e_{x1}$, it is not possible to have $e_{xx} \geq e_{x1}$. Hence, $e_{xx} < e_{x1}$.

Therefore, $e_{1x} < e_{xx} < e_{x1}$ when $\tilde{e} > \hat{e}$. Next, we consider the situation when $\tilde{e} < \hat{e}$. From (4.17) and (4.19), we get $C'(e_{1x}) > C'(e_{xx})$. So $e_{1x} > e_{xx}$. It is easy to see that $e_{1x} > e_{xx} > e_{x1}$.

Finally, we establish $\tilde{e} \neq \hat{e}$ in the equilibrium. Assume $\tilde{e} = \hat{e}$. From (4.18) and (4.19), we get $e_{x1} > e_{xx}$. From (4.17) and (4.19), we get $e_{1x} = e_{xx}$. Hence $\tilde{e} \neq \hat{e}$. This is a contradiction. Therefore, in equilibrium, $\tilde{e} \neq \hat{e}$.

Remark 4.2.3 *It is not possible to generalize Proposition 2.5 to arbitrary networks.*

Consider an arbitrary network g with n agents. Agent $i \in N\backslash\{i\}$ has n_i neighbors, where $n_i \in [1, n-2]$. Agent 1 of course has $n-1$ neighbors. The first order conditions are as follows:

$$P'(\tilde{e})[u+(n_i-1)vP(\hat{e})] = C'(e_{1i}), \quad (4.20)$$

$$P'(\tilde{e})[u+(n-1)vP(\tilde{e})] = C'(e_{i1}), \quad (4.21)$$

$$P'(\hat{e})\big[u+vP(\tilde{e})+(n_j-2)vP(\hat{e})\big] = C'\big(e_{ij}\big). \quad (4.22)$$

Using (4.20) and (21), since $n-1 \geq n_i - 1$, it is easy to verify that the inequality $e_{x1} > e_{1x}$ holds. However, from (4.21) and (4.22), since $(\tilde{e}) > P(\hat{e})$, $n-2 > n_j - 2$ and $P'(\tilde{e}) < P'(\hat{e})$, the inequality $e_{x1} > e_{xx}$ will depend on the value of parameters. The same problem exists for inequality $e_{1x} < e_{xx}$. Thus for an arbitrary network g, it is possible to claim that the non-attractive agent will expend more effort linking to the *knows-everyone* agent than this agent will spend in linking to the non-attractive agents. It is not possible to establish any other relationships between the effort levels without imposing restrictions on the parameters.

4.3 Conclusion

In this paper, we introduce the *cost-magnetic* agent and also extend the *knows-everyone* agent model of Brueckner (2006) by allowing an arbitrary number of agents. We find that the non-attractive agents expend more effort attempting to link with the attractive agent than the attractive agent expends to link with them for both these types of asymmetries when the total effort for link between non-attractive agent and attractive agent is greater than the effort for the link between two non-attractive agents. The paper shows that for an arbitrary network this ranking of effort levels depends on the parameter values. What is interesting is that our results are consistent with findings of Brueckner (2006) who only considers a small set of agents as well as with Roy and Sarangi (2009) who consider the value magnetic agent problem for regular networks. Thus for asymmetric networks (whether in costs, values, or architectures), the equilibrium effort choices regarding the links between different types of agents are robust across the various models.

CHAPTER 5: CONCLUSION

This dissertation contributes to the literature by investigating individuals' interaction in different contexts using social network analysis.. The second chapter studies a setting in which agents spend resources in both giving information to and asking information from connections to their neighbors. The third chapter empirically tests how the pattern of village structure, in terms of lineage network composition, affects people's reciprocal behavior. The last chapter analyzes friendship networks.

The second chapter generalizes the models of link formation of Rogers (2005) by combining the model of asking and the model of giving and allowing that giving and asking choices can be made separately and simultaneously by each agent. We focus on two specifications: the concave specification and the linear specification. Under the concave specification, the results show that people usually spend more effort in asking for help from those with more information, and spend more effort in offering help to those from whom they can receive more information. A social planner wants people to spend more effort in giving if they have better aggregate relationships with others. If an agent's direct neighbor benefits more from receiving information, then this agent's effort in asking information from is increasing with the effort in giving information to this neighbor. Then, we turn our attention to the linear case. In the linear case, we find people only spend resources in asking because the behavior of giving suffers more depreciation. In both the *endowment-attractive* and the *budget-attractive* cases, non-attractive agents spend all their resources connecting with the attractive agents. This conclusion is consistent with the finding of Breuckner (2006). Moreover, in both cases, efficient networks coincide with Nash networks.

The second chapter studies the relationship between lineage-based fragmentation and villagers' cooperative behavior. We find that people of types 1 and 2 villages are more likely than those of type 3 to have both within-lineage and across-lineage cooperation. We also find that the share of public goods in village budget grows faster in type 1 villages than type 3 villages from 1998 to 2002 after a series of political revolution initiated in 1998. Thus, the lineage affiliation can actually serve as a good substitute to promote intra-lineage cooperation when there is lack of formal institutions. Moreover, the existence of large lineages may promote rural development because lineage-based homogeneity supports and facilitates public-good provision.

The fourth chapter introduces the *cost-magnetic* agent and also extend the *knows-everyone* agent model of Brueckner (2006) by allowing an arbitrary number of agents. We find that the non-attractive agents expend more effort attempting to link with the attractive agent than the attractive agent expends to link with them for both these types of asymmetries when the total effort for link between non-attractive agent and attractive agent is greater than the effort for the link between two non-attractive agents. The chapter shows that for an arbitrary network this ranking of effort levels depends on the parameter values. What is interesting is that our results are consistent with findings of Brueckner (2006) who only considers a small set of agents as well as with Roy and Sarangi (2009) who consider the value magnetic agent problem for regular networks. Thus for asymmetric networks (whether in costs, values, or architectures), the equilibrium effort choices regarding the links between different types of agents are robust across the various models.

In the future, I will address other specifications of the link quality function in the asking and giving model. Another extension may be making people's giving behavior interdependent. For example, if one agent refuses to help the other agent, then the other agent's willingness to offer help will be reduced. The results in the linear case are examined for a simple network with only three agents. Thus another future direction can be based upon examining whether the results still hold in a larger universe of agents. For the empirical study, in the future I will also work on how people exhibit reciprocal behavior in the existence of favoritism and peer effect utilizing CHIPS 2002.

REFERENCES

Alesina, A., & La Ferrara, E. (2002). Who trusts others? *Journal of Public Economics*, 207-234.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic eivisions. *The Quarterly Journal of Economics*, 1243-1284.

Baker, G. R. (2004). Strategic alliance: bridges between "islands of conscious power". *Mimeo, MIT*.

Bala, V., & Goyal, S. (2000). A non cooperative model of network formation. *Econometrica*, 1181-1229.

Bandiera, O., Barankay, I., & Rasul, I. (2005). Cooperation in Collective Action. *Natural Field Experiments* , 0013.

Banerjee, A., Iyer, L., & Somanathan, R. (2005). History, social divisions and public goods in rural India. *Journal of the European Economic Association*, 639-647.

Bardhan, P. (1993). Analytics of the institutions of informal cooperation in rural development. *World Development*, 633–639.

Bardhan, P. (2002). Decentralization of governance and development. *Journal of Economic Perspectives*, 185-205.

Blalock, H. M. (1967). Toward a theory of minority-group relations. *New York: John Wiley & Sons*.

Bloch, F., & Dutta, B. (2009). Communication networks with endogenous link strength. *Games and econonomic behavior,* 39-56.

Bowles, S., & Gintis, H. (2008). Cooperation. *The New Palgrave Dictionary of Economics*.

Bramoulle, Y. a. (2007). Local public goods in networks. *Journal of Economic Theory*, 478-494.

Bramoulle, Y., Lopez-Pintado, D., Goyal, S., & Vega-Redondo, F. (2004). Network formation and anti-coordination games. *International Journal of Game Theory,* 1-19.

Brueckner, J. K. (2006). Friendship networks. *Journal of Reginal Science,* 847-865.

Coate, S., & Ravallion, M. (1993). Reciprocity without commitment: characterization and performance of informal insurance arrangements. *Journal of Development Economics*, 1-24.

Conniff, L. (2004). Village Elections: Emergence of a democratic political culture or reinforcement of a repressive regime? *Independent Study Project (ISP) Collection*, 496.

Daniel A. Hojman, A. S. (2006). Endogenous networks, social games, and evolution. *Games and Economic Behavior*, 112-130.

Easterly, W., & Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions. *Quarterly Journal of Economics*, 1203-50.

Egel, D. (2011). Tribal heterogeneity and the allocation of development resources: evidence from Yemen. *Working Paper*.

Freedman, M. (1965). Lineage organization in southeastern China. *University of London, Athlone Press.*

Galeotti, A., Goyal, S., Jackson, M. O., & Vega-redondo, F. (2010). Network games. *The Review of Economic Studies*, 218-244.

Girard, V. (2011). The impact of inter-group relationships on intra-group cooperation. A case study in rural India. *Proceedings of the German Development Economics Conference*, Berlin 2011 No. 32.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78:1360-80.

Greif, A. (1993). Contract enforceability and economic Institutions in early trade: the Maghribi traders' coalition. *American Economic Review*, 525–48.

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural biases in economic exchange. *Quarterly Journal of Economics*, 124.

GUOBANFA(Office of StateCouncil). (2002). Guan yu zuo hao 2002 nian kuo da nong cun shui fei Gaige shi dian gong zuo de tong zhi [Circular on expansion of pilot experiments on rural Tax for Fee reform in 2002]. *Office of State Council, Beijing, China (in Chinese)*.

Jackon, M., & Wolinsky, A. (1996). A strategic model of economic and social networks. *Journal of Economic Theory*, 44-74.

Jackson, M. O. (2007). Prepared for the missing links*: Formation and Decay of Economic*.

Jackson, M. O., & Rogers, B. W. (2005). The economics of small worlds. *Game Theory and Information*.

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 39-43.

Leigh, A. (2006). Trust, inequality, and ethnic heterogeneity. *The Economic Record*, 268-280.

Linxiu Zhang, R. L., & Rozelle, S. (2006). Investing in rural China: tracking China's commitment to modernization. *The Chinese Economy*, 57-84.

Luo, R., Zhang, L., Huang, J., & Rozelle, S. (2007). Elections, fiscal reform and public goods provision in rural China. *Journal of Comparative Economics*, 583–611.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. T*he Academy of Management Review*, 709-734.

Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions,and public Goods in Kenya. *Journal of Public Economics*, 89 (11–12), 2325–2368.

Newman, M., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review*, E 69.

Pan, Y. (2012). Born with the right surname. *Working Paper*.

Peng, Y. (2004). Kinship networks and entrepreneurs in China's transitional economy. *American Journal of Sociology*, 1045-74.

Posner, R. (1980). A theory of primitive society, with special reference to law. *Journal of Law and Economics*, 23(1), 1-53.

Putnam, R. D. (2007). E Pluribus Unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize lecture. *Scandinavian Political Studies*, 137-174.

Rees, A. (1966). Information networks in labor markets. *American Economic Review*, 56:559-66.

Rogers, B. (2005). A strategy theory of network status. *Mimeo*, MEDS, Northwestern University.

Tsai, L. L. (2007). Solidary groups, informal accountability, and local public goods provision in rural China. *American Political Science Review*, 355-372.

Vigdor, J. (2004). Community composition and collective action: analyzing initial mail response to the 2000 census. *The Review of Economics and Statistics*, 303-312.

Wang, J. (2006). Village governance in Chinese history. *Unpublished Dissertation*.

Wang, X. (2007). The difference between the south China and the north China in lineage inhabitation. *Research on Financial and Economic Issues*, 288(11), 20-31.

Zhang, H. (2004). An analysis of the transition of local governance in rural China . *Study of Economic Issues (In Chinese)*, 79-83.

Zhang, X., Fan, S., Zhang, L., & Huang, J. (2004). China's local governance and public goods provision in rural. *Journal of Public Economics*, 2857– 2871.

VITA

Quqiong He was born in Hubei, China. She attended Wuhan University in China where she earned a Bachelor of Art in Economics and a Bachelor of Science in Mathematics. In August, 2008, she joined Louisiana State University to pursue her graduate studies in economics. Her research focuses on social networks and applied microeconomics. Quqiong will complete the degree of Doctor of Philosophy in August 2013.