

© 2010 Dennis Charles O'Dea

THREE ESSAYS ON SOCIAL NETWORKS

BY

DENNIS CHARLES O'DEA

DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Economics  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Steven R. Williams, Chair  
Professor Daniel Bernhardt  
Professor Stefan Krasa  
Professor Mattias Polborn

# ABSTRACT

In three chapters I study the formation of social networks, and the impact the structures that arise may have in various economic settings. First, I develop a model of social network formation with heterogeneous agents and incomplete information. The model predicts an equilibrium in which agents sort themselves into “insiders” and “outsiders.” Insiders form many links to one another, and form a dense core structure in the network, while outsiders coordinate their links by connecting to an insider, and form a sparse periphery . Networks form stochastically, contingent on the private values of each agent, and include more realistic structures than networks arising among homogenous agents. I characterize the set of equilibria and identify its extremes, which have a natural interpretation as public good provision. One extreme, when agents are all insiders, is equivalent to the provision of a pure public good, and suffers from free-riding. The other extreme, when every agent but one is an outsider, the equilibrium is equivalent to the provision of an excludable public good, and suffers from coordination problems.

I next develop expand this model to study the provision local public goods, such as information, that is shared along the network. Individuals may choose to provide a public good that is not excludable among their peers in a social network. The network is formed endogenously, as agents non-cooperatively choose their social ties. I characterize the set of equilibria, and examine the relationship between public good provision and social network formation. I find that the architecture of the social network determines the strategic interaction between link formation and public good provision; for some networks, links are strategic substitutes, so that agents attempt to free-ride on their peer’s links. This leads to higher levels of public good provision, and specialization in roles: Agents either invest in the public good or form links, but not both. For other networks, however, links are strategic complements, so that agents coordinate their links by connecting to central agents. This leads to lower levels of public good provision, and less specialization; some agents will both link and invest, leading to lower welfare.

Finally I present a model of time allocation between formal and informal labor supply

where workers learn of informal job opportunities from their peers in a social network. In addition to formal income taxation and enforcement, individuals labor supply decisions depend on the number of their peers with informal jobs and the strength of social ties. Workers allocate more time to informal activities when tax enforcement is lax and job information transmission is good. More connected social networks (e.g. wheel, complete) feature lower average income but higher average utility than poorly connected social networks (e.g. star, empty). Average income may be non-monotonic in tax enforcement.

*To my mother Kathleen and my wife Meredith, for their constant love and support.*

# ACKNOWLEDGMENTS

Many thanks are due to my advisors, Steven Williams and Daniel Bernhardt, for their support and encouragement, in addition to the many suggestions and comments that improved this work. Thanks also to seminar participants at the University of Illinois and Urbana - Champaign, especially Brett Graham, Michael Dorsch, Adam Stevenson, and Jungmo Yoon. All remaining errors are of course my own.

# TABLE OF CONTENTS

|  |      |
|--|------|
| LIST OF TABLES . . . . .   | viii |
| LIST OF FIGURES . . . . .  | ix   |
| CHAPTER 1 INTRODUCTION . . . . .   | 1    |
| CHAPTER 2 HETEROGENEITY IN SOCIAL NETWORKS AND THE CLAS-<br>SICAL THEORY OF PUBLIC GOODS . . . . . | 3    |
| 2.1 Introduction . . . . .   | 3    |
| 2.2 The Model . . . . .  | 6    |
| 2.3 The Three Agent Case . . . . .   | 11   |
| 2.4 Insiders and Outsiders for $n > 3$ . . . . .   | 14   |
| 2.5 Equilibrium in Public Good Provision . . . . .   | 22   |
| 2.6 Public Good Equivalence . . . . .  | 23   |
| 2.7 Conclusion . . . . .   | 26   |
| CHAPTER 3 LOCAL PUBLIC GOODS AND NETWORK FORMATION . . . . .                                       | 27   |
| 3.1 Introduction . . . . .   | 27   |
| 3.2 The Model . . . . .  | 30   |
| 3.3 Equilibrium . . . . .  | 31   |
| 3.4 Discussion . . . . .   | 37   |
| 3.5 Welfare . . . . .  | 42   |
| 3.6 Conclusion . . . . .   | 43   |
| CHAPTER 4 INFORMAL WORK NETWORKS . . . . .   | 44   |
| 4.1 Introduction . . . . .   | 44   |
| 4.2 The Model . . . . .  | 47   |
| 4.3 Informal Work Networks . . . . .   | 51   |

|                              |     |
|------------------------------|-----|
| 4.4 Conclusions . . . . .    | 65  |
| APPENDIX A PROOFS . . . . .  | 68  |
| REFERENCES . . . . .         | 96  |
| AUTHOR'S BIOGRAPHY . . . . . | 101 |

# LIST OF TABLES

|     |  |    |
|-----|--|----|
| 3.1 | Critical Values of $k$ . . . . .   | 41 |
| 3.2 | Total Welfare . . . . .  | 42 |
| 4.1 | The sequence of events in period $t$ . . . . .   | 50 |
| 4.2 | Long Run Averages for the baseline economy: $\tau = 0.25, \alpha = 0.10, \beta = 0.70$ . .     | 62 |
| 4.3 | Long Run Averages for the alternative economy: $\tau = 0.25, \alpha = 0.10, \beta = 0.30$ . 63 |    |

# LIST OF FIGURES

|     |   |    |
|-----|---|----|
| 2.1 | A profile of actions and the resulting graph. . . . .   | 7  |
| 2.2 | Clockwise from the upper left, a center-sponsored star, a periphery-sponsored star, an incomplete periphery-sponsored star, and a redundant center-sponsored star. . . . .  | 9  |
| 2.3 | Equilibrium strategies in the three agent case. Solid lines indicate pure strategies in link formation, dashing lines indicate mixed strategies. . . . .  | 12 |
| 2.4 | Four possible realized networks for Uniform values, $c = .5$ , with 15 insiders and 15 outsiders. . . . .   | 16 |
| 2.5 | The insiders-outsiders equilibrium for Uniform values, $\theta$ , for $n = 5$ , $I = 2$ . $\theta_I$ is given by the dashing line, $\theta_O$ by the solid line. The bolded portions represent the range of $c$ over which this is an equilibrium. . . . .  | 17 |
| 2.6 | Center-sponsored equilibrium with uniform values, in bold. . . . .  | 19 |
| 2.7 | Periphery-sponsored equilibrium with uniform values, in bold. . . . .   | 22 |
| 3.1 | Linking Strategies in the Center-Sponsorship Equilibrium. Agents form no links if their private cost is above this line, or two links if it is below this line. . . . .   | 33 |
| 3.2 | Linking Strategies in the Periphery-Sponsorship Equilibrium. Outsiders form two links if their cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line. The Insider forms two links if her cost is below the dashed line, or no links if it is above the dashed line. . . . .                                      | 33 |
| 3.3 | Linking Strategies in the Hybrid Equilibrium. The Outsider forms two links if her cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line. The Insiders forms two links if her cost is below the lower dashed line, one link if it is between the two, or zero links if it is above the upper dashed line. . . . . | 34 |
| 3.4 | Linking Strategies in the Mixed Equilibrium. Each agent forms two links if her cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line. . . . .  | 34 |

|      |   |    |
|------|---|----|
| 3.5  | Linking Strategies in the Circle Equilibrium. Each agent forms two links if her cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line. . . . .                                   | 35 |
| 3.6  | Investment Strategies in the Center-Sponsorship Equilibrium. The higher curve is the investment when forming zero links and the lower curve is investment when forming two links. . . . .   | 38 |
| 3.7  | Investment Strategies in the Periphery-Sponsorship Equilibrium. The higher curve is the investment when forming zero links and the lower curve is investment when forming two links. . . . .  | 38 |
| 3.8  | Investment Strategies in the Periphery-Sponsorship Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links. . . . . | 39 |
| 3.9  | Investment Strategies in the Hybrid Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links. . . . .                | 39 |
| 3.10 | Investment Strategies in the Hybrid Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links. . . . .                | 40 |
| 3.11 | Investment Strategies in the Mixed Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links. . . . .                 | 40 |
| 3.12 | Investment Strategies in the Circle Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links. . . . .                | 41 |
| 4.1  | Network Structure Examples: Clockwise from upper left, a complete network, a wheel network, an empty network, and a star network. . . . .   | 49 |
| 4.2  | Most likely state in the long run for the Empty and Complete Networks . . . . .   | 55 |
| 4.3  | Informal Labour Supply and Arrival Probability $\alpha$ . . . . .   | 58 |
| 4.4  | Informal Labour Supply and Probability of Detection $\beta$ . . . . .   | 59 |
| 4.5  | Informal Labour Supply and Income Tax $\tau$ . . . . .  | 60 |
| 4.6  | Long Run Average Total Income and Probability of Detection $\beta$ . . . . .  | 61 |
| 4.7  | Long Run Average Welfare in the Empty Network . . . . .   | 67 |

# CHAPTER 1

## INTRODUCTION

Social connections pervade our lives; people find jobs from their friends and family, learn about political candidates from their peers, buy and sell from people they know. The structure of these networks have long been understood to have a large effect in each of these situations; workers with different social networks may have very different labor market outcomes; who your peers are may have a large effect on your political opinions, and markets may be very far from competitive, if buyers and sellers are not connected to one another. Because of the importance social networks may have, the structure of social networks has long been an important area of study.

How are these networks formed? In this dissertation I explore the source of structures in social networks, and the impact this structure may have in economic settings. In chapter 2, I develop a model of Bayesian strategic network formation, in which agent heterogeneity is modeled as a Bayesian game. In this way I am able to characterize equilibrium networks even with heterogeneity, and I find an intimate connection between the two extremes of the set of equilibria, and two different models of public good provision.

In chapter 3, I further explore the impact these extreme structures have in a setting where agents share a public good along their social ties. Different network structures provide different incentives for public good provision, with the incentive to produce the public good being strongest in networks with the least coordination in linking; that is, efficient network formation leads to inefficient public good provision. I characterize the set of equilibria, and show that only for the least coordinated networks does specialization in roles arise, where some agents form links, while others produce the public good, and no agent does both.

In chapter 4, I consider the impact network structures in the labor market have for time allocation and labor market participation in the formal and informal sectors. Characterizing the long run steady state properties of the model, we show that income may not be monotonic in labor market law enforcement; less enforcement leads to more time allocation to informal work, but at the same time allows more workers without informal jobs to expect to find

them, and so allocate more time to leisure. So, even as utility is increasing, income may be decreasing.

## CHAPTER 2

# HETEROGENEITY IN SOCIAL NETWORKS AND THE CLASSICAL THEORY OF PUBLIC GOODS

### 2.1 Introduction

There are many situations and contexts in which the detailed patterns of local interaction and the structure of interpersonal relationships will influence economic outcomes. Such networks often serve as a conduit for information, and transmit job information, fashion preferences, or even disease from one person to another (Granovetter [37], Jackson [43]). In other settings, networks are viewed directly as an asset, as in models of social insurance; in developing countries where people do not have access to formal insurance markets, their social contacts provide a substitute (Fafchamps and Lund [26]). Networks are often an important structural feature of an economic situation, such as markets where buyers and sellers must interact directly (Kranton and Minehart [51]), or markets with network constraints, such as electricity and airline networks (Hendricks, Piccione and Tan [39] and Cho [20]). This leads us to ask, how do agents choose to form these links?

I consider a setting where agents are heterogeneous in the value they derive from their connections to one another, and have incomplete information on this heterogeneity among their peers. Agents form links non-cooperatively, and receive utility for every other agent they are connected to in the resulting network, whether directly or indirectly. This value is private information, independently and identically distributed across agents. For every link they choose to form, they bear a cost. Given her personal value, agents must choose whether and with whom to form connections. The kind of heterogeneity in this model is one where each agent has an unknown, different value to every other agent; one agent does not know, *ex ante*, whether or not the information she has is of value to other agents. Because she does not know if others will form links and connect the network for her, she must weigh her own personal value of the information of others against the likely actions of the other agents.

An example of a setting this model describes is job contacts; that is, the pattern of contacts among acquaintances searching for jobs. Here one agent may contact another without her

consent; the assumption that after doing so the contacted agents value (and the value of all *her* contact agents) is available to the contacter amounts to assuming that the information those agents may have that is of value to the contacter is costlessly collected and transmitted to the contacter, without any incentive or “holdup” problems.

I characterize the set of equilibria for three agents, and identify the extremes of this set. On one extreme, agents form a link to a single central agent, while on the other, they form links to every other agent. Other equilibria are essentially combinations of these two extremes; one set of agents, the “insiders”, form many links with one another, while agents in another set, the “outsiders.”, form a single link to an insider. The structures that arise in equilibrium are “core” and “periphery” networks; the insiders form a dense core to the network, with many links to one another, while the outsiders form a periphery, with at most a single link to an insider. Network structures form stochastically, contingent on the private values of each agent; star and circle structures may still arise, but they are only two of many possible outcomes.

These extreme equilibria have a natural interpretation in terms of public good provision: When agents form links to a single insider, the equilibrium is equivalent to a cost sharing provision rule of an excludable club good. On the other extreme, where every agent forms links to everyone, the equilibrium is equivalent to the voluntary provision of a pure public good. Intermediate cases of this family of equilibria are essentially combinations of these two extremes. The strategic nature of each agent’s linking decision depends on her position in the network; link formation reproduces both pure public good provision and club good provision as special cases.

This highlights the tension between the private benefits and public externalities of social connections. Insiders wish to avoid being left unconnected when their peers free-ride. In equilibrium, this leads them to form superfluous links to one another, creating a densely connected “core” to the network. In contrast, outsiders coordinate their links and access many peers with a single link to the core; they do not free ride. Both of these incentives lead to centralized network structures, in which the path connecting any two agents is short. Clustering, the tendency of connected agents to have peers in common, arises endogenously among the insiders. This equilibrium of the network formation game thus produces two important features of observed social networks, and both of these structures are present in the same equilibrium.

This paper is a contribution to the literature of network formation. LIT The main contribution is a model of strategic network formation with incomplete information. The connections

model was first introduced by Jackson and Wolinsky [45], and studied in non-cooperative form by Bala and Goyal [8]. They focused on symmetric agents and strong solution concepts, which has the virtue of being analytically tractable. The large number of actions and outcomes is a constant problem in network theory. It is unsatisfying, however, to be unable to analyze agents that differ in their social abilities or have less than perfect information about their peers. Moreover, the structures predicted by the symmetric model are implausibly stark; circles and stars the typical outcomes. This is very different from the complicated networks that appear in the data, and there is little to suggest how agents are sorted into their network positions.<sup>1</sup>

Galeotti, Goyal and Kamphorst [66] also incorporate heterogeneity into the connections model, retaining the solution concepts of Bala and Goyal [8]. They allow for heterogeneity in the values and costs of linking, and obtain an “anything goes” result: Any minimal network is an equilibrium for some profile of values and costs. The flexibility of their model thus comes at the cost of strong predictions.

The role incomplete information may play in network formation has not been extensively considered; McBride [55] studies an environment where agents have incomplete information on the structure of the network: When agents are aware only of the local structure of social ties, equilibrium networks will be very different that when agents have full information, and will be inefficient. This paper, in contrast, considers only incomplete information on the types of one’s peers, so that conditional on types and strategies, there is complete information on the structure of the network. Complementary

The second contribution is a characterization of this model for the case of three agents, and the identification of a class of equilibria in this case that is extended to case of  $n$  agents. Core-periphery structures, as arise in my model, have been found to be important in many settings, from economic geography (Krugman [52]) and corporate structure (Mintz and Schwartz [56], Mizuchi [57]), to friendship (Adamic and Adar [1]) and academic citations (Borgatti [14]). Other works study the different forces that lead to such centralized structures. Hojman and Szeidl [40] consider an environment close to that of Bala and Goyal [8], with stronger conditions on the decay of value across links, and diminishing returns to connections. In their complete information environment, stars are the unique equilibrium. My model shows that such outcomes can arise from self-organization among ex ante identical agents.

The third contribution is a characterization of this class of equilibria in terms of public good provision. The externalities present in link formation have been recognized by network

---

<sup>1</sup>For a survey on data and methods in the economics of networks, see Jackson [44].

theorists; Bramoullé and Kranton consider a model of public good provision on networks, and the importance of the structure of the network for incentives in public good provision is key. This paper shows that not only are the links a kind of public good, but the kind of public good it is may depend on the entirety of the network, on the structure of the network itself. Goyal and Galeotti [31] endogenize link formation in a similar model, and find that many different structures may be equilibria, with many possible configurations of investment and link formation. By contrast, this paper considers only link formation, and the network itself is the public good being provided.

## 2.2 The Model

### 2.2.1 Non-cooperative Network Formation

Let  $N = \{1, \dots, n\}$  be the set of agents, with  $n \geq 3$ . Each agent  $i \in N$  values being connected to the other agents. To receive this value, she may initiate links to them. An action of an agent  $i$  is an element  $a_i \in A_i = \{0, 1\}^n$ , where  $a_{ij} = 1$  indicates that agent  $i$  initiates a link to agent  $j$ , and I adopt the convention that  $a_{ii} \equiv 0$ .<sup>2</sup> Agents choose how many links to initiate simultaneously. Two agents  $i$  and  $j$  are connected in a graph  $g$ , written  $g_{ij} = 1$ , if either  $i$  or  $j$  initiates a link to the other. Therefore,  $g_{ij} = \max(a_{ji}, a_{ij})$ . I let  $g(a)$  represent the graph that forms given the profile of actions  $a$ . Value flows both ways across a link in the network; an agent need not initiate any links to receive positive utility, if the links of others connect her to the network. In the terminology of Bala and Goyal [8], this is a model of two-way flow of value.

For example, with 3 agents, if  $a_{12} = 1$ ,  $a_{23} = 1$ ,  $a_{32} = 1$ , and  $a_{31} = 1$ , the graph that forms is  $g = \{12, 13, 23\}$ , the complete graph (Figure 2.1). Because the flow of value is two-way, I need not specify a direction for the links in the graph. Given a graph  $g$ , we can calculate the number of agents to whom agent  $i$  has a path in  $g$ . There is a path from  $i$  to  $j$  in network  $g$  if there is a sequence of distinct agents  $\{j_1, \dots, j_m\}$  such that  $g_{ij_1} = g_{j_1j_2} = \dots = g_{j_mj} = 1$ . Let  $N_i(g)$  be the number of agents to whom  $i$  has a path in  $g$ , and let  $\mu(a_i) = \sum_j a_{ij}$  be the number of links that  $i$  initiates.

Agents are heterogeneous in the benefits they receive from each agent to whom they have a path in the graph  $g$ . Each agent  $i$  has a type  $\theta_i$  that represents these benefits; she receives

---

<sup>2</sup>This reflects the convention that agent  $i$  receives no benefit from a connection to herself.

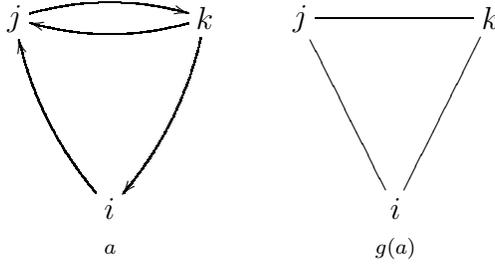


Figure 2.1: A profile of actions and the resulting graph.

$\theta_i$  for each agent to whom she has a path in  $g$ .<sup>3</sup> That is, agents differ in how they value others, but each agent values all others equally.<sup>4</sup> An agent bears a cost  $c > 0$  for each link she initiates. This cost is the same for all agents.<sup>5</sup> Agent's types are independently and identically distributed according to a distribution  $F$ , which has support  $[0, 1]$ . Each agent's type  $\theta_i$  is private information, known only to herself.

The timing of the game is as follows: (i) each agent observes her type  $\theta_i$ ; (ii) each agent simultaneously chooses which links to initiate, in ignorance of the types of the other agents; (iii) the profile of initiated links defines a graph  $g$  from which agents receive benefits, and experience the cost of the links they initiated. The utility she receives from a profile of actions  $a$  is therefore

$$U(\theta_i, a_i, a_{-i}) = N_i(g(a_i, a_{-i})) \theta_i - \mu(a_i) c.$$

A strategy is a mapping  $s_i : [0, 1] \rightarrow A_i$  that specifies an action in  $A_i$  for each type  $\theta_i$  of agent  $i$ . Given that an agent forms links knowing her own type, but in ignorance of anyone else's, the natural solution concept is *Bayesian Nash Equilibrium*. Given her own type and beliefs about the types and strategies of others, each agent chooses her links to maximize her expected utility. A profile of strategies  $s^*$  is a Bayesian Nash equilibrium if for all  $a' \in A_i$ ,  $\theta_i \in [0, 1]$ , and for all  $i$ ,

$$\mathbb{E}_{\theta_{-i}}[U(s_i^*(\theta_i), s_{-i}^*(\theta_{-i}), \theta_i) | \theta_i] \geq \mathbb{E}_{\theta_{-i}}[U(a', s_{-i}^*(\theta_{-i}), \theta_i) | \theta_i].$$

---

<sup>3</sup>Unlike other treatments of the connections model, I omit any notion of *decay*, where the value an agent derives from a path to another agent falls with the length of that path. The inclusion of small amounts of decay would have little effect on equilibrium structures, because in equilibrium paths are short. Large decay would significantly affect the results, and are beyond the scope of this paper.

<sup>4</sup>This contrasts with the case in which each agent had a distinct value for each agent she could connect to, or other, more general kinds of heterogeneity.

<sup>5</sup>Because expected utility is linear in values and costs, this model is equivalent to one where the value agents receive is a parameter, and agents are heterogeneous in the costs of linking.

Agents only care about (1) how many other agents to whom they have a path in the graph,  $N_i(g(a))$ , and (2) how many links they initiate,  $\mu_i(a_i)$ . Because of this, the sum of all agent's utilities may be the same for many different graphs. This simplifies efficiency considerations. A network is *efficient* if it maximizes the sum of all agents utilities given the realized profile of types  $\theta$ . Because each agent provides utility to other agents in the network, it is possible that an agent's utility is negative in an efficient network. This occurs when the value other agents receive from her presence is high enough that she should be connected to the network, but her own value from the network is less than the cost of any links she forms. Since it does not matter for efficiency who pays for the links to connect the network, a network where some agents have negative utility may still be efficient.

If it is efficient for the entire set of agents  $N$  to form a nonempty graph, the graph must be *minimal*. It must have exactly  $n - 1$  links, so that deleting any link would disconnect the graph into disjoint components. This is because there is no value to having multiple paths to other agents, and redundant links are costly. Furthermore, if any nonempty graph has positive value, an efficient graph must be *connected*: Every agent must have a path to every other agent. This characterizes efficient graphs.

**Proposition 1** *Any minimal connected graph is efficient if and only if  $\sum_i \theta_i \geq c$ . The empty graph is efficient if and only if  $\sum_i \theta_i \leq c$ .*

Several important architectures are both minimal and connected, and therefore efficient. A *center-sponsored star* sponsored by agent  $j$ , is a graph in which  $a_{jk} = 1 \forall k \neq j$ , and  $a_{kl} = 0 \forall k \neq j, \forall l$ . Agent  $j$  initiates links to all other agents, and no other links are initiated. Agent  $j$  is called the center of the star. A *periphery-sponsored star* centered on agent  $j$ , is a graph in which  $a_{kj} = 1 \forall k, k \neq j$ , and  $a_{kl} = 0 \forall k, \forall l \neq j$ , and  $a_{jl} = 0 \forall l$ . In this graph, all other agents form a single link to agent  $j$ , who is again the center of the star, and no other links are formed. Both action profiles give rise to the same graph  $g$ , a star. Both of these networks are minimal and connected. In fact, in any star the maximum length of a path from one agent to another is two.<sup>6</sup>

Two related architectures will be important. (1) A *redundant center-sponsored star* differs from the center-sponsored star in that there is more than one agent in the center who initiates links to every other agent. These networks are not minimal: Each agent in the center initiates links to everyone, so there are more links than are required to connect the network. (2) An

---

<sup>6</sup>Many other familiar architectures are also efficient. A line, in which every agent but one initiates a single link, and every agent but one receives a single link, is minimal and connected. In fact, every connected *tree*, a graph without loops that contains every agent, is efficient.

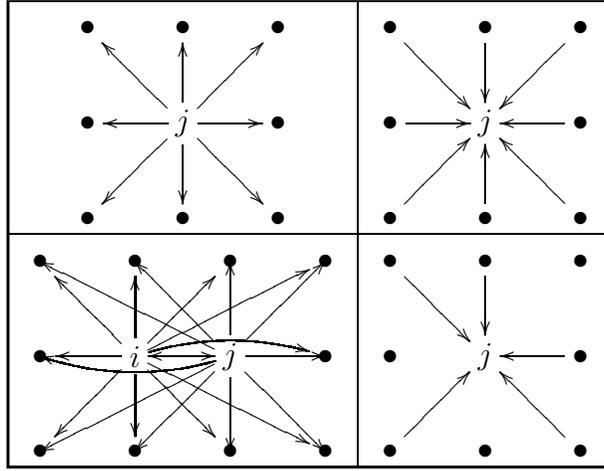


Figure 2.2: Clockwise from the upper left, a center-sponsored star, a periphery-sponsored star, an incomplete periphery-sponsored star, and a redundant center-sponsored star.

*incomplete* periphery-sponsored star differs from the periphery-sponsored star in that not every agent forms a link to the center agent: Some agents are left unconnected, and the periphery is incomplete. This network is not connected, and therefore is inefficient. These architectures are depicted in figure 2.2.

### 2.2.2 Public Good Provision

To make explicit the connection between network formation and public good provision, I analyze two well known models of non-cooperative public good provision. The same set of agents  $N$  may choose to provide a public good  $Y \in \{0, 1\}$ . This public good is binary and indivisible, and costs  $K$  to provide. I consider both pure public good goods, where the benefits of the good are nonrival and nonexcludable, and excludable public goods without congestion, where the benefits of the good are nonrival but excludable. Such a public good is commonly called a *club good*. Agent  $i$  receives value  $\theta_i$  from the good, and this value is private information. It is drawn from  $[0, 1]$ , according to  $F$ , just as in the link formation model. This value will enter utility in a different way in each of the two models of public good provision I consider.

**Voluntary Contribution to a Pure Public Good without Refunds.** In the *contribution game*, the public good is nonrival and nonexcludable. Each agent may choose to make some payment  $b_i \in \mathbb{R}_+$  to fund the good. The good is provided only if the sum of contributions exceed the cost of provision,  $K$ . If the sum is less than  $K$ , contributions are

not returned. If the sum exceeds  $K$ , excess contributions are not returned.<sup>7</sup> Agent  $i$ 's utility is

$$U_i(b_i, b_{-i}, \theta_i) = \begin{cases} \theta_i - b_i, & \sum b_j \geq K; \\ -b_i, & \sum b_j < K. \end{cases}$$

If the public good is provided, everyone benefits. A strategy is a function  $t : [0, 1] \rightarrow \mathbb{R}_+$  that specifies a contribution in  $\mathbb{R}_+$  for every type  $\theta_i$  of agent  $i$ . A profile of strategies  $t^*$  is a Bayesian Nash Equilibrium if for all  $b' \in \mathbb{R}_+$ ,  $\theta_i \in [0, 1]$ , and all  $i$ ,

$$\mathbb{E}_{\theta_{-i}}[U(t_i^*(\theta_i), t_{-i}^*(\theta_{-i}), \theta_i) | \theta_i] \geq \mathbb{E}_{\theta_{-i}}[U(b', t_{-i}^*(\theta_{-i}), \theta_i) | \theta_i].$$

**The Voluntary Cost Sharing Mechanism of an Excludable Public Good.** The second model of project provision that I consider is the *voluntary cost sharing mechanism*. The public good is now assumed to be nonrival but excludable. An agent who does not pay to fund the mechanism is excluded from its benefits. It is still binary and indivisible. Actions are now  $d_i \in \{0, 1\}$ . An agent who chooses  $d_i = 1$  indicates that she is willing to share the cost  $K$  of providing the public good; this cost is divided equally among all agents who signal 1. An agent who signals 0 pays nothing, and is excluded. This is a model of a club good, where agents must pay to join the club, and the membership benefits are nonrival. Agent  $i$ 's utility is

$$U_i(d_i, d_{-i}, \theta_i) = \begin{cases} \theta_i - \frac{K}{\sum_j d_j}, & d_i = 1; \\ 0, & d_i = 0. \end{cases}$$

A strategy is a function  $r : [0, 1] \rightarrow \{0, 1\}$  that specifies an action in  $\{0, 1\}$  for every type of agent  $i$ . A profile of strategies  $r^*$  is a Bayesian Nash Equilibrium if for all  $d' \in \{0, 1\}$ ,  $\theta_i \in [0, 1]$ , and all  $i$ ,

$$\mathbb{E}_{\theta_{-i}}[U(r_i^*(\theta_i), r_{-i}^*(\theta_{-i}), \theta_i) | \theta_i] \geq \mathbb{E}_{\theta_{-i}}[U(d', r_{-i}^*(\theta_{-i}), \theta_i) | \theta_i].$$

Efficiency in the provision of public goods is characterized simply: If the sum of all agents' values for the good exceeds the cost of provision, then it is efficient to provide the good. This holds for both the contribution game and the voluntary cost sharing mechanism. In the contribution game, efficiency in cost allocation requires that no agent over-contribute to the good when it is provided: Excess contributions above the minimum required to fund the good are wasted. If it is efficient not to provide the good, agents must make no contributions

---

<sup>7</sup>The lack of refunds is not essential to the result. The same equilibrium exists in the model with and without refunds, but notation is needlessly complicated with refunds.

at all: Positive contributions are wasted. The voluntary cost sharing mechanism does not have this problem: By construction, agents cannot over-contribute to funding the good, and if the good is provided, some agent has agreed to bear the cost, so agents cannot under-contribute either. Efficiency only requires that if the sum of all agents' values exceeds the cost of provision, then at least one agent must signal a willingness to pay, and every agent must be included in its benefits. Because there is no congestion or rivalry in the utilization of the public good, it is never efficient to exclude anyone in the voluntary cost sharing mechanism. It may be that an agent would prefer not to participate in the voluntary cost sharing mechanism, because her benefit  $\theta$  may be less than her share of the cost of provision. Efficiency still requires that she be included, however, because the total cost of provision does not depend on the number of agents included, and including her increases total welfare. I summarize with the following proposition.

**Proposition 2** *If  $\sum \theta_i \geq K$ , it is efficient to undertake the public good, and if  $\sum \theta_i < K$  it is efficient not to. If the good is provided, efficiency in the contribution game requires that  $\sum b_i = K$ , and efficiency in the cost sharing mechanism requires that all agents be included in the project.*

## 2.3 The Three Agent Case

It is clear that the link formation model may have many different equilibria. To illustrate the possibilities, I exhaustively characterize all equilibria of link formation model for the case of three agents, when types are drawn from the uniform distribution on  $[0, 1]$ . This simple case still has many equilibria, including symmetric and asymmetric equilibria, involving pure and mixed strategies.<sup>8</sup> These equilibria vary both in the willingness of agents to form links, as a function of their private value for connections, and the architecture these links will be arranged in, if initiated.

**Proposition 3** *There are exactly 5 classes of equilibrium strategies for the three agent case with uniform values, which the following architectures:*

1. *A circle, with two possible orientations.*

---

<sup>8</sup>This is not a symmetric game in the strictest sense, because each agent has a different set of feasible actions; agent  $i$  can link to agent  $k$ , but agent  $k$  cannot link to herself. The meaning of symmetric is just that each agent face the same decision problem, and uses the same cutoff rule in equilibrium.

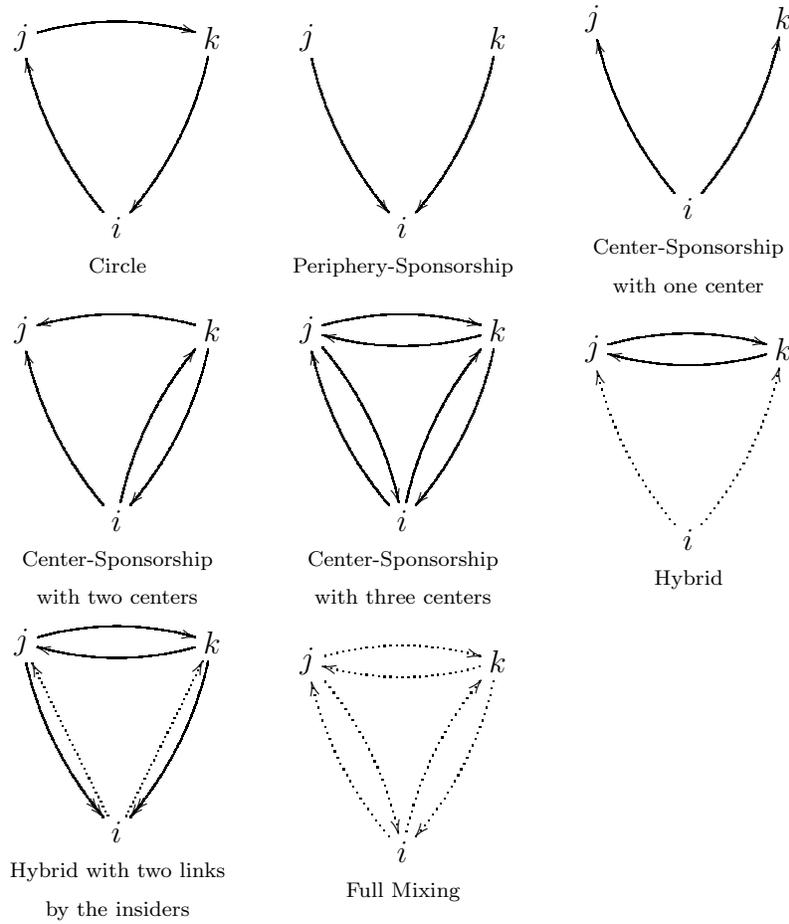


Figure 2.3: Equilibrium strategies in the three agent case. Solid lines indicate pure strategies in link formation, dashed lines indicate mixed strategies.

2. A periphery-sponsored star, with agent 1, 2 or 3 in the center.
3. A center-sponsored star, with one, two or three agents in the center.
4. A hybrid structure, where two agents link each other and the third links mixes between forming a link to either of them. There are two such equilibria, one where the agents linking each other also form a link to the third agent if their value is high enough, and one where they do not.
5. A full mixing structure, where each agent mixes in forming a single link to one of the other agents.

Figure 2.3 depicts these strategies.<sup>9</sup> In the circle equilibrium, each agent forms a single link to another, forming a circle. This is only an equilibrium when the number of agents is low; when there are more agents, the probability of having a path to an agent on the far side of the circle is low, and agents would prefer to deviate and ‘link across’ the circle. With three agents, there is no need, because each agent is only one link away from every other agent.

In the center-sponsored star, agents choose either to link to every other agent, or not link at all. This is an equilibrium for one, two or three agents in the center, choosing to form links. If only two agents choose to form links, for example, the third agent is passive, and the others form links with higher probability. In the periphery-sponsored star, two agents choose whether or not to link to the third agent, who forms no links. This is an equilibrium for each agent  $i$ ,  $j$ , or  $k$  in the center.

In the hybrid equilibrium, two agents form a single link to one another, while the third agent mixes in forming a single link to either of them. There are two such equilibria, one where the non-mixing agents also forms a link to the mixing agent if their value is sufficiently high, and one where they do not. In the full mixing structure, each agent mixes in forming a single link to one of the other two agents.

While this is a large number of seemingly diverse equilibria, there is a pattern among them. We can view the three agent center-sponsored star, the periphery-sponsored stars, and the hybrid equilibria without second links as members of a family of equilibria. To see this, split the set of agents into one nonempty group  $I \subseteq N$  of insiders, and  $O = N/I$  of outsiders. The following is an equilibrium strategy profile. For agents  $i \in I$ , let

$$s_I(\theta) = \begin{cases} a_{ik} = 0 \ \forall k, & \theta < \theta_I; \\ a_{ik} = 1 \text{ for } k \in I, a_{ik} = 0 \text{ else,} & \theta > \theta_I. \end{cases}$$

That is, agents in  $I$  will link to every other agent in  $I$  if their value is above a cutoff  $\theta_I$ , and will otherwise form no links. For agents  $i \in O$ , let

$$s_O(\theta) = \begin{cases} a_{ik} = 0 \ \forall k, & \theta < \theta_O; \\ a_{ik} = 1 \text{ with probability } \frac{1}{|I|} \ \forall k \in I, a_{ik} = 0 \ \forall k \notin I & \theta > \theta_O. \end{cases}$$

This strategy calls for agents in  $O$  to link to a single agent in  $I$  if their value exceeds  $\theta_O$ , with equal probability for each. If their value is below  $\theta_O$ , they form no links. This profile of strategies is an equilibrium for any partition of the set for agents with at least one agent

---

<sup>9</sup>That is, these figures describe strategies, *not* realized networks.

in  $I$ .

The equilibrium architecture of this strategy profile is a set of insiders, who are connected to each other in a redundant center-sponsored star, and a set of outsiders, each of whom forms a single link to one of the insiders. This is the insiders-outsiders family of equilibria. The three-agent center-sponsored star is the special case of the insider-outsider equilibrium where  $I = N$ , and every agent is an insider. The periphery-sponsored star is the case where  $I = \{k\}$ , for a particular agent  $k$ . Here, everyone but  $k$  is an outsider. The hybrid is the case where  $I = \{j, k\}$ , two insiders and one outsider. The center-sponsored stars with fewer than 3 agents are not special cases of the insider-outsider equilibrium, nor is the hybrid equilibrium with second links or the mixing equilibria.

The three agent center-sponsored star and the periphery-sponsored star form two extremes of the set of equilibria, in the following sense: They represent the most and the least free-riding by agents who form links. All other equilibria are intermediate cases. The extent to which an agent may free-ride is determined by whether or not other agents will connect her to the rest of the network, without any action on her part. In the circle equilibrium, there is one agent who may do so. For agents forming links in the one or two-agent center-sponsored equilibrium, there are zero or one agent who may do so, respectively.

In the three-agent center-sponsorship equilibrium, each of an agent's peers may connect her to the network, so the potential for free-riding is greatest. In the periphery-sponsorship equilibrium, neither of an agent's peers will connect her to the network, so there is no potential for free-riding.

## 2.4 Insiders and Outsiders for $n > 3$

I focus only on the insiders-outsiders family of equilibria for larger populations, for several reasons. First, the predicted architecture has a core-periphery structure, and this architecture in many studies of social networks, from the structure of corporate directorates (Mintz and Shwartz [56], Mizruchi [57]) and friendship (Adamic and Adar [1]) to labor market contacts (Cross et al. [62]). Indeed, the structure of the equilibrium networks is exactly the idealized core-periphery graph as defined by Borgatti and Everett [14], who find empirical support for this classification in the structure of citation networks (Baker [7]).

Secondly, while stars are a common equilibrium architecture of many network formation games, they are rarely played in laboratory settings (Falk and Kosfeld [27], Kosfeld [50]).

Goeree et. al. [34] find that changes in the informational structure of the game and private information lead to more star and "star-like" structures in experiments, in a model similar to that studied here.

Finally, this family contains the two extremes of the set of equilibria, the center-sponsorship equilibrium and the periphery sponsorship equilibrium. It is in these two equilibria that the public good nature of network formation is most evident, and these two equilibria illustrate the competing incentives present in the link formation game.

The insiders-outsiders profile of strategies is an equilibrium for general distributions of value  $F$ , and arbitrary numbers of agents. I now characterize the general case. Let  $I \subseteq N$  be the set of insiders, which can have size  $|I| = 1, \dots, n$ . Define  $\theta_I$  and  $\theta_O$  implicitly by

$$F(\theta_I)^{|I|-1} \left( \frac{n-|I|}{|I|} (1-F(\theta_O))\theta_I \right) + (1-F(\theta_I))^{|I|-1} ((|I|-1)\theta_I + (n-|I|)(1-F(\theta_O))\theta_I) = (|I|-1)(\theta_I - c) + (n-|I|)(1-F(\theta_O))\theta_I, \quad (2.1)$$

$$F(\theta_I)^{|I|} \left( \theta_O + \frac{n-|I|-1}{|I|} (1-F(\theta_O))\theta_O \right) + (1-F(\theta_I))^{|I|} (|I|\theta_O + (n-|I|-1)(1-F(\theta_O))\theta_O) - c = 0. \quad (2.2)$$

$\theta_I$  and  $\theta_O$  are the critical type of insiders and outsiders, respectively, that are indifferent between forming and not forming a link. That is, these conditions equate the equilibrium benefit of a link to the cost of that link. There is one condition for insiders, and one for outsiders. These critical types determine the cutoffs agents use in equilibrium.

**Proposition 4 (Insider-Outsider Equilibrium)** *For  $i \in I$  the strategy  $s_I(\theta)$  given by*

$$s_I(\theta) = \begin{cases} 0^{n-1}, & \theta < \theta_I; \\ a_{ik} = 1 \text{ for } k \in I, a_{ik} = 0 \text{ else,} & \theta > \theta_I. \end{cases}$$

*and for  $i \notin I$ , the strategy  $s_O(\theta)$  given by*

$$s_O(\theta) = \begin{cases} 0^{n-1}, & \theta < \theta_O; \\ a_{ik} = 1 \forall k \in I, a_{ik} = 0 \forall k \notin I, \text{ with probability } \frac{1}{|I|}. & \theta > \theta_O, \end{cases}$$

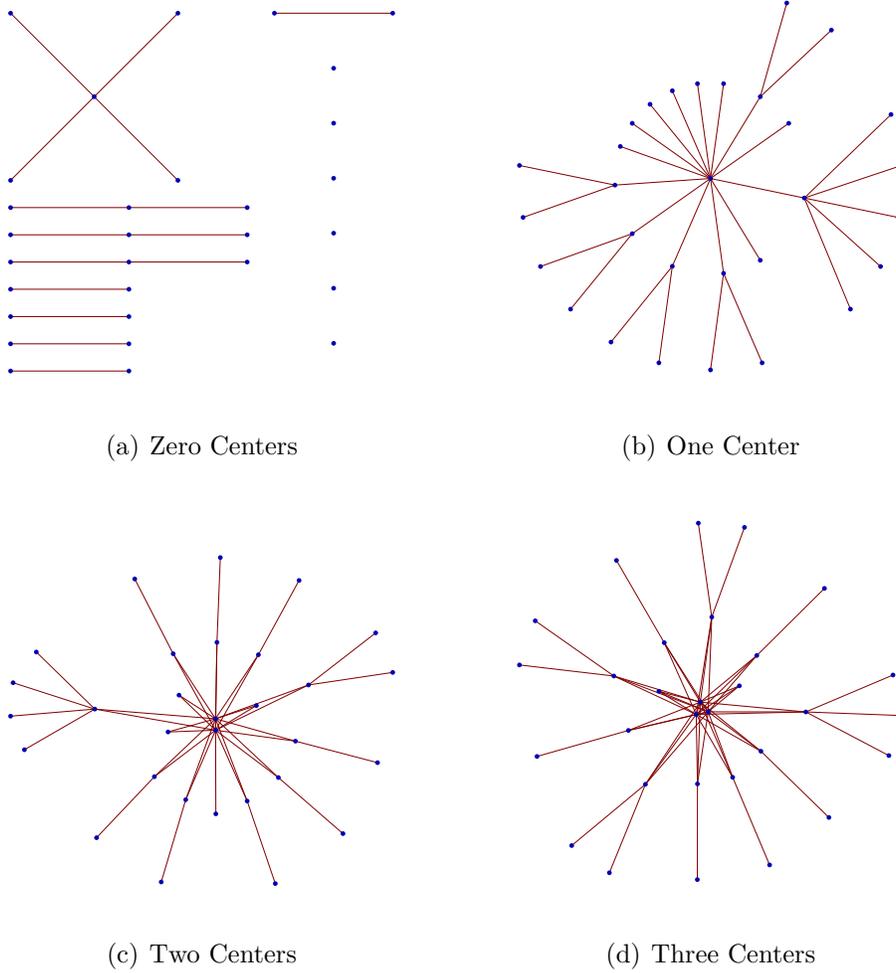


Figure 2.4: Four possible realized networks for Uniform values,  $c = .5$ , with 15 insiders and 15 outsiders.

is a Bayesian Nash equilibrium of the network formation game, if

$$c > F(\theta_O) + (1 - F(\theta_O)F(\theta_I))^{|I|-1} \left( 2 + \frac{n - |I| - 1}{|I|} (1 - F(\theta_O)) \right).$$

This is an equilibrium for any  $n$ . The range of  $c$  is restricted to ensure that an equilibrium exists. Links cannot be so inexpensive that insiders wish to deviate and form direct links to outsiders, and outsiders must not wish to form direct links to other outsiders, or multiple links to insiders. In addition, the cost of links cannot be so high that no one wishes to link to anyone. These restrictions are satisfied by many different distributions  $F$ .

**Example 1 (Uniform Values)** When value are distributed uniformly on  $[0, 1]$ ,  $F(\theta) = \theta$ ,

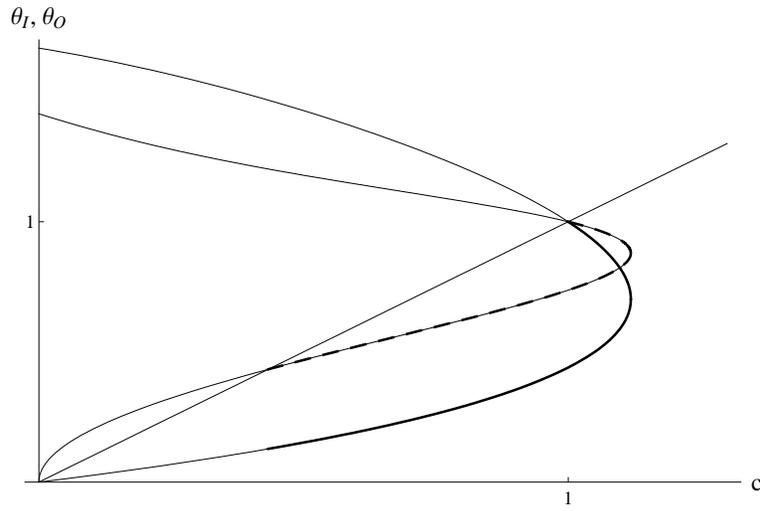


Figure 2.5: The insiders-outsiders equilibrium for Uniform values, , for  $n = 5$ ,  $I = 2$ .  $\theta_I$  is given by the dashed line,  $\theta_O$  by the solid line. The bolded portions represent the range of  $c$  over which this is an equilibrium.

the cutoffs  $\theta_I$  and  $\theta_O$  solve

$$\theta_I^{|I|-1} \left( \frac{n-|I|}{|I|} (1-\theta_O)\theta_I \right) + (1-\theta_I^{|I|-1}) ( (|I|-1)\theta_I + (n-|I|)(1-\theta_O)\theta_I ) = (|I|-1)(\theta_I - c) + (n-|I|)(1-\theta_O)\theta_I$$

and

$$\theta_I^{|I|} \left( \theta_O + \frac{n-|I|-1}{|I|} (1-\theta_O)\theta_O \right) + (1-\theta_I^{|I|}) ( |I|\theta_O + (n-|I|-1)(1-\theta_O)\theta_O ) - c = 0.$$

Figure 2.5 depicts  $\theta_I$  and  $\theta_O$  as a function of  $c$  for  $n = 5$  and  $I = 2$ . Four possible realized networks are depicted in 2.4.

There are multiple solutions to equations (2.1) and (2.2) because outsiders' actions are complements: The more likely outsiders are to link, the higher is the expected utility from such a link. For uniform values, there are two different equilibrium cutoffs for outsiders, corresponding to a high and low link likelihoods. These two cutoffs are both equilibria for  $c > 1$ ; for  $c < 1$  the high cutoff is not an equilibrium, because links are too cheap, and agents would deviate and form more. Multiple cutoffs for outsiders leads to multiple cutoffs for insiders: When outsiders are highly likely to link, insiders have a higher incentive to free ride, and so a different cutoff than when outsiders are less likely to link. This model thus delivers a natural feature of social networks. The more friends someone has, the more

valuable she is as a friend, and the more people wish to be her friend. In contrast, an agent with few friends is less attractive as a friend; this phenomenon arises as possible equilibria among ex ante identical agents, typical in coordination games.

The conclusions that Galeotti, Goyal and Kamphorst [66] draw from their insiders-outsiders model arise endogenously here. They partition the set of agents into two groups, who have a low cost when linking within their own group, and a higher cost when linking across groups. This ex ante asymmetry between agents induces an asymmetry in the equilibrium structures similar to the structures I find. First, they find that one group will form a core that is internally connected, while the other group only links to the core, and not to each other. Second, they find that this structure holds generally, regardless of the number of agents, so that even a large network has a low diameter. Third, they find that centrality is important: All links are oriented towards some central player. Each of these properties arises in this model, but I do not impose the restriction that agents are exogenously partitioned into groups with different linking costs. In my model, the asymmetries arise endogenously among ex ante symmetric agents. The partitioning of agents into groups is an *equilibrium* phenomenon that is self-supporting.

Insiders and outsiders have different views of the strategic nature of their decision. To make these differences explicit, I consider the two extremes of the insider-outsider equilibrium; (i) the case where every agent is an insider and (ii) the case where every agent but one is an outsider. In the first case, we have the following:

**Corollary 1 (Center-Sponsored Equilibrium)** *Suppose that  $F(c) < 1$ . Then there exists  $\theta_I$  defined implicitly by*

$$\theta F(\theta)^{n-1} = c. \tag{2.3}$$

Let

$$s_I(\theta_i) = \begin{cases} \{0\}^{n-1}, & \text{if } \theta_i \leq \theta_I; \\ \{1\}^{n-1} & \theta_i > \theta_I \end{cases}$$

*Then for  $c \in [0, 1]$ ,  $\{s_I\}$  is a Bayesian Nash equilibrium of the network formation game.*

Here, every agent is an insider,  $I = N$ . Each agent chooses to either link to every other agent, or to no one. There is only a single solution to the the insider cutoff condition, which in this case reduces to equation (2.3). This is because links are not complements for insiders,

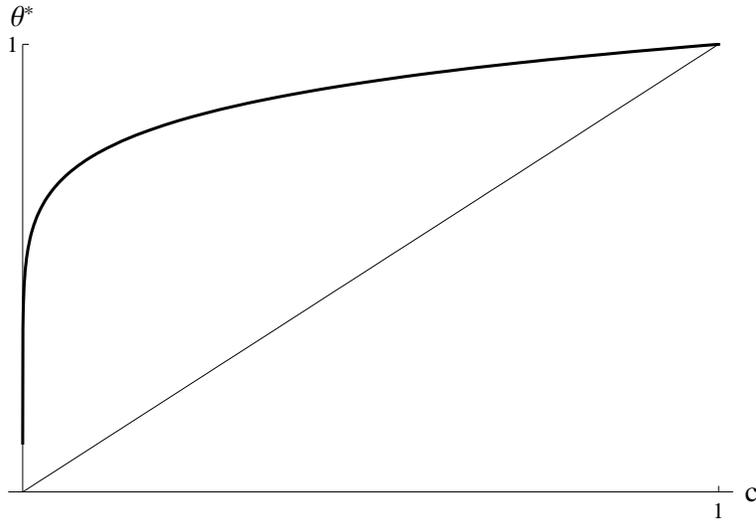


Figure 2.6: Center-sponsored equilibrium with uniform values, in bold.

only for outsiders. For insiders, the more likely others are to link, the less the expected utility from linking is; their actions are strategic substitutes. This leads to a unique equilibrium cutoff.

**Example 2 (Uniform Values)** For the case of uniform values,  $F(\theta) = \theta$ , the cutoff  $\theta_I$  solves

$$\theta_I(\theta_I)^{n-1} = c,$$

so that  $\theta_I = c^{\frac{1}{n}}$ .

Figure 2.6 the equilibrium cutoff when values are distributed uniformly on  $[0, 1]$  is depicted in. If an agent's type  $\theta_i$  is above the depicted curve, she is above the cutoff and forms a link to every other agent. If her type is below the curve, she is below the cutoff, and forms no links. Note that it is possible for an agent's value to be above  $c$ , the cost of linking, but below the equilibrium cutoff. That is, it may be that agents can profitably link to another, but choose not to—they free-ride. It is possible that in equilibrium, every agent's value is above  $c$  but below  $\theta_I$ , so that the network is empty. It is also possible that more than one agent's value is above  $\theta_I$ , so that multiple agents form links to every other agent. Over-connection and under-connection are both possible in equilibrium, and the network generally takes the form of a redundant center-sponsored star. The more agents whose value exceeds  $\theta_I$ , the more centers in the star, and the more superfluous links.

On the other extreme, when  $I = \{k\}$ , everyone but agent  $k$  is an outsider.

**Corollary 2 (Periphery-Sponsored Equilibrium)** Define  $\theta_O^{n-1}$  and  $\theta_m^{n-1}$  implicitly by

$$\theta \left( 1 + (n-2)(1 - F(\theta)) \right) = c, \quad (2.4)$$

$$\theta = \frac{1 + (n-2)(1 - F(\theta))}{(n-2)f(\theta)}.$$

For a selected agent  $k$ , consider the following strategy profile:

1. Agent  $k$  forms no links.
2. All other agents follow a cutoff strategy:

$$s_O(\theta_i) = \begin{cases} \{0\}^n, & \text{if } \theta_i < \theta_O^{n-1}; \\ (k), & \text{if } \theta_i > \theta_O^{n-1}. \end{cases}$$

The action  $(k)$  denotes a single link to agent  $k$ . That is, link to agent  $k$  if  $\theta_i > \theta_O^{n-1}$ , and otherwise form no links.

For  $\frac{(n-1)\theta_O^{n-1}}{(n-2)\theta_O^{n-1}+1} \leq c \leq \theta_m^{n-1} \left( 1 + (n-2)(1 - F(\theta_m^{n-1})) \right)$ , this strategy profile is a Bayesian Nash Equilibrium of the network formation game, for every  $k$ .

I index the cutoff  $\theta_O^{n-1}$  by  $n$ . This index will be important when I later compare this equilibrium to models of public good provision. In this equilibrium, agent  $k$  is a passive insider and forms no links, while every other agent chooses whether or not to link to her.<sup>10</sup> The upper restriction on  $c$  ensures that the cutoff exists, while the lower restriction eliminates deviations.

**Example 3 (Uniform Values)** For the case of uniform values,  $F(\theta) = \theta$ , the cutoff  $\theta_O$  solves

$$\theta(1 + (n-2)(1 - \theta)) = c. \quad (2.5)$$

Equation (2.5) has two solutions. There is a high cutoff solution and a low cutoff solution. Let  $\theta_O^H$  be the high cutoff, and  $\theta_O^L$  be the low cutoff. They are given by

---

<sup>10</sup>This is similar to the equilibrium structure of Hojman and Szeidl [40], but is driven by strategic behavior rather than decay in value.

$$\theta_O^L = \frac{(n-1) - \sqrt{(n-1)^2 - 4c(n-2)}}{2(n-2)} \quad (2.6)$$

$$\theta_O^H = \frac{(n-1) + \sqrt{(n-1)^2 - 4c(n-2)}}{2(n-2)}. \quad (2.7)$$

Figure 2.7 depicts these cutoffs. The lower and upper arcs of the parabola correspond to solutions to (2.6) and (2.7), respectively.

The ranges of  $c$  over which these solutions represent equilibria differ for the high and low cutoffs. The low cutoff equilibrium requires that

$$c \geq \frac{(n-1)\theta_O^L}{(n-2)\theta_O^L + 1} \Rightarrow c \geq 0.$$

The high cutoff equilibrium requires that

$$c \geq \frac{(n-1)\theta_O^H}{(n-2)\theta_O^H + 1} \Rightarrow c \geq 1.$$

The low cutoff equilibrium exists even for small  $c$  because the expected utility from a single link to the center agent is so high that there is no cost small enough to make additional links worthwhile. The high cutoff equilibrium, in contrast, only exists for costs so high that no agent can profitably connect to isolated agents on the periphery. This is represented by the 45Ph.D. line in figure 2.7 that connects the upper and lower arcs of the curve; only the portions of the curve to the left of the 45Ph.D. line represent an equilibrium.

The structure of the equilibrium network in this case is an incomplete periphery-sponsored star. Each agent whose value is above the cutoff forms a link to the center, and each agent whose value is below the cutoff does not. Unlike the center-sponsored equilibrium, the periphery-sponsored equilibrium is never empty when every agent's value is above  $c$ ; when that happens, every agent's value is above the low equilibrium cutoff. This equilibrium suffers from a different kind of under-connection, however: Some agents choose not to link to the center and are thus unconnected to the network.

The periphery-sponsored equilibrium exists for wider ranges of costs  $c$  than the center-sponsored equilibrium. This is because in the periphery-sponsored equilibrium, agents receive all of their value via a single link. Even if the cost of that link is high, they may find it worthwhile to link to an insider. In the center-sponsored equilibrium, each link carries only a small part of the total value each agent receives, so that when the costs of linking are too

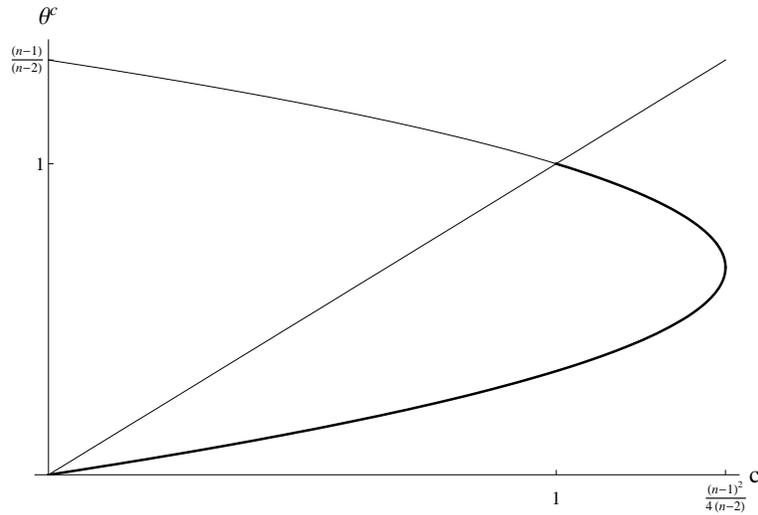


Figure 2.7: Periphery-sponsored equilibrium with uniform values, in bold.

high, agents prefer to form no links at all.

These two examples illustrate the two extremes of the insider-outsider equilibrium, and the different network structures that arise. Intermediate cases are essentially combinations of these two structures; insiders organize themselves in a redundant center-sponsored star, while outsiders organize themselves in an incomplete periphery-sponsored star, centered not on a single agent, but on the entire group of insiders. Insiders thus form a densely connected core to the network, while outsiders form a sparsely connected periphery.

## 2.5 Equilibrium in Public Good Provision

I now characterize equilibria in the contribution game and the voluntary cost sharing mechanism.

**Proposition 5 (Contribution Game)** *Suppose that  $F(c) < 1$ . Then there exists  $\theta_I$  that solves*

$$\theta F(\theta)^{n-1} = K.$$

Let

$$t(\theta) = \begin{cases} 0, & \theta \leq \theta_I \\ K, & \theta > \theta_I \end{cases}$$

Then for  $0 < K < (n-1)$  the strategy profile  $\{t\}$  is a symmetric Bayesian Nash equilibrium

strategy of the contribution game.

The equilibrium of the contribution game is to contribute the full amount necessary for undertaking the project,  $K$ , or else to contribute nothing and free ride. The incentive to free ride is strongest in this model, because if even a single peer contributes, the good will be provided.

**Proposition 6 (Voluntary Cost Sharing Mechanism)** *Let  $\theta_O^n$  solve*

$$\theta \left( 1 + (n-1)(1 - F(\theta)) \right) = c,$$

and let  $\theta_m^n$  solve

$$\theta = \frac{1 + (n-1)(1 - F(\theta))}{(n-1)f(\theta)}$$

Let

$$r(\theta) = \begin{cases} 0, & \text{if } \theta < \theta_O^n; \\ 1, & \text{if } \theta > \theta_O^n. \end{cases}$$

That is, agree to bear part of the cost if  $\theta_i > \theta_O^n$ , and otherwise opt out. Then the strategy profile  $\{r\}$  is a symmetric Bayesian Nash equilibrium of the voluntary cost sharing game for  $0 \leq c \leq \theta_m^n \left( 1 + (n-2)(1 - F(\theta_m^n)) \right)$ .

I index the equilibrium by  $n$ , for comparison with the periphery-sponsored equilibrium of the link formation model. Here, agents must choose whether or not to indicate willingness to bear the cost of providing the good. If they do not agree to share the cost of providing the good, then they are excluded from its benefits. An agent must therefore forecast how many other agents will contribute to the good, and decide if her personal benefit from the good is likely to be high enough for her to bear his expected share of the cost. In this equilibrium, agents do not free-ride, and know that the more of their peers agree to share the cost, the better.

## 2.6 Public Good Equivalence

I now derive the explicit relationship between the link formation game and public good provision. Say that two models  $A$  and  $B$  satisfy *equilibrium expected utility equivalence* if

there exists a pair of order-preserving functions  $h : X_A \rightarrow X_B$  and  $g : A_i \rightarrow B_i$  such that (i) if  $S_A(\theta, X_A)$  is a profile of equilibrium strategies for model  $A$ , then  $S_B = g(S_A(\theta, h(X_A)))$  is an a profile of equilibrium strategies for model  $B$ , and (ii)

$$\mathbb{E}_{\theta_{-i}}[U(f(X_A), g(S_A(\theta, h(X_A))), \theta_i) | \theta_i] = \mathbb{E}_{\theta_{-i}}[U(X_A, S_A(\theta), X_A), \theta_i | \theta_i].$$

That is, if we relabel parameters  $X$  and actions  $A$ , an equilibrium in one model is an equilibrium in the other model, and an agent's expected utility is the same in these equilibria, given the relabeling. I now prove that the two extremes of the insider-outsider family of equilibria, the center-sponsored equilibrium and the periphery-sponsored equilibrium, are equilibrium expected utility equivalent to the contribution game and the voluntary cost sharing mechanism, respectively.

In the center-sponsorship equilibrium of the link formation game, agents choose whether or not to link to every other agent. This decision is effectively whether or not to provide every agent with a path to every other agent, so that they will all receive the benefits of the network. This is essentially the choice agents must make in the contribution game, of whether or not to unilaterally provide the public good.

**Proposition 7** *The center-sponsorship equilibrium of the network formation game and the contribution game satisfy equilibrium expected utility equivalence, with*

$$h((\theta, c, n)) = \left( \frac{\theta}{n-1}, \frac{K}{n-1}, n \right)$$

$$g(a) = \mu(a)c.$$

Proposition 7 shows that the value and cost of a link in the link formation game is equivalent to  $\frac{1}{n-1}$  of the value and cost of the public good in the contribution game, and that choosing to form links in the link formation game is equivalent to choosing to provide the good in the contribution game. In the link formation game, when there is a chance of being connected to the network without having to form any links, there is an incentive to free ride. This incentive is the strongest in the center-sponsored extreme of the insider-outsider equilibria, because every other agent may have a high enough value to choose to connect the entire network. Proposition 7 reveals that this incentive is so strong that the model is equivalent to a model of provision of a pure public good.

The periphery-sponsored equilibrium, in contrast, is equilibrium expected utility equivalent to the voluntary cost sharing mechanism of an excludable public good. Each agent

chooses for herself whether to participate in the network by linking to the center agent. If she does not link, no one else will connect her to the network; she is excluded. This is essentially the choice an agent faces in the voluntary cost sharing mechanism.

**Proposition 8** *The periphery-sponsorship equilibrium of the network formation game and the equilibrium of the voluntary cost sharing mechanism satisfy equilibrium expected utility equivalence, with*

$$h((\theta, c, n)) = \left( \frac{\theta}{(n-1)(1-F(\theta_O^n))}, \frac{K}{(n-1)(1-F(\theta_O^n))}, n+1 \right)$$

$$g(a) = (a_{ik}).$$

Here, the value and cost of a link is  $(n-1)(1-F(\theta_O^n))$  times the value and cost of the public good in the voluntary cost sharing mechanism. An agent's decision to link to the center in the link formation game is the same as her decision to signal that she is willing to pay in the voluntary cost sharing mechanism. There is no chance of being connected to the network by another agent's links, so agents have no incentive to free ride. They forecast their expected utility from forming a link against the cost, and initiate the link if that expected utility is sufficiently high. In the periphery-sponsored extreme of the insider-outsider equilibrium, every agent beside the central agent behaves this way. Unlike the case of the center-sponsored equilibrium and the contribution game, the periphery-sponsored equilibrium with  $n$  agents is equivalent to the voluntary cost sharing mechanism with  $n-1$  agents. This is because in the periphery-sponsored equilibrium, there is a center agent who is passive, so outsiders only forecast the behavior of  $n-1$  other agents. The equivalence of this equilibrium is with the  $n-1$  voluntary cost sharing mechanism: There is "one less" agent than in the link formation game. Another difference between the periphery-sponsored equilibrium of the link formation game is that there is no lower bound on  $c$ , besides 0. This is because agents have no deviations that need to be disallowed: They cannot unilaterally provide the goods themselves – they can only work through the given provision mechanism.

In this way I decompose the link formation equilibria into models of public good provision. Insiders seek to free-ride on each other's links, and in the extreme their decision problem is equivalent to the provision of a pure public good. Outsiders coordinate their links by linking to insiders, and in the extreme their decisions are equivalent to the provision of a club good.

Other equilibria are essentially combinations of these two extremes. In the intermediate cases, insiders choose whether or not to provide a pure public good, where the outsiders represent some "free" expected utility above the value of the good. Outsiders play a cost

sharing game, but with a portion of their benefit coming not from cost reduction when their peers contribute, but rather from a “free” group of insiders. The incentives in these equilibria can be characterized by the extent to which one incentive or the other dominates; whether agents are likely to be connected to the network by their peers, so that the free-riding incentive of the pure public good dominates, or whether they can access many peers with a single well-chosen link, so that the incentive of the club good dominates.

## 2.7 Conclusion

In this paper I show that the connections model of link formation contains two very different strategic incentives, and that both incentives may be present in equilibrium. I prove that we can interpret link formation as a more general kind of public good, where the nature of the public good and its method of provision are determined endogenously.

On the one hand, there is the classic free-riding incentive, in which to avoid being left unconnected agents form many links of their own. This results in a densely connected structure among insiders. On the other, there is the classic coordination problem, in which agents must forecast the behavior of their peers in choosing whether or not to link. This leads to a sparsely connected periphery structure among outsiders.

Both of these incentives lead to centralized network structures, but for very different reasons. Whereas insiders become the center of their own stars, outsiders coordinate their links on an insider. The result is that in equilibrium all agents have short paths to one another. Clustering arises as a byproduct of the excessive linking by insiders; since multiple insiders may choose to become a center of a redundant center-sponsored star, any two insiders who choose to link will share many peers in common.

# CHAPTER 3

## LOCAL PUBLIC GOODS AND NETWORK FORMATION

### 3.1 Introduction

There are many situations in which people rely on their peers for information; consumers may observe the choices of their friends and families before making decisions (Feick and Prices [28] ) and voters rely on their peers for information on candidates (Katz and Lazarsfeld [48] and Beck et. al. [10]). Patterns of research and development depend on the structure of professional relationships (Valente [70]), and managers obtain information from their personal contacts with one another (Cross and Parker [62]).

In each case, one agent has undertaken some costly activity, which benefits his peers. These peers may have to bear some costs in maintaining a relationship with one another, but given this relationship, they have access to each other's information. Information is a *local public good*, whose benefits are not excludable among peers, but are nonetheless costly to undertake. Individuals must decide what relationships to maintain, and how much value to create themselves.

One such situation is the problem of strategic experimentation.<sup>1</sup> If there is some underlying state of the world agents wish to learn, they may undertake costly experiments, or observe the result of their peer's experiments. Experiments will have declining marginal value, as later results are not as informative as early results. Thus, agents must choose how much experimentation to do themselves, and which peers to attempt to learn from.

This paper presents a model of such local public goods and social network formation when agents are heterogeneous. Individuals may either produce a public good themselves, or form a costly link to their peers to access their production, or both. In order to learn from another agent, they must be connected, whether directly or indirectly, in a social network. My results show first that higher degrees of coordination in linking behavior, which results

---

<sup>1</sup>See Bikhchandani et. al. [11] for a review of the social learning literature.

in the formation of centralized structures such as stars, leads to lower levels of public good provision. Second, it is only in networks with no coordination of links that agents will specialize their roles, performing either link formation or public good provision, but not both. Networks with coordination in linking always feature investment while linking, so that some agents perform both roles.

Taken together, this shows that welfare is lowest in networks with more coordination, and highest in networks with the least coordination. One might think that centralized network structures lead to better outcomes, by allowing agents to be connected to one another more easily. Indeed the importance of well-connected central agents has been emphasized in many studies of communication networks. It is this very ease of communication that reduces the incentives to invest in the public good. In equilibrium the ability to economize on link formation by coordinating links leads to lower welfare.

Social network formation with heterogeneity is an important problem that can be difficult to handle, due to the large numbers of actions and outcomes. Previous work has avoided this by focusing on symmetric agents and strong solution concepts, but it is important to be able to accommodate environments where agents may differ in their social abilities, or lack complete information about their peers. I introduce heterogeneity by treating it as private information, and can draw strong conclusions even in a complex environment. I characterize the equilibrium networks that arise, and the equilibrium levels of public good provision.

This paper contributes to a growing literature on network formation and public good provision in networks. The network formation model closest in spirit is that of Bala and Goyal [8]. Networks are formed non-cooperatively, so at least one agent in each relationship undertakes some costly effort to maintain that relationship. This cost represents the time and effort of social activities, and it is in this cost that agents are heterogeneous. This is due to natural variation in the social skills of different people, which are private information. Given their own private cost of forming social ties, individuals forecast the linking and investment decisions of their peers, and make their own linking and investing decision.

In equilibrium, there are only a few possible network structures. These differ in the extent to which agents are able to coordinate their links, and access more peers via a well-chosen link to an intermediary. The ability to do so depends on the global architecture of the network. On one extreme, there is no useful intermediary, and agents must either link to everyone, or no one. This network features the most free-riding on link formation, because if one's peers link at all, they will connect the entire network. Other equilibria feature varying degrees of coordination. Agents are willing to form a link to an intermediary, if their peers

do so as well. The extreme case of coordination is a circle equilibrium, where agent's single links are arranged along a circle, and each agent's peers are linking in exactly the way to maximize the value of a single link. The interaction between the investment decision and the linking decision is different in these diverse equilibrium network architectures.

A key question is whether agents will specialize in their roles: Will some individuals form connections, and make no investment, while others invest in the public good, and form no links? Kranton and Bramoulle [17] study the incentive to provide public goods when the network structure is exogenously given. They find that specialization is always an equilibrium, and that incentives for efficient provision of the good are stronger when there are fewer links; the ability to free-ride on your peers reduces efficiency. Goyal and Galleotti [31], consider a model of non-cooperative link formation and public good provision when agents have perfect information on the actions of their peers, and there is no heterogeneity. They find that specialization in linking and investment is possible, but there are many equilibrium profiles without specialization.

In my model specialization hinges on the cost of public good provision and the global structure of the social network. For every equilibrium architecture, if the cost of public good provision is low enough, then all agents will make positive investments. When the cost of the public good is larger, however, specialization depends on the network architecture. For the uncoordinated equilibrium, agents fully specialize. Agents will either link to all of their peers and make no investment, or invest in the public good and form no links. It is in this architecture that the interaction between linking and investment is weakest, and agents perform one role or the other, but not both. For other equilibrium network architectures, however, the answer is no. No matter what the cost of public good production is, some agents will both link and invest in the public good.

Both public good provision and link formation can be thought of as a kind of public good; in both cases individuals prefer that their peers undertake the costly action of investing in the public good and connecting them to the network. Link formation, however, is more complicated.<sup>2</sup> For some network architectures, linking decisions are complements; the more likely my peers are to link to an intermediary, the more I also wish to link to that intermediary.

While incentives to link to the intermediary increase, incentives to invest in the public good fall. If no one invests while linking the intermediary, however, coordination has no value; there is no point in being connected to someone who is not investing. For these architectures to be equilibria, agents must invest and link at the same time. Thus, specialization in

---

<sup>2</sup>See Jackson [43] for a comprehensive review of the economic networking literature.

roles is only possible for one specific network architecture, where linking decisions are never complements: the center-sponsorship equilibrium.

In the following section I present the model. In section 3.3 I characterize the equilibrium linking and investment decisions, which I discuss further in section 3.4. Section 3.6 concludes.

## 3.2 The Model

Let  $N = \{i, j, k\}$  be the set of agents; that is, we consider a network of only three agents, the smallest non-trivial network. Each agent has the ability to produce a public good, and access the public good production of others. In order to do so, they may invest in public good provision themselves, or form costly links to their peers.

If they have access to  $y$  units of the public good, either from their own investment or from the investment of others, via the social network, they value this at  $u(y)$ . The marginal benefit of the public good is diminishing in  $y$ . Agents must choose a level  $y_i$  of the public good to produce themselves. The production has a linear per-unit cost of  $k$ , which is common to all players.

In addition to producing the public good themselves, agents may choose to initiate links to one another, in order to access the public good production of others. For each link an agent chooses to initiate, she bears a cost  $c_i$ . A linking choice of player  $i$  is an element  $a_i \in A_i = \{0, 1\}^n$ , where  $a_{ij} = 1$  indicates that agent  $i$  initiates a link to agent  $j$ , and we adopt the convention that  $a_{ii} \equiv 0$ .<sup>3</sup> Agents will choose how many links to initiate simultaneously. Two agents  $i$  and  $j$  are connected in a graph  $g$ , written  $g_{ij} = 1$ , if either agent  $i$  or agent  $j$  initiate a link to the other. Therefore,  $g_{ij} = \max(a_{ji}, a_{ij})$ , and call the graph that forms given the profile of actions  $a$  as  $g(a)$ , or  $g(a_i, a_{-i})$ . Value flows both ways across a link in the network; an agent need not initiate any links to receive positive utility, if the links of others connect her to the public good investment of others. In the terminology of Bala and Goyal [8], this is a model of two-way flow of value.

For example, with 3 agents, if  $a_{12} = 1$ ,  $a_{23} = 1$ ,  $a_{32} = 1$ , and  $a_{31} = 1$ , the graph that forms is  $g = \{12, 13, 23\}$ , the complete graph (Figure 2.1). Because the flow of value is two-way, I need not specify a direction for the links in the graph. Given a graph  $g$ , we can calculate the number of agents to whom agent  $i$  has a path in  $g$ . There is a path from  $i$  to  $j$  in network  $g$

---

<sup>3</sup>This reflects the convention that agent  $i$  need not pay any linking cost to benefit from her own public good investment.

if there is a sequence of distinct agents  $\{j_1, \dots, j_m\}$  such that  $g_{ij_1} = g_{j_1j_2} = \dots = g_{j_mj} = 1$ . Let  $N_i(g)$  be the number of agents to whom  $i$  has a path in  $g$ , and let  $\mu(a_i) = \sum_j a_{ij}$  be the number of links that  $i$  initiates.

Let  $g(a)$  be the network formed by the profile of linking decisions  $a$ . Let  $N_i(g)$  be the set of agents to whom agent  $i$  has a path in the network  $g$ , and let  $\mu_i(a)$  be the number of links agent  $i$  initiates at the profile of actions  $a$ . Then the total value that agent receives is

$$U(y_i, y_{-i}, a_i, a_{-i}) = u\left(\sum_{j \in N_i(g(a))} y_j + y_i\right) - ky_i - \mu_i(a)c_i.$$

That is, agents receive value from their own investment in the public good,  $y_i$ , as well as the value of the investment done by any agents they have a path to in the network that forms as a result of the linking decisions. They bear a linear cost  $k$  for any investment they undertake, and a per link cost of  $c_i$  for the links they form. Agents are heterogeneous in this cost of linking; each agent has her own private cost of linking  $c_i$ , and she views the costs of her peers as coming from a common distribution of costs  $F(c)$ , with support  $[0, \infty]$ . Note that there is no decay in value across links in this model; it does not matter how many links a path has; an agent receives the same value regardless.

The solution concept is Bayesian Nash Equilibrium. Given her type  $c_i$ , each agent chooses how much to invest in production of the public good, and whether and with whom to form links, taking as given the strategies of the other agents. A profile of strategies  $(y^*, a^*)$  is a Bayesian Nash equilibrium if for all  $(y, a)$ ,  $c_i$ , and for all  $i$ ,

$$\mathbb{E}_{c_{-i}} U(y^*, a^*) \geq \mathbb{E}_{c_{-i}} U(y, y_{-i}^*, a, a_{-i}^*).$$

## 3.3 Equilibrium

### 3.3.1 Networks

Equilibrium strategies in the linking decision take the form of cutoff rules. Contingent on her private cost of linking, she will either link to both of her peers, neither of her peers, or one of her peers. For low private costs, below  $c^{\text{Low}}$ , she will link to both, for high private costs above  $c^{\text{High}}$  she will link neither, and for intermediate private costs, she may choose to link to one of her peers. There are only a limited number of possible equilibrium arrangements

of that single link.

**Proposition 9** *Equilibria in linking actions take the form of cutoffs rules in each agent's private cost of linking. In any equilibrium, agents will form three links if their private cost of linking is sufficiently low, and no links if it is sufficiently high, and possibly a single link for intermediate costs. The only equilibrium architectures are the following:*

1. *A center-sponsorship equilibrium where agents either form no links or two links.*
2. *A periphery-sponsorship equilibrium, where one central agent either forms no links or two links, and the other two agents either form no links, a single link to the central agent, or two links.*
3. *A hybrid equilibrium, where two agents either form no links, a single link to one another, or two links, while a third agent either forms no links, a single link at random to one of the other two, or two links.*
4. *A mixing equilibrium, where each agent either forms no links, a single link to one of the other agents with equal probability, or two links.*
5. *A circle equilibrium, where each agent either forms no links, a single link to a given peers, forming a circle, two links.*

**Example 4 (Logarithmic Utility and Exponential Costs)** *To illustrate the equilibrium linking decision, consider the following example:*

$$u(y) = \log(y + 1)$$

$$F(c) = 1 - e^{-c}$$

*This utility function has several attractive properties. First, the marginal utility of the public good is bounded above by 1; this is necessary if we are to have any specialization in investment and link formation. If the marginal utility of the public good went to infinity, all investment levels would always be positive. For this reason, it would be especially interesting if, despite this functional form, agent did not fully specialize. In addition, for this utility function agents never invest if  $k > 1$ , so we need only consider  $k \in [0, 1]$  to characterize the equilibrium.*

*The equilibrium linking cutoffs for this example are depicted in figures 3.1 to 3.5.*

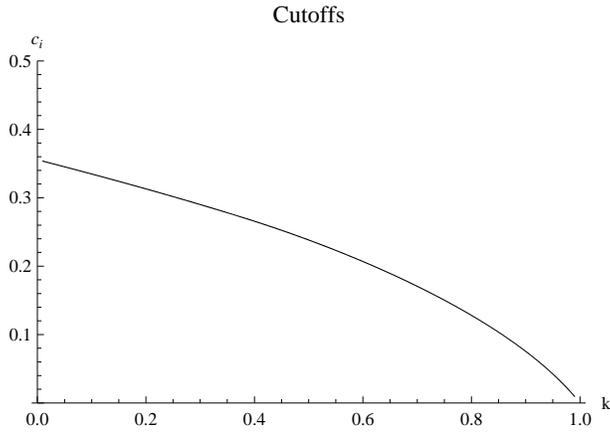


Figure 3.1: Linking Strategies in the Center-Sponsorship Equilibrium. Agents form no links if their private cost is above this line, or two links if it is below this line.

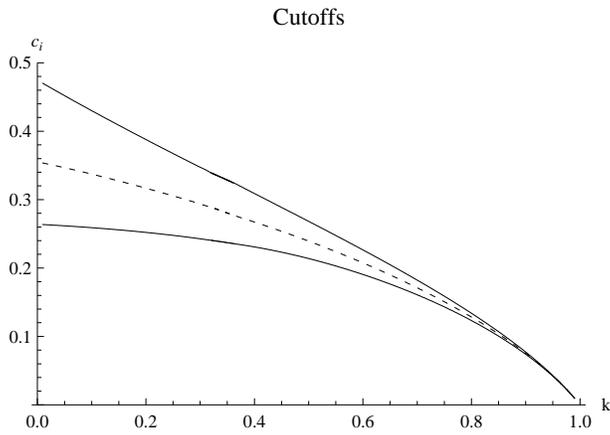


Figure 3.2: Linking Strategies in the Periphery-Sponsorship Equilibrium. Outsiders form two links if their cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line. The Insider forms two links if her cost is below the dashed line, or no links if it is above the dashed line.

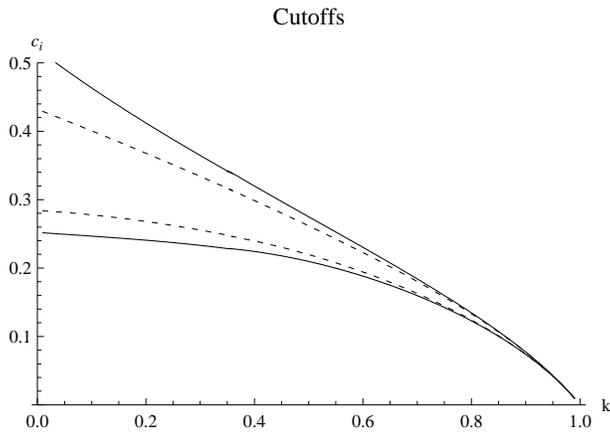


Figure 3.3: Linking Strategies in the Hybrid Equilibrium. The Outsider forms two links if her cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line. The Insiders forms two links if her cost is below the lower dashed line, one link if it is between the two, or zero links if it is above the upper dashed line.

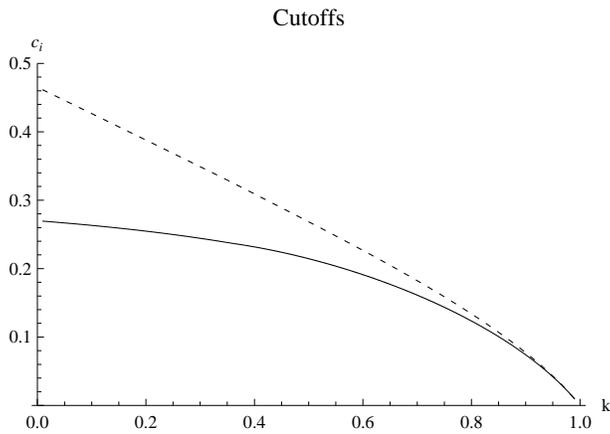


Figure 3.4: Linking Strategies in the Mixed Equilibrium. Each agent forms two links if her cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line.

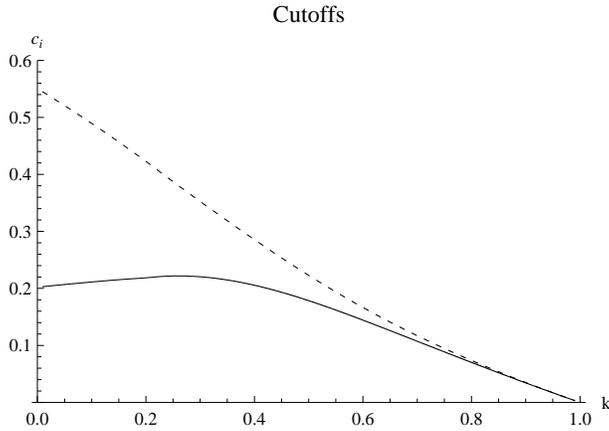


Figure 3.5: Linking Strategies in the Circle Equilibrium. Each agent forms two links if her cost is below the lower solid line, one link if it is between the two, or zero links if it is above the upper solid line.

It is important to note that the center-sponsorship equilibrium and the mixed equilibrium are the only symmetric equilibrium. The rest involve asymmetric strategies, where agents with different positions in the network are treated differently by their peers. It is this asymmetry that allows for *coordination* among the agents; any equilibrium in which some agents direct their links to a specific individual will be asymmetric. In essence, by coordinating their links, agents are able to profitably connect to one another for higher private costs of linking, and economize on link formation. We will see that it is precisely this coordination that leads to underprovision of the public good.

For the center-sponsored equilibrium, agents never form only one link; they either form two or zero. In the periphery-sponsorship equilibrium, one agent never forms a single link, while the other two coordinate their single links towards her. It is this coordination that is possible in asymmetric equilibria, and for this reason, agents are willing to form a single link at lower costs, and two links at higher costs, than in the center-sponsorship equilibrium.

**Proposition 10** *The linking cutoff in the center-sponsorship equilibrium  $c_I$  satisfies  $c^{Low} < c_I < c^{High}$  for all equilibria where agents form a single link.*

In the hybrid equilibrium, two agents form their single link to one another, while the third forms her single link to one of them at random. There is still more coordination in this equilibrium, as the two “inside” agents coordinate by linking each other, and the third agent coordinates by linking one of them.

These three equilibria can be considered members of a single “family” of equilibria. They each consist of one set of agents, the “insiders,” who form many links to one another,

and another set of agents, the “outsiders,” who form a single link to one of the insiders. The center-sponsorship equilibrium is the case where every agent is an insider, the hybrid equilibrium is the case where two agents are insiders, and one agent is an outsider, and the periphery sponsorship equilibrium is the case where one agent is an insider and two are outsiders.

In the mixed equilibrium, each agent forms her single link at random, to one of the others with equal probability. There is no explicit coordination of links in this equilibrium, but because the other agents are likely to be linked to one another, a single link is worth forming.

Finally, in the circle equilibrium, each agent forms her single link to the next agent along a circle; agent  $i$  links to  $j$ , who links to  $k$ , who links back to  $i$ . This equilibrium features the most coordination. Agent  $j$  and  $k$ 's links are arranged in just the correct fashion to give agent  $i$  the most incentive to link to  $j$ .

### 3.3.2 Investment

Contingent on the equilibrium linking strategies being played, and her position in that network, each agent chooses her investment to maximize her expected utility, given the conjectured strategies of her peers. Due to the concavity of the utility function, she will always invest less when forming more links, because she expects to have access to more of her peer's investment.

**Proposition 11** *Equilibrium investment strategies take the following form:*

$$y_i(c_i) = \begin{cases} y_i^{High}, & c_i > c^{High}; \\ y_i^{Mid}, & c^{High} > c_i > c^{Low}; \\ y_i^{Low}, & c_i < c^{Low}; \end{cases}$$

Where  $y_i^{High} > y_i^{Mid} \geq y_i^{Low} \geq 0$ .

This is both a function of her private cost of linking  $c_i$ , and her name,  $i$ , since agents with different roles in the network will have different incentives to invest. Because in some equilibria agents only form zero or two links, these agents do not make a  $y_i^{Mid}$  investment

It is clear that for  $k$  sufficiently small, all investment levels in every equilibrium will be positive; this follows from the fact that marginal utility is never 0. For larger  $k$ , however,

it may be that some of these investment levels are zero. If  $y_i^{\text{Mid}} = y_i^{\text{Low}} = 0$  for some agent  $i$  in some equilibrium, we say that agent is using a *specialized strategy*; when she links she does not invest, and when she invests she does not link. In fact, in every equilibrium, for every agent, there is some critical  $k^*$  above which agents do not invest when they form links to both of their peers.

**Proposition 12** *There exists a critical  $k^*$  for every strategy, in every equilibrium network, such that for  $k > k^*$ ,  $y_i^{\text{Low}} = 0$ , and for  $k < k^*$ ,  $y_i^{\text{Low}} > 0$ .*

So for  $k$  sufficiently large, agents do not invest when linking both of their peers. This implies that for the center-sponsorship equilibrium, for  $k$  sufficiently large, agents *do* use specialized strategies. Individuals with a low private cost of linking specialize in network formation and form links to each of their peers. Individuals with a high private cost of linking specialize in public good provision and form no links.

Is the same true for the other equilibrium architectures? Is it ever the case that  $y_i^{\text{Mid}} = 0$  for sufficiently large  $k$ ? The answer is no: For these equilibria, for any agent who forms one link,  $y_i^{\text{Mid}} > 0$

**Proposition 13** *In any equilibrium in which agents may form a single link, the level of investment undertaken when forming that link is positive, for any cost of the public good  $k$ . That is,  $y_i^{\text{Mid}} > 0$  for all  $k$ .*

**Example 5** *The equilibrium investment strategies, for each equilibrium and for each role in that equilibrium, are depicted for logarithmic utility and exponential costs in figures 3.6 to 3.12*

## 3.4 Discussion

### 3.4.1 Specialization

The various network architectures differ in the extent to which agents are able to free-ride on one another's links, and the extent to which agents face a coordination problem when linking. This in turn leads to differences in investment behavior in each network. The critical

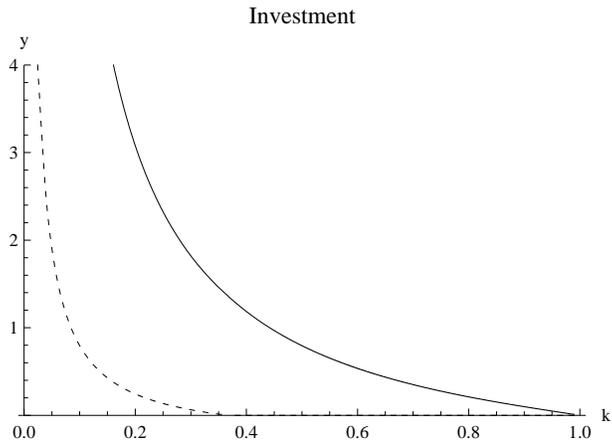


Figure 3.6: Investment Strategies in the Center-Sponsorship Equilibrium. The higher curve is the investment when forming zero links and the lower curve is investment when forming two links.

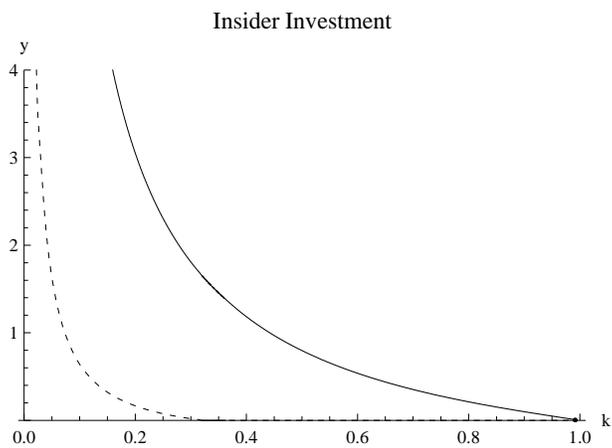


Figure 3.7: Investment Strategies in the Periphery-Sponsorship Equilibrium. The higher curve is the investment when forming zero links and the lower curve is investment when forming two links.

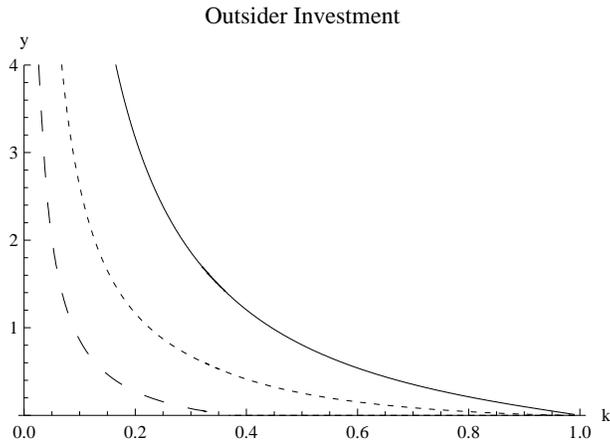


Figure 3.8: Investment Strategies in the Periphery-Sponsorship Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links.

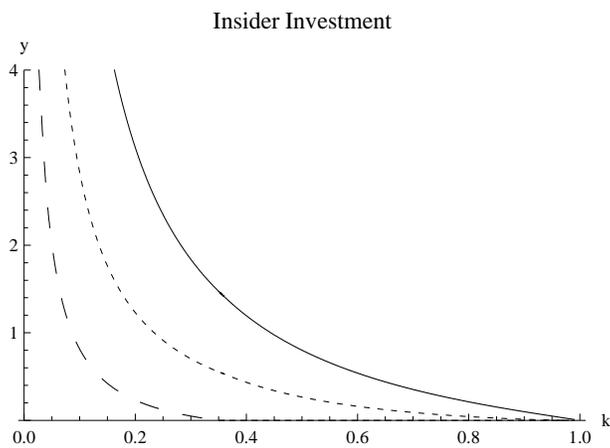


Figure 3.9: Investment Strategies in the Hybrid Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links.

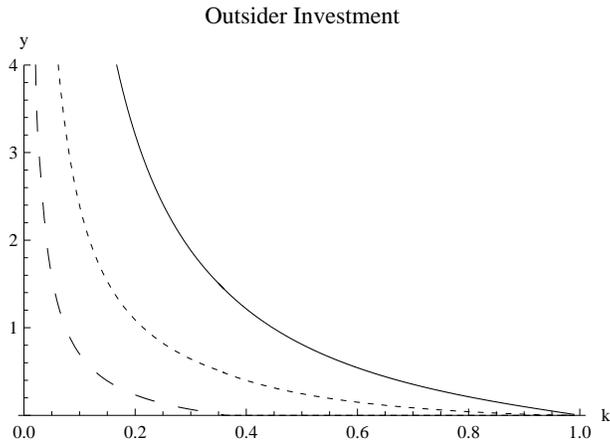


Figure 3.10: Investment Strategies in the Hybrid Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links.

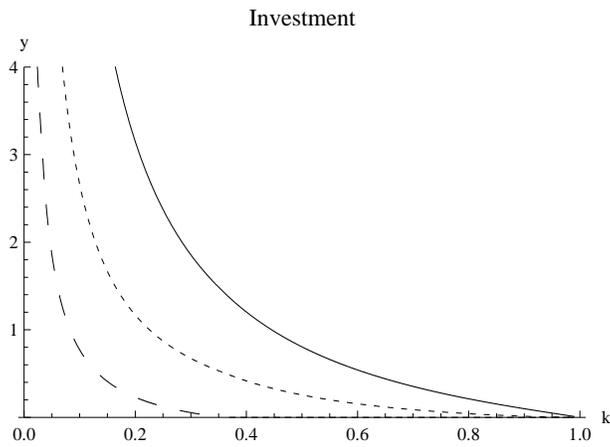


Figure 3.11: Investment Strategies in the Mixed Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links.

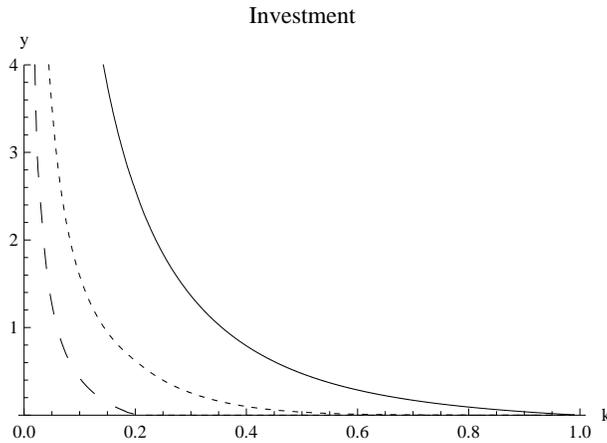


Figure 3.12: Investment Strategies in the Circle Equilibrium. The higher curve is the investment when forming zero links, the middle curve is investment when forming one link, and the lower curve is investment when forming two links.

| Network Architecture  | $k^*$ |
|-----------------------|-------|
| Center-Sponsorship    | 0.367 |
| Periphery-Sponsorship | 0.364 |
| Mixing                | 0.357 |
| Hybrid                | 0.351 |
| Circle                | 0.202 |

Table 3.1: Critical Values of  $k$

cost of investment  $k^*$  is one manifestation of this difference. The value of  $k^*$  for logarithmic utility and exponential costs for each network in table 3.1.

What determines the thresholds is the externality being exerted by the network architecture: In the center-sponsorship equilibrium, where link formation most strongly exhibits free-riding, agents are relatively unable to rely on their peers to invest for them; for this reason they begin to invest in the public good, even when linking, at a relatively high cost of the public good. The remaining equilibria exhibit less free-riding in link formation, and more free-riding in public good provision. Because they are able to access one another via a single link, they do so at lower private costs of linking. This is possible because of the coordination in their equilibrium linking strategies. This comes at the cost of public good provision; because they are linked to one another more easily, they have a lower incentive to invest in the public good, and so the equilibrium investment strategies are interior at a lower critical  $k^*$ . The extreme case in the circle equilibrium, with the most coordination in linking, and a very low cost  $k$  below which agents invest even when forming two links.

| Network Architecture  | $k = .2$ |
|-----------------------|----------|
| Center-Sponsorship    | 3.363    |
| Periphery-Sponsorship | 3.336    |
| Hybrid                | 3.326    |
| Mixing                | 3.307    |
| Circle                | 1.691    |

Table 3.2: Total Welfare

In addition to this difference in linking behavior, none of the coordinated equilibria have specialized investment strategies. The existence of such a specialized equilibrium depends on the network architecture: It exists only for the center-sponsorship equilibrium, which exhibits the most free-riding in link formation. For all other architectures are at most *partially specialized*; when forming many links, agents do not invest, but when forming only one, they do. This is because coordination requires that the agents coordinating be making a positive investment, so that there is some reason to coordinate. Thus, the only way agents can coordinate in equilibrium is if they do not specialize.

### 3.5 Welfare

The welfare properties of the different equilibrium network architectures, for a particular value  $k$ , for logarithmic utility and exponential costs, are presented in table 3.2. The center-sponsorship equilibrium has the highest welfare, for any value of  $k$ . This is because this equilibrium most efficiently separates the roles of investor and connector; a large amount of free-riding on link formation means agents are relatively unable to free-ride on public good investment, and must make their own investment more often. This leads to higher welfare.

The mixing equilibrium has lower welfare than the hybrid equilibrium for low  $k$ , but higher welfare for high  $k$ ; This is because the inefficiencies investment, due to randomness in linking, disappear at a high  $k$ , when some investment is 0. The inefficiencies in the hybrid equilibrium, due to complementarity in linking, becomes more evident in this case. The circle equilibrium is again exceptional; because of the extreme coordination of links, there is extreme free-riding in public good provision, and low welfare.

## 3.6 Conclusion

In this paper I develop a model of network formation and public good provision, and characterize the equilibrium networks and investment profiles. I find that specialization is only possible for one equilibrium network, the center-sponsored star, and that for the others, due to the complementarity of link formation, agents do not specialize. If investment in the public good were complements, rather than substitutes, we may see that coordinated structures have higher welfare. It would be interesting to see if the logic of the equilibrium with private information, that the strategic nature of the linking decision has a strong influence on the investment decision, can be extended to dynamic settings, or if the simultaneity in these decisions is what drives the results.

# CHAPTER 4

## INFORMAL WORK NETWORKS

### 4.1 Introduction

In this chapter we develop a model of time allocation to formal and informal work, where the informational structure of the informal sector leads to peer effects in labour provision, and this effect is stronger in weaker institutional settings. The formal sector is competitive and subject to taxation, while informal jobs are not taxed, but are subject to informational frictions. Informal activities are driven by the social networks of workers. The informational structure of the informal market follows Calvó-Armengol and Jackson [19]. Workers learn about informal job opportunities, and may pass this information to their peers. We characterize the equilibrium time allocations, where formal and informal labour provision are functions of taxation, enforcement, and the properties of the social network.<sup>1</sup>

In this setting, we introduce a very specific kind of heterogeneity among workers, both in the transmission of job information from one worker to another, and heterogeneity over time in the employment status of each worker's peers. These two peer effects drive our results. Peer effects are observed in the time allocation of those workers without an informal job. Their informal time allocation is increasing in the probability that they receive informal job information from one of their peers. Workers who already have an informal job do not take social ties into account when choosing informal labour supply and allocate more time to informal work than a worker without an informal job. Because workers who already have an informal job may pass job information to their peers, formal labour supply is decreasing in the number of peers with an informal job, while informal labour supply is increasing, for workers without an informal job.

This model allows for the traditional explanations of informal work, i.e., taxation and

---

<sup>1</sup>This work will be published as “Informal Work Networks,” by Marcelo Arbex and Dennis O’Dea, in the *Canadian Journal of Economics*.

institutional and enforcement conditions.<sup>2</sup> We model the strength of the institutional environment by the evasion detection probability. A lower detection probability reflects a weaker institutional environment or, equivalently, a higher probability workers will keep their informal jobs. This may be due to corruption or the lack of appropriate instruments to enforce tax obligations. In this environment informal work is more attractive, and peer effects are stronger. Workers' time allocation becomes more sensitive to the possibility of receiving job information from their peers, and their chances of finding informal employment rise much more from stronger social ties, than when institutions are stronger. In any environment, stronger social ties improve the transmission of job information and increase utility.

We consider four network structures (empty, complete, star and wheel) to compare different social structures and information transmission processes. We find that different networks lead to differences in labour allocation, as well as differences in long run average income and utility. Long run average income depends on two components: income earned in each state of the economy and the long run probability of those states. The effect of institutional conditions are reflected in changes to these two components. When the probability of being detected working in the informal sector is low, workers increase the time they allocate to that sector, which will increase the income earned in every state, for any network. It will also change the likelihood of these states. This has an opposing effect on long run average income. It becomes easier for agents to transition from no informal job to employment in the informal sector and this increases the likelihood that, in any state, workers with informal jobs start the period without them. Because agents who start a period without an informal job allocate less time to informal work, this decreases long run average income. In well connected networks, this later effect dominates the former effect because, in these networks, informal jobs are easier to get and workers are more able to become employed in the informal sector. Hence, even as weaker institutions lead workers to work more in the informal sector and workers' utility rises, in well connected networks long run average income falls.

It has long been understood that social ties are important in labour markets (see Granovetter [36], and Ioannides and Loury [42] for a recent survey). Indeed, a large proportion

---

<sup>2</sup>Several works have studied the effects of government interventions, such as taxation and labour market regulations (Banerjee and Andrew [9]; Johnson et al. [46], Friedman [29]; Schneider and Enste [67]; Fugazza and Jacques [30]) and the impact of bureaucracy, corruption and other institutional and enforcement conditions on informal labour (Busato and Chiarini [18]; Choi and Thum [21]; Dabla-Norris [23]). Other studies have argued that the heterogeneity of firms and entrepreneurs and limited access to capital markets are key to explain the emergence of informal activities (Dessy and Pallage [24]; Gordon and Li [35]; Amaral and Quintin [5]; Antunes and Cavalcanti [6]). This paper is also related, although not explicitly, to a great deal of literature on tax evasion and social norms (see, for instance, Kirchler [49] and Lemieux and Frechette [53]).

of people (about 50% on average) hear about or obtain jobs through friends and relatives (see Holzer [41], Montgomery [58], Topa [69], for the U.S., Gregg and Wadsworth [38], for the U.K. and Addison and Portugal [2], for Portugal). Numerous studies have shown the importance of referrals and word of mouth for job search (See Bradshaw [16], Bortnick and Ports [15], Blau [12], and Blau and Robins [13]). Such referrals come from a worker's peers, family, and co-workers. The nature of this social network can have a large impact on job search (see, for instance, Munshi [59] for a study of Mexican immigrants in the United States). It is also clear that usage of social networks varies along a variety of dimensions, such as gender, age, education, ethnicity (see Bradshaw [16], Ports [61], Elliott [25]). The importance of social connections for job search also varies by occupation (see Ioannides and Loury [42], Rees [64] and JORGE and Valadão [47]). While many different forces might lead to distinct usage of social networks, we focus in this paper on differences that are the result of frictions in job search created by the different legal status of formal and informal labour.

Workers use alternative sources of information when searching for jobs. These can be broadly classed as formal sources of information, such as job postings and placement agencies, and informal sources of information, such as family, peers and coworkers. While formal work may be found through formal and informal sources of information, there are no formal sources of information for informal work. We interpret this asymmetry between the two sectors in our model to be the result of legal frictions and the nature of informal activities that prevent firms and workers from finding one another.<sup>3</sup> Firms and workers must instead rely on word of mouth, referrals and happenstance to find each other. This process will depend on the social network of contacts, colleagues, friends and family of the worker. That is, information on informal job opportunities arrives from peers and is passed from one agent to another in a social network. It is natural, therefore, to model the market for informal work differently than the market for formal work.

Social capital is especially important for informal activities. Close ties to neighbours, friends and family members often play a role in the existence of such activities. Moreover, informal networks are a potential source of economic support and are more likely to exist in areas with stronger social networks (See Gaughan and Ferman [32], Losby [54] and Gerxhani [33]). The informal economy consists of both self-employment and wage employment. Informal self-employment is usually part-time employment that provides a supplemental income to people's primary employment Alden [3]. This often involves switching between economic sectors (formal and informal) during the same workday. Schneider and Enste [67] argue that

---

<sup>3</sup>Informal activities are defined as all income generating activities which do not comply with tax obligations, tax evasion and non-compliance with legislation.

the informal use of labour may consist of a second job after (or even during) regular working hours. Ratner [63] argues that informal opportunities may derive from an agent being formally employed and using the equipment and tools available, as well as access to consumers, in its formal employment for informal work outside the hours of employment. Pedersen [60] provides evidence that, in Denmark, Norway, Sweden, Germany and the United Kingdom, employment in the formal sector determines whether people carry out informal activities, suggesting that workers build up an informal network of contacts in their formal job.

A common feature of the existing research on this topic is the absence of explicit and formal treatment of informal activities as driven by social ties, ignoring peer effects or social interactions effects. Despite some evidence, the economic literature of informal activities and tax evasion is largely based on models of individuals deciding whether to carry out informal activities where the taxpayer is assumed to be completely individualistic and amoral (see Allingham and Sandmo [4], Cremer and Gahvari [22]; Lemieux and Frechette [53]; Slemrod and S. [68]; Sandmo [65] and Cremer and Gahvari [22]).

The paper proceeds as follows: We present the basic model in section 4.2. In section 4.3, we describe the model's equilibria and we analyze a particular parameterization of our model to illustrate our results. We also discuss the government's optimal choice of tax-enforcement policy instruments to maximize social welfare. Section 4.4 concludes.

## 4.2 The Model

Let  $N = \{1, \dots, n\}$  be the set of workers. Each worker has a utility function  $u(c, h)$ , where  $c$  is consumption and  $h$  is leisure. Each worker  $i$  must choose how much time to allocate to work in the formal sector,  $l_i$ , and how much time to allocate to work in the informal sector,  $\gamma_i$ . There are no technological differences between the two sectors of the economy. The only differences are their informational structure and how the worker's labour income is taxed. There is an informational asymmetry between the two markets. The formal labour market is competitive, so that workers can always find a formal job if they want one. The informal market, on the other hand, is mediated by social networks. A worker must hear of an informal job from his peers. Information about formal jobs is public and abundant, while information about informal jobs is only passed from one person to another. If a worker hears about a job, he will either take it himself, if he doesn't already have an informal job, or he will pass the job to one of his peers. We assume that, except for taxes, formal and informal

jobs pay the same wage.<sup>4</sup> Proceeds from taxation and enforcement are dissipated.

Time allocated to the informal sector may not necessarily be spent in work. If a worker is not matched to an informal job or loses an informal job he already has, he may unwillingly spend it on leisure. Our approach follows Calvó-Armengol and Jackson [19] and we interpret the timing as one where job break-up occurs, essentially, at the beginning of the period. We track a worker's informal employment status by a state variable  $S_{it}$ , where  $S_{it} = 1$  for a worker  $i$  employed in the informal sector and  $S_{it} = 0$  for a worker without an informal job. We write  $S_t = \{S_{1t} \dots S_{nt}\}$  as the vector containing every agent's employment status at time  $t$ .

A job contact network, or social network, is described by a symmetric matrix  $g$ , where  $g_{ij} \in \{0, 1\}$  denotes whether a link exists between agents  $i$  and  $j$ . The symmetry of  $g$  reflects the reciprocity of social ties. We will focus on a number of cases (Figure 1), where networks have different structures and different forces shape the strength of social ties, as follows:

**Case 1 (Empty)** *Let  $g_{ij} = 0$  for all pairs of workers  $i$  and  $j$ . Job information received by an agent will be either used by himself or discarded.*

**Case 2 (Complete)** *Let  $g_{ij} = 1$  for all pairs  $ij$ . Every agent is directly connected to every other agent. Informal job information each agent passes goes to each of his unemployed peers with equal probability. This network features the strongest peer effects.*

**Case 3 (Star)** *Let  $g_{ik} = 0$  for all agents  $k \neq j$ ; let  $g_{ij} = 1$  for every agent  $i \neq j$ . This network is centered on agent  $j$ , who plays a central role in connecting otherwise unrelated agents. Periphery agents can only receive job information from the center of the star, whose ties are equally strong to all her peers. The center agent may receive information from any of the periphery agents, and no other information is passed.*

**Case 4 (Wheel)** *Order agents from 1 to  $n$ , and let  $g_{ik} = 1$  if  $i = k \pm 1$  modulo  $n$ , and  $g_{ik} = 0$  otherwise. Agents are connected to only two peers, forming a wheel. Agents can only receive job information from their immediate neighbors in the network. They have no access to job information from agents farther away in the social network, but also face little competition for the information of their neighbors.*

---

<sup>4</sup>The reason for this is to highlight the role of peer effects in time allocation between formal and informal labour markets. Differences in wages between the two sectors lead to differences in time allocation, and could easily be included in the model. This only clouds the impact of social networks, however.

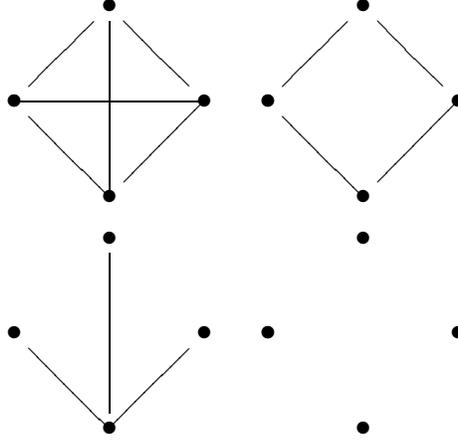


Figure 4.1: Network Structure Examples: Clockwise from upper left, a complete network, a wheel network, an empty network, and a star network.

The informal job transmission process is described by a function  $p_{ij} : \mathbf{R}^n \rightarrow [0, 1]$ , where  $p_{ij}(S_t)$  gives the probability that a job originally heard by worker  $j$  is eventually received by worker  $i$  when the state is  $S_t$ . We will assume that  $p_{ij}(S_t)$  is nondecreasing in every element of  $S_t$ , i.e., the more workers have informal jobs, the more likely they are to pass information about jobs. This might be, for example, because they pass jobs they themselves do not need. The function

$$p_{ij} = \begin{cases} (1 - S_{it-1})\alpha, & i = j \\ (1 - S_{it-1})S_{jt-1}\alpha \frac{g_{ij}}{\sum_{k|S_{k;t-1}=0} g_{jk}}, & i \neq j \\ 0, & \text{otherwise.} \end{cases}$$

describes the case where a worker  $i$  hears about an informal job himself with probability  $\alpha$  and is passed information from another worker  $j$  if  $j$  already has an informal job, hears about one, and then passes him that information. The probability  $j$  does so is given by the relative strength of the social ties between  $i$  and  $j$ , given by a function  $g_{ij}$ . This is the transmission function considered in Calvó-Armengol and Jackson [19]. Job information is only passed “one step” along the job contact network in this example. Information that is not used by its first recipient is lost. This is a simple kind of social network with peer effects, where the prevalence of informal jobs is given by  $\alpha$ , and the strength of the social ties between  $i$  and  $j$  is given by a social network  $g$ .

If workers choose to work in the informal sector, they face some chance of detection by the tax authorities. We model this by a parameter  $\beta$ , which gives the probability of detection, determined by audit policies and monitoring procedures. In the event of detection, the worker

loses his informal job and the opportunity to work in the informal sector that period.<sup>5</sup> This parameter can also be interpreted as the informal job break-up probability.

Workers face the following budget constraint:

$$c_{it} \leq l_{it}(1 - \tau) + S_{it}\gamma_{it}, \quad (4.1)$$

where consumption is produced in the formal and informal sector according to the constant returns to scale production function  $y = \sum_i (l_{it} + S_{it}\gamma_{it})$ . There is no savings in this model.

The sequence of events in a given period  $t$  is summarized in Table 4.1. The time allocation decision is made at the beginning of period  $t$  and cannot be revised. Workers may hear of new informal jobs, which they will either take themselves or pass to their peers. Workers who either started the period with an informal job, or heard of one, are then detected with probability  $\beta$  and may lose the chance to work in the informal sector, instead spending that allocated time on leisure. Finally, workers work in the formal and informal sector, if they have not been detected, and consume their income. If agent  $i$  is unemployed in the informal sector at the start of period  $t$ ,  $S_{it-1} = 0$ . He will become employed if he hears of an informal job and is not detected, in which case  $S_{it} = 1$ . Otherwise,  $S_{it} = 0$ .

Table 4.1: The sequence of events in period  $t$

|   |   |
|---|---|
| 0 | $S_{it-1}$ is given   |
| 1 | Workers choose formal labour time allocation $l_{it}$ and informal time allocation $\gamma_{it}$  |
| 2 | Workers hear about informal jobs with probability $1 - \prod_j (1 - p_{ij}(S_{it-1}))$  |
| 3 | Unemployed workers who hear of an informal job become matched with that job, and workers carry out their labour.  |
| 4 | All matched informal jobs are detected with probability $\beta$ ; $S_{it}$ is now determined.   |
| 5 | Workers with an informal job who are not detected receive income $l_{it}(1 - \tau) + \gamma_{it}$ .<br>Workers without an informal job or are detected, receive income $l_{it}(1 - \tau)$ .                                   |
| 6 | Workers consume their income and experience leisure.<br>Workers with an informal job and not caught experience leisure $1 - l_{it} - \gamma_{it}$<br>Workers unmatched with an informal job experience leisure $1 - l_{it}$ . |

---

<sup>5</sup>Although a more complex enforcement scheme could be considered, it would not change the results substantively.

## 4.3 Informal Work Networks

### 4.3.1 Network Structure and Information Transmission

An equilibrium of this model is a profile of time allocation choices  $\{l_i, \gamma_i\}_{i=1}^n$  such that, given  $l_{-i}$  the profile of actions of workers besides  $i$ , and the state of nature  $S_t$ ,  $c_i, l_i, \gamma_i$  maximize expected utility, with respect to  $S$ , for every worker  $i$ . Let  $S_t = \{S_{it}\}_{i=1}^n$  be the vector of employment states. We now write the worker's dynamic problem as a function of this state  $S$ .

$$\begin{aligned}
 V(S_{t-1}) &= \max_{c_{it}, l_{it}, \gamma_{it}} \mathbb{E}_S U(c_{it}, (1 - l_{it} - S_{it}\gamma_{it}) + \delta \mathbb{E}_S [V(S_t)]) \\
 &\text{subject to} \\
 c_{it} &\leq l_{it}(1 - \tau) + S_{it}\gamma_{it} \\
 S_t &= M(S_{t-1}, l_{it}, \gamma_{it}, l_{-it}, \gamma_{-it}),
 \end{aligned} \tag{4.2}$$

where  $M$  is the law of motion for  $S_t$ . Let  $K$  be the probability an agent has an informal job in hand, prior to job breakup. Regardless of the social network, for a worker who starts the period with an informal job in hand, called an ‘‘informal worker’’,  $K = 1$ , and so his peer's employment status and the structure of the social network have no effect on his time allocation decision. All an informal worker is concerned with is the possibility of detection,  $\beta$ , and the tax rate,  $\tau$ .

For a worker who starts the period without an informal job, called a ‘‘formal worker’’,  $K$  will in general depend on  $\alpha$ , the employment status of the other agents, and the structure of the social network  $g$ . Thus, an agent without an informal job, who has many peers who may pass him information, and who has few competitors for that information, will allocate more time to the informal sector than an agent with only a few peers to pass information or many competitors for that information. His choices will exhibit peer effects.

The structure of social ties among workers determines the probability a formal worker is passed job information. Conditional on being unemployed in the informal sector, the probability worker  $i$  receives at least one offer is  $K = 1 - \prod_j (1 - p_{ij})$ . The probability that a worker unemployed in the informal sector finds an informal job and is not detected is  $(1 - \beta)K$ . Workers with different social networks and in different employment states face decision problems that differ only through the function  $K$ .

For each of the networks we investigate, conditional on being unemployed in the informal

sector,  $K$  is given by

$$K = 1 - (1 - \alpha) \prod_{j \neq i} \left( 1 - S_{jt-1} \alpha \frac{g_{ij}}{\sum_{k|S_{k;t-1}} g_{jk}} \right). \quad (4.3)$$

That is, the probability of being passed information depends on the structure of the network  $g$ . For the *empty network*,  $g_{ij} = 0$  for every pair of workers  $ij$ , where  $i \neq j$ . That is, workers are not connected to one another. The probability a worker becomes employed in the informal sector is simply  $\alpha$ , the probability he hears of a job himself.

On the other extreme, the *complete network*,  $g_{ij} = 1$  for all pairs of workers and social ties are of equal strength between everyone. This network has the most information transmission. If  $m$  is the number of peers of worker  $i$  who are informal workers, we have

$$p_{ij} = p_{Complete}^* = \begin{cases} \alpha, & i = j; \\ S_{jt-1} \frac{\alpha}{n-m-1}, & i \neq j. \end{cases}$$

The probability he receives information from an employed peer (a peer with an informal job) depends on the number of competitors he has for that information,  $n - m - 1$ . Therefore,

$$K_{Complete}^* = 1 - (1 - \alpha) \left( 1 - \frac{\alpha}{n - m - 1} \right)^m,$$

for every worker without an informal job. Information may be passed from up to  $m$  other workers.

In the *star network* there is a central agent with links to every other agent, and no other links. Clearly, the center and the periphery agents will behave differently. The central agent has access to the information of every periphery agent, and faces no competition for that information, whereas periphery agents only have access to the center and may have many competitors. For the center agent, the probability he receives informal job information from some agent  $i$  on the periphery is given by

$$p_{Center}^* = \begin{cases} \alpha, & \text{If } i \text{ has an informal job;} \\ 0, & \text{Otherwise.} \end{cases}$$

If an agent on the periphery hears of an informal job, and if he does not use it himself, he

can only pass it to the center. Therefore,

$$K_{Center}^* = 1 - (1 - \alpha)(1 - \alpha)^m,$$

where  $m$  is the number of informal workers on the periphery. The probability he becomes employed is much higher than, for example, any agent in the *empty network*. He may be passed information from any of these  $m$  agents. For an agent on the periphery, however,

$$P_{Periphery}^* = \begin{cases} \frac{\alpha}{n-1-m}, & \text{If the center is an informal worker;} \\ 0, & \text{Otherwise.} \end{cases}$$

The probability he is passed information from the center is lower, and is only positive in the case where the center already has an informal job.<sup>6</sup> When the center does not have an informal job, no one else is passed any information. This is because the center will keep any information he receives for himself. If multiple periphery agents have job information to pass, those opportunities may be wasted in the *star network*. The equilibrium probability a periphery agent receives job information is

$$K_{Periphery}^* = \begin{cases} 1 - (1 - \alpha)(1 - \frac{\alpha}{n-1-m}), & \text{If the Center has an informal job;} \\ \alpha, & \text{Otherwise.} \end{cases}$$

In the *wheel network*, the relevant state space of the network is larger. The probability an agent receives job information depends not only on how many of his peers have informal jobs, but how those peers are arranged. For instance, for the case  $n = 4$ , if  $m = 0$  or  $m = 3$ , where  $m$  is the number of a worker's peers who are informal workers, there is no ambiguity: there is only one way to arrange 0 or 3 agents on a circle, and every formal worker is in the same situation. If  $m = 1$ , then an informal worker may be one of the two agents adjacent to that informal worker, or not. If  $m = 2$ , these two agents may be adjacent to each other, or not. These two arrangements will have different implications for formal workers hoping to receive informal job information. If the two informal workers are adjacent, then each informal worker may be passed information from only one of them, but he faces no competition for that information. If the two informal workers are separated, then each formal worker may get information from either of them, but now faces competition from the

---

<sup>6</sup>When there are four agents, a periphery agent can only have at most  $m = 1$  peers who may pass her information, but may have 0, 1 or 2 competitors for that information.

other formal worker. Therefore,

$$p_{Wheel}^* = \begin{cases} 1 - (1 - \alpha)^2, & \text{If } m = 3; \\ \alpha, & \text{If } m = 2 \text{ and they are adjacent;} \\ \frac{2\alpha(1-\alpha)}{2} + \frac{\alpha^2}{4}, & \text{If } m = 2 \text{ and they are separated;} \\ \frac{\alpha}{2}, & \text{If } m = 1 \text{ and the worker is adjacent;} \\ 0, & \text{If } m = 1 \text{ and workers are separated;} \\ 0, & \text{If } m = 0. \end{cases}$$

And the probability an agent without an informal job learns about one is given by

$$K_{Wheel}^* = \begin{cases} 1 - (1 - \alpha)^3, & \text{If } m = 3; \\ 1 - (1 - \alpha)^2, & \text{If } m = 2 \text{ and they are adjacent;} \\ (1 - \alpha)\left(\frac{2\alpha(1-\alpha)}{2} + \frac{\alpha^2}{4}\right) + \alpha, & \text{If } m = 2 \text{ and they are separated;} \\ (1 - \alpha)\frac{\alpha}{2} + \alpha, & \text{If } m = 1, \text{ the formal worker is adjacent;} \\ \alpha, & \text{If } m = 1, \text{ the formal worker is separated;} \\ \alpha, & \text{If } m = 0. \end{cases}$$

### 4.3.2 Network Structure and Long Run Behavior

The state of the economy specifies how many agents have informal jobs, and how they are arranged in the social network  $g$ . There are only finitely many such states, with a transition matrix  $T$  determined by  $\alpha$ ,  $\beta$ , and the network structure  $g$ . Therefore, there is a unique invariant distribution  $\mu$  of the state.<sup>7</sup>

We illustrate the invariant distribution of these transition matrices, for the *empty* and *complete* networks, in Figure 2. The regions of the unit square, corresponding to different values of  $\alpha$  and  $\beta$ , are shaded according to which state is likeliest for those values. State behavior is similar in all networks. If the probability of detection  $\beta$  is close to 1, zero informal workers is the most likely state. As  $\alpha$  increases, the cutoff at which this state becomes the likeliest increases. For lower values of  $\beta$ , states with more informal workers are more likely. While these appear similar across networks, there are differences in these state probabilities. For example, four informal workers are more likely in the *complete network* than the *empty*

---

<sup>7</sup>Such a distribution is guaranteed to exist because the transition matrix of the state is irreducible for  $\alpha, \beta \in (0, 1)$ .

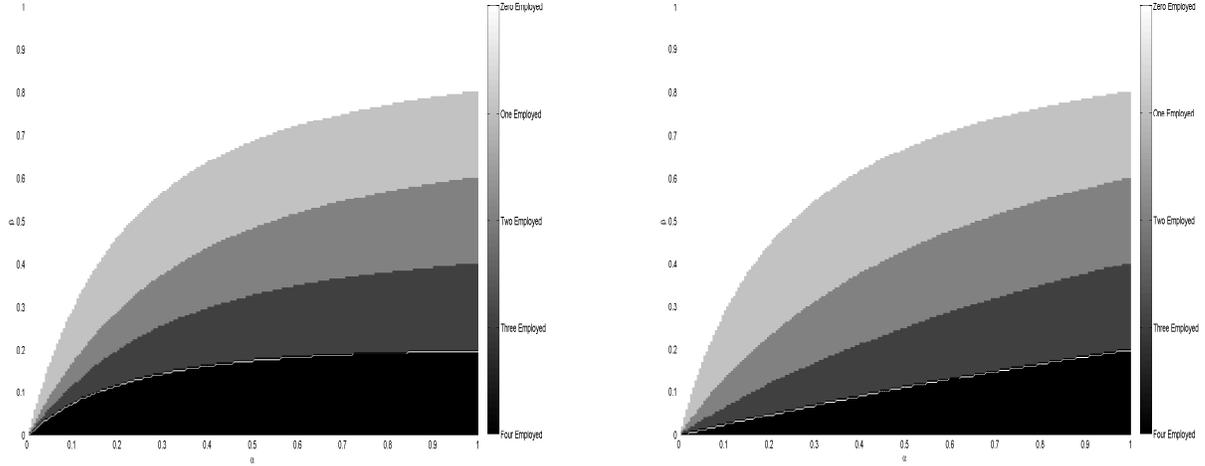


Figure 4.2: Most likely state in the long run for the Empty and Complete Networks

*network* for all values of  $\alpha$  and  $\beta$ .

Because there is no savings, labour allocations depend only on the state of the economy, so in each of these states a certain amount of time is allocated to the formal and informal sectors. Taking the length of one period in the model to be a week, let  $F$  to be the vector of time allocated by all agents to formal labour in each state  $k$ . Then,  $[(1/n)(1 - \tau)F\mu \times 52]$  is the long run average yearly income from the formal sector. Note that this depends on  $\alpha$ ,  $\beta$ ,  $\tau$  and  $g$  through both the invariant distribution  $\mu$  and the labour choices  $F$ .

Similarly, in every state, agents allocate a certain amount of time to the informal sector. Unlike formal labour, however, the outcome of the informal labour market is uncertain. The transition probabilities from state  $d$  are given by the column of the transition matrix,  $T_j$ . For each initial state  $d$ , and final state  $k$ , a certain amount of time is allocated to informal work, and a certain amount of income from informal work is actually earned. Let  $I$  be the matrix of total informal labour performed for each possible state transition. Then the average yearly income from the informal sector is given by  $[(1/n)(I \otimes T)\mu \times 52]$ , where  $(I \otimes T)$  gives the Hadamard product of  $I$  and  $T$ , a vector whose elements are the sum of the elementwise product of the columns of  $I$  and  $T$ , which tells us the expected informal income in each state. Multiplying this by  $\mu$ , we have the long run average informal income. We can calculate the long run average yearly utility of agents in a similar way, as well as the long average total income agents earn.

Our model thus illustrates the importance of peer effects in worker's decision to work in the informal sector. When job information on informal work passes from worker to workers,

through a social network, then peer effects will influence time allocation.

### 4.3.3 Time Allocation, Income, Utility and Network Structure

In order to illustrate the effects of changes in institutions and social networks on the economy, we specify a particular utility function and social structure. We solve the model to find the optimal time allocation to the formal and informal sectors, and illustrate the dynamics of the economy. This allows us to evaluate the impact of network structures on time allocation, income, and utility. We demonstrate that stronger social networks allow agents to better exploit sources of information, leading to higher utility. Peer effects are stronger in weaker institutional environments. We show that the weakening the enforcement of tax obligations, while increasing the income a worker receives from informal work, may actually reduce average income in the long run. The implicit welfare discussion in this section is strictly from a worker's point of view.

We solve the model assuming that the utility function is of the form

$$u(c, h) = c^{1/2}h^{1/2},$$

We set  $n = 4$  and we consider the simplest setting with four distinct networks.

The job transmission function  $p$  depends on a job arrival probability  $\alpha$  and the strength of social ties between workers,  $g_{ij}$ . The effects of changes in  $\alpha$  and  $g$  can be discerned by their effect on  $p_{ij}$ . A rise in  $\alpha$  will increase  $p_{ij}$ , and so increase time allocated to the informal sector. If the social ties between  $i$  and  $j$  become stronger, information is more likely to be passed from  $j$  to  $i$ . If another worker  $k$ 's ties to  $j$  become stronger, then because  $i$  and  $k$  are competitors for  $j$ 's information,  $p_{ij}$  decreases, and so time allocated to informal work falls.

The comparative statics of the equilibrium labour choices are examined for different networks and different states. It is clear that the relative values of  $\tau$  and  $\beta$  and the attitude of workers toward risk is relevant for these choices, because workers are balancing the tradeoff between working in the formal sector, subject to the tax  $\tau$ , and working in the informal sector, subject to the risky lottery of the job matching process  $(\alpha, \beta, g)$ . We observe that, regardless of the network structure, time allocated to formal (informal) work is decreasing (increasing) in the arrival probability  $\alpha$  and in the income tax  $\tau$ . Formal labour supply is increasing in the detection probability  $\beta$ , while hours worked in the informal sector decrease as the probability of detection increases. Workers who start the period with an informal job

in hand allocate more time to informal work than a worker who does not. Because they are sure to be able to engage in informal work, they set aside more time to do so.

Differences in the networks lead to differences in  $K$ , the probability an unemployed agent learns of a job opportunity, which leads to differences in the equilibrium time allocation of informal workers. Let  $l_E$  be the formal labour supply of a worker employed in an informal job and  $l_U(m)$  is the formal labour supplied by a worker not employed in an informal job when he has  $m$  peers employed in the informal sector. In each network structure, informal workers always work less in the formal sector than formal workers.

The informal time allocation of an informal worker,  $\gamma_E$ , and of a formal worker without any peers who may pass information (that is, when  $m = 0$ ),  $\gamma_U(0)$ , represent the highest and lowest time allocation to informal work, respectively. For all other networks and states, time allocation will be between these two extremes. The more peers a formal worker has who may pass him informal job information, the more his time allocated to informal labour approaches the time allocation of a worker who already has an informal job. That is, formal worker's time allocations exhibit peer effects.

Increasing the probability workers are caught evading taxes ( $\beta$ ) will discourage informal work and lead agents to allocate more time to formal work. Formal workers allocate less time to informal work, and more to formal work. Differences in labour time allocation between the networks are larger. Consider a formal worker with one peer employed in the informal sector. Besides the fact that he works less in the formal sector, he works much less if there are no peer effects (as in the *empty network*) than when there are strong peer effects (as in the *complete network*). This shows that the effects of the network structure and the information transmission mechanism are stronger when institutions are weaker. When the detection probability  $\beta$  is low, a worker is more likely to keep his informal job and so he cares relatively more about informal job opportunities. Thus, in an economy where enforcement is weaker, social ties matter more and workers alter their behavior in order to influence their chance of getting an informal job. That is, the informal job information transmission process becomes more important in an environment where informal jobs do not break up easily.

The effects of changes in income taxation on time allocated to formal and informal work are simple: when taxes are lower, agents will allocated more time to the informal sector. For any network, and any state, as  $\tau$  goes to one, time allocation to formal work goes to zero. Conversely, as  $\tau$  goes to zero, time allocation to informal work goes to 0.5.<sup>8</sup> The various networks and states differ only in how quickly agents reduce their time allocation to formal

---

<sup>8</sup>This is an artifact of the utility function, so that 0.5 is the most time ever allocated to any sector.

work. The more peers an agent has who may pass information, and the more conducive to information transmission the network is, the more responsive time allocations are to changes in  $\tau$ . Figures 3 – 5 illustrate the informal labour supply choices with respect to the arrival probability, probability detection and income tax, respectively.

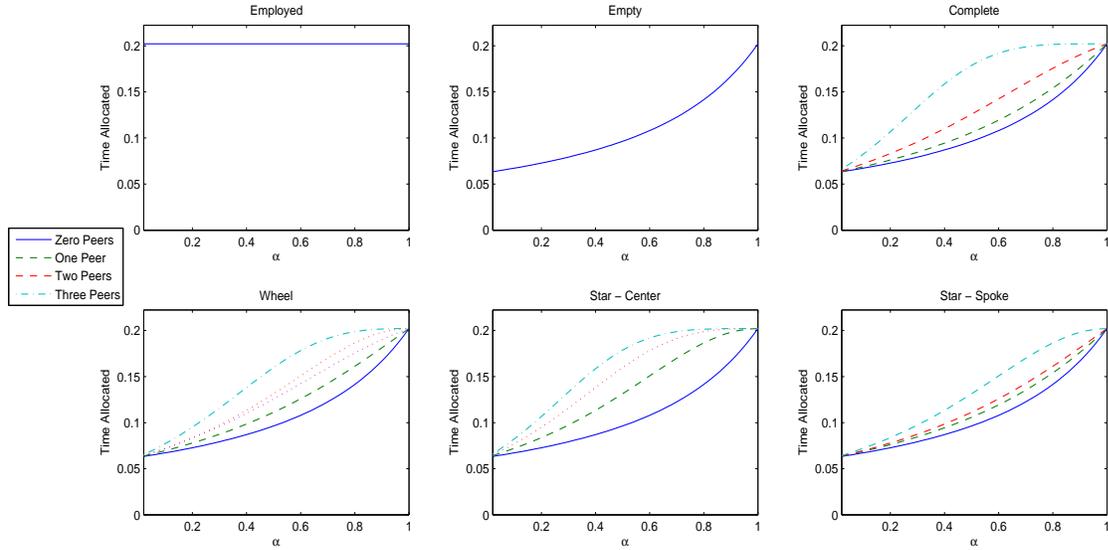


Figure 4.3: Informal Labour Supply and Arrival Probability  $\alpha$

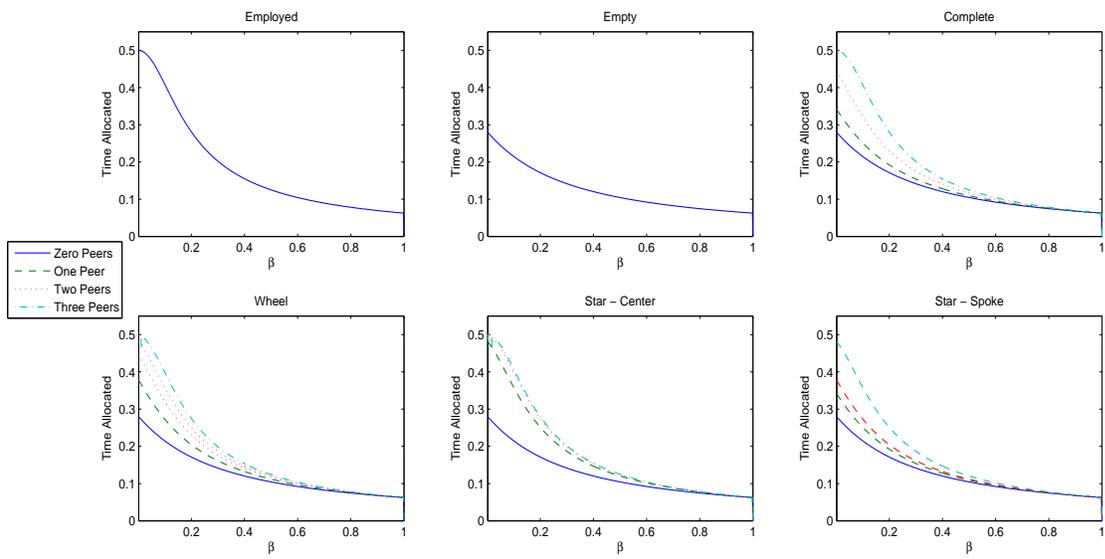


Figure 4.4: Informal Labour Supply and Probability of Detection  $\beta$

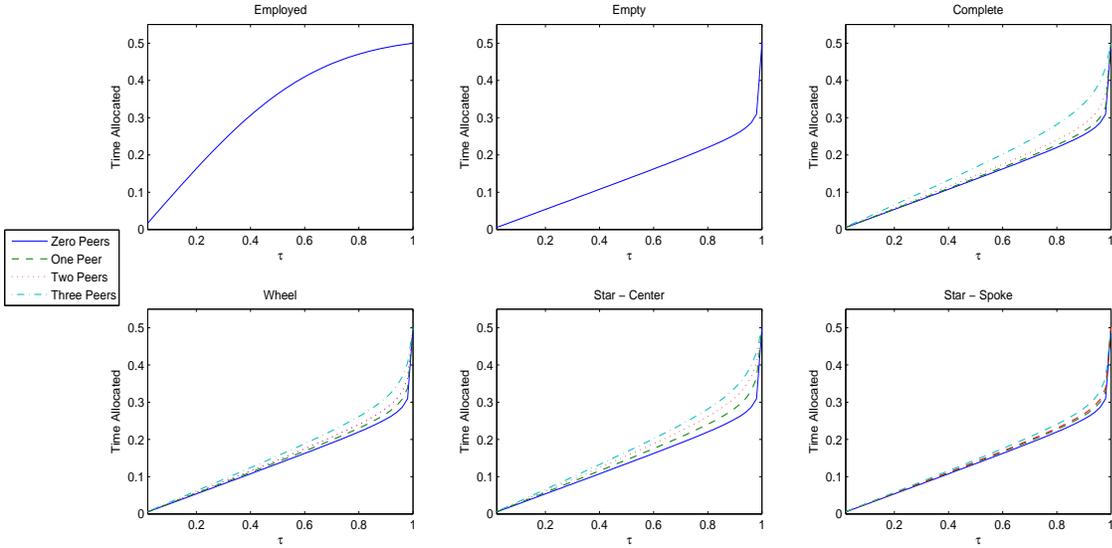


Figure 4.5: Informal Labour Supply and Income Tax  $\tau$

If a worker's peers do not have informal jobs (that is, if  $m = 0$ ), formal and informal time allocations are the same for every network. Without peers who may pass you information, every structure is equivalent to the *empty network*. As more peers become employed in the informal sector, formal workers allocate less time to formal work, and more time to informal work. In other words, for formal workers (those who start the period without informal jobs), formal labour supply is decreasing in the number of peers with an informal job, and conversely, informal labour supply is increasing in the number of peers with an informal job. If the number of peers working in the informal sector is small, a formal worker is in competition with other peers for the information about new informal job opportunities.

We also investigate the impact of changes in  $\alpha$ ,  $\beta$  and  $\tau$  on long run average total formal and informal income, and utility for each of the network structures. Total and formal incomes and utility are decreasing in income tax  $\tau$ , while the informal labour income increases as the income tax rate increases. As taxes increase, workers allocated less time to the formal sector, and earn less income from the time they do allocate. At the same time, they allocate more time to the informal sector, which increases their informal income. This increase does not fully replace their lost formal income, leading to a drop in total income and utility.

Income and utility are relatively constant with respect to the arrival probability  $\alpha$ . Formal income is increasing in the probability of detection. Facing an increasing probability of being detected working in the informal sector, a worker reduces his informal labour supply and increases his time allocated to formal work. On the other hand, informal income is more

sensitive to an increase in this probability and falls sharply as the chances of being caught working in the informal sector increases, reducing the worker's total income and utility. Figure 6 shows that depending on the network structure total income is not monotonically decreasing in the detection probability. This is the case for the complete and wheel networks.

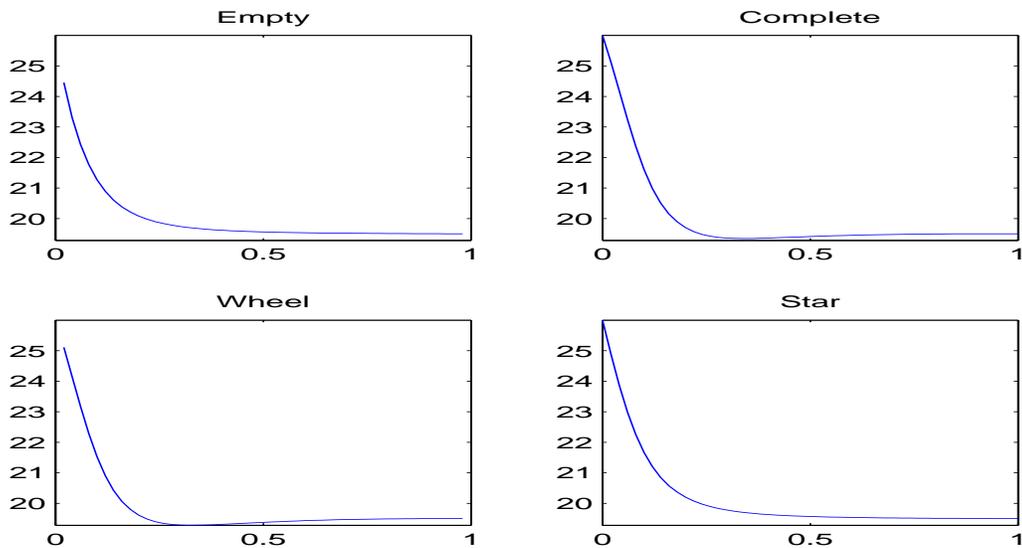


Figure 4.6: Long Run Average Total Income and Probability of Detection  $\beta$

The effects of changes in  $\alpha$ ,  $\beta$ ,  $\tau$  and network structure  $g$  on long run average income and utility depend on the interaction between changes in labour allocation, the distribution over states, and state transition probabilities. The dynamics of the economy and the importance of peer effects are especially apparent in the long run behavior of the model.

Consider, for instance, a baseline economy where the formal tax rate, job arrival probability and detection probability are as follows:  $\tau = 0.25$ ,  $\alpha = 0.10$  and  $\beta = 0.70$ , respectively. These values meant to resemble an economy with strong tax enforcement environment. Given labour allocations in each state, and the long run probabilities of those states, the average long run yearly formal income, informal income, total income, and utility are presented in Table 4.2 for each of the four networks.

Note that while total income is the highest in the *star network*, utility is the lowest. Income is the lowest in the *complete network*, but utility is the highest. To understand this counterintuitive result, compare the *complete network* when two agents have an informal job, to the *star network* when two periphery agents have an informal job. Consider a transition to a state where, again, only two agents have informal jobs. This may be accomplished in a number of ways. Either the agents who start with the informal jobs keep them, or some

Table 4.2: Long Run Averages for the baseline economy:  $\tau = 0.25$ ,  $\alpha = 0.10$ ,  $\beta = 0.70$ .

|                 | Empty   | Star    | Wheel   | Complete |
|-----------------|---------|---------|---------|----------|
| Total Income    | 19.5209 | 19.5218 | 19.4703 | 19.4806  |
| Formal Income   | 19.3671 | 19.3620 | 19.3556 | 19.3654  |
| Informal Income | 0.1538  | 0.1598  | 0.1146  | 0.1151   |
| Utility         | 22.5277 | 22.5217 | 22.5281 | 22.5281  |

of the agents without informal jobs find one, and agents who start with a job lose it. In the *star network*, the latter scenario is much more unlikely than it is in the *complete network*. Only the center agent can pass information to the two unemployed periphery agents, and only the center agent can profit from the information of the employed periphery agents. In the *complete network*, however, both unemployed agents (i.e., without an informal job) may receive information from both of their employed peers.

This change in state transition probabilities means that when the *star network* is in a state with many agents employed in the informal sector, it is very likely that those agents started the period already employed in the informal sector. Because agents who start employed in the informal sector allocate more time to that sector, expected average informal income will be higher than in the *complete network*. Due to the easier transmission of information in the *complete network*, agents who find themselves working in the informal sector may have started that period without an informal job, and so have allocated less time to informal work. But, the changes in state transition probabilities also change the long run distribution over states,  $\mu$ . The economy spends more time in states where agents have informal jobs under the *complete network* than it will under the *star network*. Average long run informal income is lower, even as workers are more likely to have informal jobs.

The overall effect is that the more connected networks (the *wheel* and the *complete network*) feature lower income but higher utility than the less connected networks (the *star* and *empty network*). The *star network*, despite being advantageous for the central agent, is little better than the *empty network* in how it transmits information.

We also investigate the impact of different institutions, such as weaker enforcement, on long run average income and utility. Consider an alternative economy, where we keep the job arrival probability and tax rate as in our baseline economy ( $\alpha = 0.10$ ,  $\tau = 0.25$ ), but the detection probability is lower ( $\beta = 0.30$ ). This captures an environment where enforcement is more lax due to corruption, weaker institutions or lack of appropriate policy instruments to fight informal activities. The average long run yearly income and utility are presented in Table 4.3 for each of the four networks.

Table 4.3: Long Run Averages for the alternative economy:  $\tau = 0.25$ ,  $\alpha = 0.10$ ,  $\beta = 0.30$ .

|                 | Empty   | Star    | Wheel   | Complete |
|-----------------|---------|---------|---------|----------|
| Total Income    | 19.7412 | 19.7788 | 19.2913 | 19.3617  |
| Formal Income   | 18.1504 | 17.9429 | 17.8474 | 17.9330  |
| Informal Income | 1.5908  | 1.8360  | 1.4439  | 1.4287   |
| Utility         | 22.6323 | 22.6296 | 22.6552 | 22.6550  |

First, note that formal income is lower in the alternative economy than in the baseline economy, informal income is higher, and utility is higher, for any network. This is because in the alternative economy, due to lax enforcement, workers allocate more time to informal labour and less time to formal labour. Because they face a smaller chance of detection in the informal sector, their utility must necessarily rise in equilibrium. Since detection is still possible, however, the informal sector is still risky, and the lost income from the formal sector is not fully replaced by the work performed in the informal sector, and total income may fall. In the *wheel* and *complete network*, this is what occurs. For precisely the same reason that income is lower in the *complete network* than in the *star*, while utility is higher, income is lower in the alternative economy for the *wheel* and *complete* networks, while utility is higher. Lower detection probabilities make it easier for agents with jobs to keep them, but do not make it easier for agents without jobs to find them. And, in any given state, it is still likely that those who end up with informal jobs are those who started with them. Thus, a more lax enforcement environment, while associated with more time allocated to a higher earning sector, may lead to less income.

Long run utility rises not only due to the greater ease of obtaining informal work, and the resulting higher income that may be earned by a worker without an informal job,<sup>9</sup> but also due to a greater ability to smooth labour choices, and enjoy leisure. In states with the greatest information transmission, the *complete* and *wheel network*, agents face the least uncertainty in their labour decision. The change in utility from the baseline to the alternative economy is the largest in the *wheel* and *complete network*, and the smallest in the *star* and *empty network*. Weaker institutions matter more in more connected social networks. In the *wheel* and *complete network*, agents are more able to benefit from lax enforcement of labour laws than in the *star* and *empty* networks.

---

<sup>9</sup>As was shown previously, long run *average* income falls, but for a particular agent without an informal, easier job transmission will *increase* her expected income.

### 4.3.4 Government Policy Instruments and Network Structure

In the analysis to this point, tax and enforcement policy instruments have been treated as given. But as policy instruments, the government may choose them optimally to maximize agents' welfare. In this section, we allow the social planner to choose the best combination of tax and enforcement instruments  $(\tau, \beta)$  that maximize social welfare, defined here as the sum of agents' utility in this economy.

Let  $(c_i(\tau, \beta), l_i(\tau, \beta), \gamma_i(\tau, \beta))$  denote the solution of a social planner's problem for a fixed tax-enforcement policy  $(\tau, \beta)$  and  $U(\tau, \beta)$  denotes the corresponding values of the utility function. Consider the following problem of optimal tax-enforcement policy:

$$\max_{\tau, \beta} \sum_i \mathbb{E}_S[U(c_i(\tau, \beta), l_i(\tau, \beta), \gamma_i(\tau, \beta))].$$

Let  $(\tau^*, \beta^*)$  and  $U^* = U(\tau^*, \beta^*)$  be the solution of this problem and the corresponding value of the utility function, respectively. We introduce functions  $\tau(\beta)$  and  $\beta(\tau)$  as follows. For any fixed  $\tau \in [0, 1]$ , define  $\beta(\tau)$  to be the best level of detection probability, i.e.

$$\max_{\beta} \sum_i \mathbb{E}_S[U(\tau, \beta)] = \sum_i \mathbb{E}_S[U(\tau, \beta(\tau))].$$

And, similarly, for any fixed  $\beta \in [0, 1]$ , define  $\tau(\beta)$  to be the best level of tax rate, i.e.

$$\max_{\tau} \sum_i \mathbb{E}_S[U(\tau, \beta)] = \sum_i \mathbb{E}_S[U(\tau(\beta), \beta)].$$

These functions satisfy  $\beta(\tau^*) = \beta^*$  and  $\tau(\beta^*) = \tau^*$ .

We characterize numerically the optimal tax-enforcement policy  $(\tau^*, \beta^*)$  when the government has to raise an exogenously given revenue and the cost of auditing the informal sector varies. Let  $\beta A$  the cost of auditing, where  $A$  is a fixed unit cost of audits. This cost can be interpreted as the government using some of the tax revenue raised from the formal sector in enforcement. If tax-enforcement is costless ( $A = 0$ ), the optimal policy is to set the detection probability to its maximum level, i.e.,  $\beta^* = 1$ , and the formal tax rate to the minimum amount needed to raise the required revenue. In this case, given that auditing is costless and formal taxation is distortionary, the optimal policy would require taxing inelastic variables more heavily and so a low  $\tau$  - high  $\beta$  policy is welfare enhancing. This result is robust regardless of the network structure and workers' risk aversion. For all network structures we study, the optimal detection probability is  $\beta^* = 1$ . That is, it is socially optimal for the

planner to shut down the informal sector.

Figure 7 shows the social welfare  $U(\tau, \beta)$  for various levels of the tax rate  $\tau$  and probability of detection  $\beta$  when the audit cost is zero ( $A = 0$ ), for the *empty* network. In any network, for different levels of expected revenue, there are different combinations of  $(\tau, \beta)$  that raise that expected revenue, but only one tax-enforcement policy pair that maximizes welfare, i.e., the optimal combination  $(\tau^*, \beta^*)$ . In the figure we highlight the set of  $(\tau, \beta)$  that are optimal for different levels of revenue for the *empty* network. For instance, to raise 0.2 in expected revenue,  $(\tau^*, \beta^*) = (0.1, 1.0)$  is the optimal policy (this  $(\tau^*, \beta^*)$  pair is indicated by the arrow in Figure 7).

However, when audit costs are positive ( $A > 0$ ), auditing everybody is no longer optimal. In fact, the optimal detection probability  $\beta$  declines and the optimal tax rate  $\tau$  increases as the audit cost rises. In this case, for an *empty* network with audit costs of  $A = 0.05$ , the optimal policy is to set  $\tau^* = 0.12$  and  $\beta^* = 0.98$ , if the planner has to raise an expected revenue of 0.20. If auditing is costly, the agent's risk aversion is relevant for the determination of the optimal tax-enforcement policy. If agents are risk averse, a small probability of detection would be enough to discourage agents from working in the informal sector. This allows the planner to set a smaller  $\beta$  and a lower  $\tau$ , in order to maximize the social welfare. On the other hand, if agents are risk neutral, a larger tax rate and a higher detection probability are the optimal policy.<sup>10</sup>

## 4.4 Conclusions

The importance of social ties to labour market outcomes has been long understood. This paper formally models the effect of network structure on time allocation and informal work. It is flexible enough to accommodate many different social structures, but simple enough to generate empirical predictions. In our social network model, we explore how peer effects interact with two well emphasized determinants of informal activities - tax burden and institutional quality. When enforcement is weaker, social ties matter more and workers are more responsive to the strength of social ties. The informal job information transmission process is more important in an environment with weaker enforcement. The institutional environment determines the importance of peer effects for labour time allocation. Workers allocate more time to informal activities in the presence of lax enforcement and better job

---

<sup>10</sup>Results for the other networks are available from the authors upon request.

information transmission. Further analysis of the different network structures may provide insights into different social institutions. Features such as social norms and fairness could be easily introduced in our analysis, for instance, via frictions in the transmission of information about informal activities opportunities. We leave this for future research.

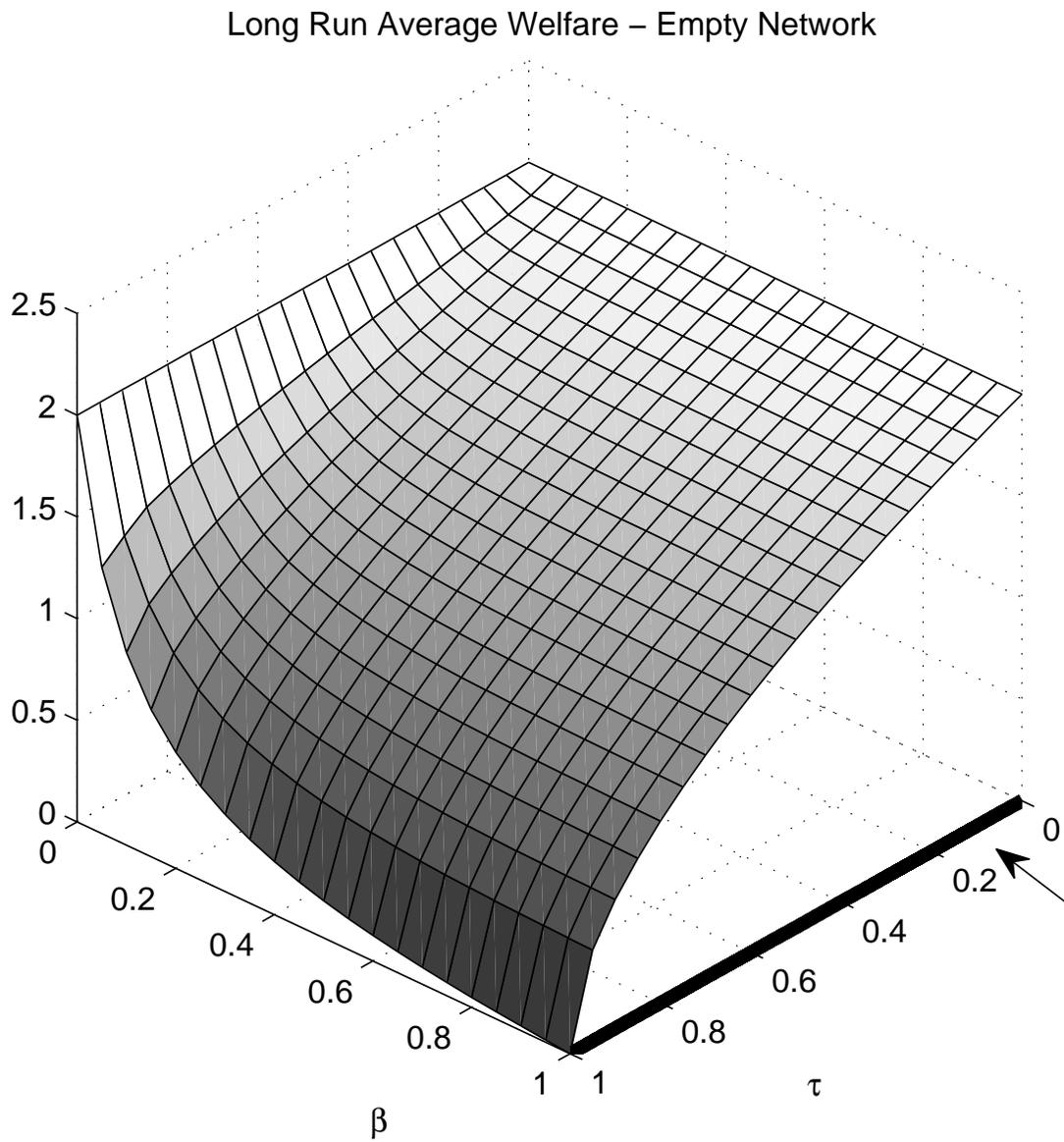


Figure 4.7: Long Run Average Welfare in the Empty Network

# APPENDIX A

## PROOFS

[Proofs]

### Proof of Proposition 1

There are three cases to consider.

1.  $\sum_i \theta_i > c$ . Suppose that a graph  $g$  is efficient, and that it is not a minimal connected graph. First, suppose it is not minimal. Then there is some component  $C$  of  $g$  that contains a link that could be deleted without disconnecting the component. Because utility does not decay across links, the value received by the agents in the component would not decrease when this link is deleted, but the agent who is initiating it would see her costs decrease. This would increase utility, contradicting the network being efficient. Now suppose the graph is not connected. Then let there be  $m \geq 2$  components of the graph, that are disconnected. Let  $C_1, \dots, C_m$  be these components, and let  $n_1, \dots, n_m$  be the number of agents in each one. Consider connecting these components, each via a single link. The change in utility is

$$\sum_{j=1}^m \left( \sum_{k \neq j} n_k \right) \sum_{i \in C_j} \theta_i - (m-1)c. \quad (\text{A.1})$$

That is, each agent  $i$  in component  $C_j$  now has a path to every agent in the other components. There are  $\sum_{k \neq j} n_k$  such agents, and  $i$  receives  $\theta_i$  for each one. This is then summed over components. Let  $n_{max}$  be the size of the largest component. Then,

from (A.1),

$$\begin{aligned}
\left(\sum_{j=1}^m \left(\sum_{k \neq j} n_k\right) \sum_{i \in C_j} \theta_i\right) - (m-1)c &= \left(\sum_{j=1}^m (n - n_j) \sum_{i \in C_j} \theta_i\right) - (m-1)c \\
&= \left(\sum_{j=1}^m \sum_{i \in C_j} n \theta_i\right) - \left(\sum_{j=1}^m \sum_{i \in C_j} n_j \theta_i\right) - (m-1)c \\
&= n \sum_i \theta_i - \left(\sum_{j=1}^m n_j \sum_{i \in C_j} \theta_i\right) - (m-1)c.
\end{aligned}$$

The size of component  $C_j$ ,  $n_j$ , is necessarily less than  $n_{max}$ . The maximum size of  $n_{max}$  is  $n - (m - 1)$ . This is largest possible component if there are  $m$  components. Therefore,  $n_j \leq n - (m - 1)$ . We therefore have

$$\begin{aligned}
n \sum_i \theta_i - \left(\sum_{j=1}^m n_j \sum_{i \in C_j} \theta_i\right) - (m-1)c &\geq n \sum_i \theta_i - \left(\sum_{j=1}^m (n - (m - 1)) \sum_{i \in C_j} \theta_i\right) - (m-1)c \\
&= n \sum_i \theta_i - (n - (m - 1)) \sum_i \theta_i - (m-1)c + nc - nc \\
&= n \left(\sum_i \theta_i - c\right) - (n - (m - 1)) \left(\sum_i \theta_i - c\right) > 0.
\end{aligned}$$

The final inequality follows because  $\sum_i \theta_i > c$  and  $n > n - (m - 1)$ . Connecting these components increased total utility, so the original graph  $g$  could not have been efficient.

Conversely, if a graph  $g$  is minimal and connected, it must be efficient. It cannot be that adding links increase utility, as this does not change the value any agent receives from the network, and only increases linking costs. Deleting a link will necessarily disconnect the network, since it is minimal, and the argument above shows that any disconnected network is not efficient. Therefore, a network is efficient if and only if it is a minimal connected graph.

2.  $\sum_i \theta_i = c$ . If  $\sum_i \theta_i = c$ , the total utility for a minimal connected graph is zero, as is the value of an empty graph. To see that such a graph is efficient, suppose a graph  $g$  is efficient, and is neither empty nor a minimal connected graph.  $g$  must contain a non-empty component  $C$  with  $m \leq n$  agents, and from the no decay assumption,  $C$  is

connected with exactly  $m - 1$  links. The total utility of the agents in  $C$  is therefore

$$(m - 1)\left(\sum_{i \in C} \theta_i - c\right) < (m - 1)\left(\sum_i \theta_i - c\right) = 0,$$

so that aggregate utility is less than either the empty network or a minimal connected network. Thus,  $g$  could not have been efficient. Conversely, this argument shows that if a graph is either empty or a minimal complete graph, it is efficient.

3.  $\sum_i \theta_i < c$ . If  $\sum_i \theta_i < c$ , a network is efficient if and only if it is empty. Suppose a network is efficient, but not empty. Then there is a component  $C_j$  with  $n_j > 1$  members. The aggregate utility of all the agents in this component is at most

$$(n_j - 1) \sum_{i \in C_j} \theta_i - (n_j - 1)c.$$

This can be simplified to

$$\begin{aligned} (n_j - 1) \sum_{i \in C_j} \theta_i - (n_j - 1)c &= (n_j - 1) \left( \sum_{i \in C_j} \theta_i - c \right) \\ &< (n_j - 1) \left( \sum_i \theta_i - c \right) < 0, \end{aligned}$$

where the last inequality follows by hypothesis. Thus no efficient graph can be nonempty. In contrast, if a graph is empty, it must be efficient. Adding a link would create a non-trivial component, and the argument above shows that any non-trivial component is necessarily inefficient. Therefore a network is efficient if and only if it is the empty graph.

### Proof of Proposition 3

I exhaustively characterize the equilibria of the 3 agent case in five lemmas. I establish in Lemma 1 that the only equilibria in which agents use mixed strategies are the two hybrid equilibria, and the full mixing equilibrium. In Lemma 2, I show that if agents use only pure strategies, then the only equilibrium in which all three agents, for some positive measure of types, form only a single link is the symmetric circle equilibrium. In Lemma 3, I show that if agents use only pure strategies, then the only equilibrium in which exactly two agents form a single link, for some positive measure of types, is the periphery-sponsorship equilibrium.

In Lemma 4, I show that if agents use only pure strategies, then there is no equilibrium in which only a single agent forms a single link for some positive measure her types. In Lemma 5, I show that the only equilibria in which agents use only pure strategies and no agent ever forms only a single link, are the various center-sponsorship equilibria. This exhaustively characterizes all the equilibria of the three agent case with uniform values.

It is easy to see that equilibrium strategies must be monotone in types; if it a best response for an agent of type  $\theta$  to form two links, it cannot be a best response for the same agent, when her type is  $\theta' > \theta$ , to form only one link. This implies the existence of critical types, cutoffs where agents are indifferent between forming zero or one link, and one or two links. In some strategies, agents may form zero links for a low range or types, one link for an intermediate range, and two links for a higher range or her type. We will say an agent's strategy *places positive weight* on an action if for some positive measure of her type space she performs that action.

Let  $\delta_i^{01}$  denote the cutoff used by agent  $i$  between forming zero links, and one link, and  $\delta_i^{12}$  denote the cutoff used by agent  $i$  between forming one link and two links; that is, these are the critical types of an agent who in equilibrium must be indifferent between those two actions. Because types are drawn from the uniform distribution, the ex ante probability agent  $i$  forms zero links is  $F(\delta_i^{01}) = \delta_i^{01}$ , the probability she forms 1 link is  $F(\delta_i^{12}) - F(\delta_i^{01}) = \delta_i^{12} - \delta_i^{01}$ , and the probability she forms two links is  $1 - F(\delta_i^{12}) = 1 - \delta_i^{12}$ . Let  $p_{ij}$  be the probability that agent  $i$  links to agents  $j$ , when agent  $i$  is mixing in forming a single link to either  $j$  or  $k$ . By definition,  $p_{ik} = 1 - p_{ij}$ , so in what follows, only one of these variables will be referred to. I will only refer to  $p_{ij}$ ,  $p_{jk}$  and  $p_{ki}$ .

**Lemma 1** *The only equilibria in which agents mix are hybrid equilibria, one in which the nonmixing agents place positive weight on forming two links, and one in which she does not, and the full mixing equilibria, in which all three agents mix when forming their single link.*

**Proof.**

**Step 1.** I first characterize the mixing equilibrium, when all three players mix. If all three players mix, then 6 indifference conditions must be satisfied. Each agent must be indifferent between their actions when mixing, forming a single link to one or the other agent, and there is a cutoff value for each agent where she is indifferent between zero links and one link. This

leads to the following system of equations:

$$\begin{aligned}
1 - \delta_k^{01} + (1 - \delta_j^{12}) \delta_k^{01} + (-\delta_j^{01} + \delta_j^{12}) \delta_k^{01} p_{jk} &= 1 - \delta_j^{01} + \delta_j^{01} (1 - \delta_k^{12}) + \delta_j^{01} (-\delta_k^{01} + \delta_k^{12}) (1 - p_{ki}), \\
1 - \delta_i^{01} + \delta_i^{01} (1 - \delta_k^{12}) + \delta_i^{01} (-\delta_k^{01} + \delta_k^{12}) p_{ki} &= 1 - \delta_k^{01} + (1 - \delta_i^{12}) \delta_k^{01} + (-\delta_i^{01} + \delta_i^{12}) \delta_k^{01} (1 - p_{ij}), \\
1 - \delta_i^{01} + \delta_i^{01} (1 - \delta_j^{12}) + \delta_i^{01} (-\delta_j^{01} + \delta_j^{12}) (1 - p_{jk}) &= 1 - \delta_j^{01} + (1 - \delta_i^{12}) \delta_j^{01} + (-\delta_i^{01} + \delta_i^{12}) \delta_j^{01} p_{ij}, \\
\delta_i^{01} &= \frac{c}{(-\delta_j^{12} \delta_k^{01} + 2\delta_j^{01} \delta_k^{12} - \delta_j^{01} \delta_k^{01} p_{jk} + \delta_j^{12} \delta_k^{01} p_{jk} - 2\delta_j^{01} \delta_k^{12} p_{jk} + 2\delta_j^{12} \delta_k^{12} p_{jk} + (\delta_k^{01} - \delta_k^{12}) (\delta_j^{01} (3 - 2p_{jk}) + 2\delta_j^{12} p_{jk}) p_{ki}}, \\
\delta_j^{01} &= \frac{c}{(2\delta_i^{12} \delta_k^{01} - \delta_i^{01} \delta_k^{12} + 2\delta_i^{01} \delta_k^{01} p_{ij} - 3\delta_i^{12} \delta_k^{01} p_{ij} + \delta_j^{01} \delta_k^{01} p_{ij} + (-\delta_k^{01} + \delta_k^{12}) (\delta_i^{01} - 2\delta_i^{12} (-1 + p_{ij}) + 2\delta_i^{01} p_{ij}) p_{ki}}, \\
\delta_k^{01} &= \frac{c}{(-\delta_i^{12} \delta_j^{01} + 2\delta_i^{01} \delta_j^{12} - \delta_i^{01} \delta_j^{01} p_{ij} + \delta_i^{12} \delta_j^{01} p_{ij} - 2\delta_i^{01} \delta_j^{12} p_{ij} + 2\delta_i^{12} \delta_j^{12} p_{ij} + (\delta_j^{01} - \delta_j^{12}) (\delta_i^{01} (3 - 2p_{ij}) + 2\delta_i^{12} p_{ij}) p_{jk}}.
\end{aligned}$$

The first of these equations is an indifference condition; agents of the critical types must be indifferent between forming no links and mixing in forming a single link, and agents of whatever type who mix in forming a single link must be indifferent among whom she may form that link to. There are many solutions to this system that are consistent with equilibrium. For example, there is a symmetric equilibrium, in which each agent uses the same mixing probability and the same cutoff. In this case the mixing probabilities are each  $\frac{1}{2}$ , and the equilibrium cutoff is

$$\delta^{01} = \frac{(9c + \sqrt{3}\sqrt{1 + 27c^2})^{1/3}}{3^{2/3}} - \frac{1}{3^{1/3} (9c + \sqrt{3}\sqrt{1 + 27c^2})^{1/3}}.$$

Agents form a single link at random if their value exceeds this cutoff, and form no links if their value is below it. It is not possible for any agent to place positive weight on forming two links in these equilibria. If any agent does, then the others will prefer to form two links themselves rather than one; that is, they will prefer to deviate to center-sponsored equilibrium.

**Step 2.** Next, I eliminate the possibility that only two agents mix when forming a single link, in equilibrium. By contradiction, suppose that exactly two agents do mix. When two agents are mixing, they must each be indifferent between linking to the other mixing agent, and linking to the agent who is not mixing. Suppose that  $i$  and  $j$  are mixing, and consider their indifference conditions when mixing:

$$\begin{aligned}
\delta_j^{01} &= \frac{\delta_j^{12} \delta_k^{01} (1 - p_{jk})}{\delta_k^{12} p_{ki} + \delta_k^{01} (1 - p_{jk} - p_{ki})} \\
\delta_i^{01} &= \frac{\delta_i^{12} \delta_k^{01} p_{ij}}{\delta_k^{01} p_{ij} - \delta_k^{01} p_{kj} + \delta_k^{12} p_{kj}}
\end{aligned}$$

Because by hypothesis agent  $k$  is not mixing, then either  $p_{ki} = 0$ ,  $p_{jk} = 0$ , or both. That is,

agent  $k$  can form a single link to one or the other, or not form a single link at all. But this means that at least one of  $p_{ki}$  and  $p_{jk}$  must be zero. Suppose without loss of generality that  $p_{ki} = 0$ . Substituting this into the above indifference condition,  $\delta_j^{01} = \delta_j^{12}$ , which contradicts the hypothesis that agent  $j$  places positive weight on forming a single link. Therefore, this cannot be an equilibrium.

**Step 3.** Finally, suppose only one agent uses mixed strategies in equilibrium. Without loss of generality, let this be agent  $i$ . I will show that the architecture of this equilibrium must be the hybrid structure. In this structure all three agents place positive weight on forming a single link. Agent  $i$  mixes in forming her single link to the others, while agent  $j$  and agent  $k$  form a single link to each another.

First, suppose that all three agents place positive weight on forming a single link. The possible arrangements of these links are a periphery-sponsorship structure, where  $k$  and  $j$  form their single link to agent  $i$ , a chain, where  $k$  links to  $j$ , who links to  $i$ , and the hybrid structure. I will first show that the periphery sponsorship structure cannot be an equilibrium. Suppose that agents  $j$  and  $k$  are both directing their single link to agent  $i$ , who is mixing between the two of them with her single link. For agent  $j$  to prefer to link to agent  $i$  over  $k$ , it must be that  $i$  does not place any weight on linking to agent  $j$ . This is because, otherwise, a link to agent  $k$  would be preferable: It would yield a certain connection to  $k$  and a possible connection to  $i$ , while a link to  $i$  yields a certain connection to  $i$  and a possible connection to  $k$ , but the likelihood of this latter connection is smaller than the possible connection to  $i$ . To see this, note that  $j$ 's expected utility from a link to  $i$  is

$$\theta_j - c + \theta_j(1 - \delta_k^{01}(\delta_i^{01} + (\delta_i^{12} - \delta_i^{01})p_{ij})),$$

while the utility from a link to agent  $k$  is

$$\theta_j - c + \theta_j(1 - \delta_k^{01}\delta_i^{01}).$$

She strictly prefers a link to  $k$ , unless either  $(\delta_i^{12} - \delta_i^{01}) = 0$ , which contradicts the hypothesis that  $i$  places positive weight on forming a single link, or  $p_{ij} = 0$ , which contradicts the hypothesis that  $i$  is mixing.

To eliminate the chain architecture as a possible equilibrium, suppose without loss of generality that agent  $k$  forms her link to agent  $j$ , who forms her link to agent  $i$ , who mixes. For  $i$  to be indifferent between linking to either of the other two agents, it must be that  $\delta_j^{01} = 1$ , which contradicts our hypothesis that agent  $j$  places positive weight on forming one

link. To see this, note that the expected utility to agent  $i$  of a link to agent  $j$  is

$$\theta_i - c + \theta_i(1 - \delta_k^{01}),$$

and the expected utility from a link to  $k$  is

$$\theta_i - c + \theta_i(1 - \delta_k^{01}\delta_j^{01}).$$

The result is immediate. This holds whether the other agents form one or two links. Therefore the chain architecture cannot be an equilibrium; this leaves only the hybrid, where two agents form a single link to each other, and the third mixes in his single link to them.

**Step 4.** To conclude, I characterize the hybrid equilibria. Let agent  $i$  be the mixing agent. I now show that all three agents cannot place positive weight on forming two links. Suppose otherwise; such an equilibrium must satisfy the following indifference conditions, corresponding to each agent's indifference between zero and one link, one and two links, and the mixing agent's indifference:

$$\begin{aligned} \delta_i^{01} &= \left( \frac{c}{2\delta_j^{12}\delta_k^{12}} - \delta_j^{01}\delta_k^{01} \right), \\ \delta_i^{12} &= \frac{c}{\delta_j^{01}\delta_k^{01}}, \\ \delta_j^{01} &= \left( \frac{c}{\delta_i^{12}\delta_k^{01}p_{ik} - \delta_i^{01} + \delta_i^{12}\delta_k^{01} + \delta_i^{01}\delta_k^{12} - \delta_i^{01}\delta_k^{01}p_{ik}} \right), \\ \delta_k^{01} &= \frac{c}{-\delta_i^{01} - \delta_j^{01} + 2\delta_i^{12}\delta_j^{01} + \delta_i^{01}\delta_j^{12} + \delta_i^{01}\delta_j^{01}p_{ik} - \delta_i^{12}\delta_j^{01}p_{ik}}, \\ \delta_j^{12} &= \frac{c}{\delta_i^{01}}, \\ \delta_k^{12} &= \frac{c}{\delta_i^{01}}. \end{aligned}$$

This system of equations has three candidate solutions.<sup>1</sup> First, there is one that requires that  $\delta_k^{01} = 1$ , which contradicts our hypothesis that agent  $k$  is forming links, and cannot be an equilibrium. Second, there is one that requires that agent  $i$ 's mixing probabilities be imaginary, which also cannot be an equilibrium.

In the third candidate solution, I now show that there is no mixing probability  $p_{ik}$  that satisfies the solution of this system, and is consistent with equilibrium. This solution requires

---

<sup>1</sup>This was determined using Mathematica.

that

$$\delta_k^{01} = \frac{(1 - p_{ik})}{(1 - 2p_{ik} + 2p_{ik}^2)}.$$

This requires that  $p_{ik}$  be at least  $\frac{1}{2}$ , in order for  $\delta_k^{01}$  to be less than one. The solution also requires that either

$$\delta_j^{01} = \frac{1}{\delta_k^{01}}$$

or

$$\delta_j^{01} = 1 - \delta_k^{01} + 2\delta_k^{01} p_{ik}$$

which both require that  $p_{ik}$  be at most  $\frac{1}{2}$ , so that  $\delta_j^{01}$  be less than one. So  $p_{ik} = \frac{1}{2}$ , but at this value  $\delta_j^{01} = \delta_k^{01} = 1$ , which contradicts the hypothesis that a positive measure of types of both agents  $j$  and  $k$  form a single link. Hence, this solution cannot be an equilibrium. Therefore, there is no solution to this system of equations that is consistent with equilibrium.

Therefore, it cannot be that all three agents place positive weight on forming two links. Suppose now that only agent  $i$ , who is mixing, places positive weight on forming two links, and the others do not. Such an equilibrium must satisfy the following indifference conditions:

$$\begin{aligned} \delta_j^{01} &= \frac{c}{\delta_k^{01} (\delta_i^{12} - \delta_i^{01} p_{ik} + \delta_i^{12} p_{ik})}, \\ \delta_k^{01} &= \frac{c}{\delta_j^{01} (-\delta_i^{01} + 2\delta_i^{12} + \delta_i^{01} p_{ik} - \delta_i^{12} p_{ik})}, \\ \delta_i^{01} &= \left( \frac{c}{2 - \delta_j^{01} \delta_k^{01}} \right), \\ \delta_i^{12} &= \frac{c}{\delta_j^{01} \delta_k^{01}}. \end{aligned}$$

This system of equations has only one solution, which requires that  $\delta_i^{01} = \delta_i^{12} = c$ . This contradicts our hypothesis that agent  $i$  places positive weight on forming one link, so this solution cannot be an equilibrium.

It is clear that it cannot be that agent  $i$  that places positive weight on forming two links, while only one of agents  $j$  or  $k$  do; agent  $i$  will strictly prefer to link to the other agent, rather than mix. Suppose next that either agent  $j$  or  $k$ , and possibly both, place positive weight on forming two links, while the mixing agent does not. Then an equilibrium must

satisfy the following equations:

$$\begin{aligned}\delta_i^{01} &= \frac{c}{2(\delta_j^{12})^2 - (\delta_j^{01})^2}, \\ \delta_j^{01} &= \delta_k^{01} = \frac{-2c}{-3\delta_j^{01} + 2\delta_i^{01} + \delta_j^{01}\delta_i^{01} - 2\delta_j^{12}\delta_i^{01}}, \\ \delta_j^{12} &= \delta_k^{12} = \frac{c}{\delta_i^{01}}.\end{aligned}$$

This requires that both agents  $j$  and  $k$  place positive weight on forming two links. There are multiple solutions to this system that represent equilibria. This characterizes the hybrid equilibria where the non-mixing agents place positive weight on forming two links.

Suppose finally that no agent places positive weight on forming two links. The system of indifference equations such an equilibrium must satisfy, representing indifference between forming zero and one link, and indifference in mixing:

$$\begin{aligned}\delta_i^{01} &= \frac{c}{2 - \delta_j^{01}\delta_k^{01}}, \\ \delta_j^{01} &= \frac{c}{\delta_k^{01}(1 + p_{ik} - \delta_i^{01}p_{ik})}, \\ \delta_k^{01} &= \frac{c}{\delta_j^{01}(2 - \delta_i^{01} - p_{ik} + \delta_i^{01}p_{ik})}.\end{aligned}$$

The solution to this requires that  $p_{ik} = \frac{1}{2}$ ,  $\delta_i^{01} = \frac{6-c-\sqrt{36-36c+c^2}}{4}$  and

$$\delta_i^{01} = \frac{3\delta_j^{01}\delta_k^{01}}{4 - \delta_j^{01}\delta_k^{01}}.$$

There is a continuum of such equilibria. This characterizes the hybrid equilibria where the non-mixing agents do not place positive weight on forming two links. This shows that when all three agents place positive weight on forming one link, the only equilibrium structure with mixing is a hybrid equilibrium.

**Step 5.** I next eliminate the cases where only one or two agents place positive weight on forming one link. I must show that when only one agent is mixing, it cannot be that only two agents, the mixer and one other, are placing positive weight on forming one link in equilibrium. Suppose first that agent  $j$ , the non-mixing agent, is placing positive weight on forming a single link, and this link is formed to agent  $k$ . I now show that none of these agents may place positive weight on forming two links in equilibrium. If agent  $i$  places

positive weight on forming two links, the following indifference conditions, corresponding to indifference for critical types and indifference in mixing, must hold:

$$\begin{aligned}\delta_i^{01} &= \frac{c}{(2\delta_j^{12} - \delta_j^{01}) \delta_k^{02}}, \\ \delta_i^{12} &= \frac{c}{\delta_j^{01} \delta_k^{02}}, \\ \delta_j^{01} &= \frac{c}{\delta_k^{02} (\delta_i^{12} - \delta_i^{01} p_{ik} + \delta_i^{12} p_{ik})}.\end{aligned}$$

These imply that  $\delta_j^{12} = \delta_j^{01}$ , a contradiction. If  $j$  places positive weight on forming two links, the following indifference conditions must hold:

$$\begin{aligned}\delta_i^{01} &= \frac{c}{(2\delta_j^{12} - \delta_j^{01}) \delta_k^{02}}, \\ \delta_i^{12} &= \frac{c}{\delta_j^{01} \delta_k^{02}}, \\ \delta_j^{12} &= \frac{c}{\delta_i^{01} \delta_k^{02}}.\end{aligned}$$

These also imply that  $\delta_j^{12} = \delta_j^{01}$ . Using this, we write the indifference conditions that must be satisfied if agent  $k$  places positive weight on two links:

$$\begin{aligned}\delta_i^{01} &= \frac{c}{(2 - \delta_j^{01}) \delta_k^{02}}, \\ \delta_j^{01} &= \frac{c}{\delta_k^{02} (1 - \delta_i^{01} p_{ik} + p_{ik})}, \\ \delta_k^{02} &= \frac{2c}{\delta_i^{01} (1 + \delta_j^{01} p_{ik}) - \delta_i^{012} - \delta_j^{01} (-2 + p_{ik})}.\end{aligned}$$

There are three candidate solutions to this system of equations, none of which can be an equilibrium. In the first,  $p_{ik} = -1$ , which is not an equilibrium. In the second,  $p_{ik} = 1$  or  $0$ , which contradicts agent  $i$  mixing. In the third, either  $\delta_j^{01} = 1$ , which is a contradiction, or  $\delta_j^{01} = \frac{1-5p_{ik}}{3(p_{ik}-1)p_{ik}}$ , which does not depend on  $c$ , and so cannot be an equilibrium for sufficiently high  $c$ . Therefore all agents place zero weight on forming two links. But in this case, agent  $k$  prefers to deviate and form a link to agent  $i$ . To see this, note that the indifference conditions by  $i$  and  $j$  imply

$$\delta_i^{01} = \frac{c}{(2 - \delta_j^{01})} \text{ and } \delta_j^{01} = \frac{c}{(1 - \delta_i^{01} p_{ik} + p_{ik})}.$$

If agent  $k$  is of type 1, that is, the highest value, then for every value of  $c$  and mixing

probability  $p_{ik}$  by  $i$ , forming one link to  $i$  yields greater utility than forming zero links. To see this, I calculate agent  $k$ 's incremental utility from forming one link to  $i$  rather than zero links, using the fact that none of the other agents place positive weight on forming two links, and the indifference conditions of agents  $i$  and  $j$ . Using these, one can show that agent  $k$ 's incremental utility from forming one link is

$$\frac{1}{8p_{ik}^2(1+p_{ik})}(-c^2(-1+p_{ik})^2(1+p_{ik})-2(-1-3p_{ik}+6p_{ik}^2)\left(-2-2p_{ik}+\sqrt{c^2(-1+p_{ik})^2+4(1+p_{ik})^2-4c(1+p_{ik})^2}\right)+c\left(4-\sqrt{c^2(-1+p_{ik})^2+4(1+p_{ik})^2-4c(1+p_{ik})^2}+p_{ik}\left(10+p_{ik}\left(4-34p_{ik}+\sqrt{c^2(-1+p_{ik})^2+4(1+p_{ik})^2-4c(1+p_{ik})^2}\right)\right)\right)).$$

This expression is positive for all feasible  $p_{ik}$  and  $c$  less than 1. Agent  $k$  will prefer to deviate and form a link to agent  $i$ , and complete the circle architecture, rather than form zero links. Therefore, this cannot be an equilibrium.

Now suppose instead that agent  $j$  directs her single link to agent  $i$ , the mixing agent, rather than to agent  $k$ . Agent  $i$ 's utility from forming a link to agent  $j$  is

$$\theta_i - c + \theta_i(1 - \delta_k^{02}\delta_j^{12}),$$

and her utility from forming a link to agent  $k$  is

$$\theta_i - c + \theta_i(1 - \delta_k^{02}\delta_j^{01}).$$

In order to mix between the two, she must be indifferent between them: This requires that  $\delta_j^{12} = \delta_j^{01}$ , which contradicts the hypothesis that agent  $j$  places positive weight on forming one link. Hence, this cannot be an equilibrium.

**Step 6.** Finally, I show that there cannot be an equilibrium in which only one agent mixes, and is the only agent placing positive weight on forming one link. This will complete the lemma.

To see this, first note that if agents  $j$  and  $k$  are not placing positive weight on forming one link, then agent  $i$  must not place positive weight on forming two links. Otherwise, the only solution to the indifference conditions for critical types of agent  $i$  would require that he place zero weight on forming one link, contradicting the hypothesis that she forms one link. To see this, note that the indifference conditions for the two cutoffs are identical.

$$\delta_i^{01} = \frac{c}{\delta_j^{02}\delta_k^{02}} = \delta_i^{12},$$

Therefore, agent  $i$  cannot place any weight on forming 2 links in equilibrium. Now suppose that agent  $j$  places positive weight on forming 2 links. This cannot be an equilibrium. To see this, note the cutoff indifference conditions for agent  $i$  and agent  $j$ :

$$\delta_i^{01} = \frac{c}{\delta_j^{02} \delta_k^{02}}, \text{ and } \delta_j^{02} = \frac{2c}{\delta_k^{02} (1 + \delta_i^{01} + p_{ik} - \delta_i^{01} p_{ik})}.$$

There are two candidate solutions to these equations. One requires that  $p_{ik} = -1$ , and cannot be an equilibrium, while the other requires that  $\delta_i^{01} = 1$ , which contradicts the hypothesis that  $\delta_i^{01} < 1$ . A similar calculation shows that agent  $k$  cannot place positive weight on two links in equilibrium.

Hence, no agent can place positive weight on two links. It is easy to show now that there is no equilibrium of this form at all. Such an equilibrium requires that  $\delta_i^{01} = c$ . But then both  $i$  and  $j$  would prefer to deviate and form links if their value  $\theta$  is sufficiently high. For example, if either agent has value  $\theta = 1$ , then they strictly prefer to form two links rather than zero.

■

This establishes that the only equilibria in which agents mix are the hybrid and mixing equilibria. Therefore, for the following lemmas, assume that agents use only pure strategies.

**Lemma 2** *The only equilibrium in which every agent places positive weight on forming one link, when there is no mixing, is the symmetric circle equilibrium.*

**Proof.**

**Step 1.** I must first show that if each agent places positive weight on forming a single link in equilibrium, then these links must be arranged in the the circle architecture. Suppose that agent  $j$  is forming a link to agent  $k$ , who is forming a link to agent  $i$ . Agent  $i$ 's expected utility from forming a link to  $j$  is

$$\theta - c + (1 - \delta_j^{01} \delta_k^{01})\theta.$$

Her expected utility from a single link to agent  $k$  is

$$\theta - c + (1 - \delta_j^{01})\theta.$$

Since by hypothesis  $\delta_k^{01} < 1$ , forming a single link to  $j$  is strictly preferred by agent  $i$  to forming a single link to  $k$ . That is, she prefers to complete the circle.

Suppose next that agents  $j$  and  $k$  each form a single link to one another. Then agent  $i$  strictly prefers to link to whomever  $k$  is less likely to link. I have eliminated the cases where  $i$  mixes, so suppose  $i$  links to  $j$ . Then by the argument above, this cannot be an equilibrium, because  $k$  prefers to link to  $i$  and complete the circle, rather than link to  $j$ . Thus, the three links must be arranged in a circle.

**Step 2.** I now show that that if all three agents place positive weight on forming a single link, they must place zero weight on forming two links. Suppose not; let agent  $i$  put positive weight on two links. Agent  $i$ 's indifference condition between forming one and two links at her critical type is:

$$\delta_i^{12} \delta_k^{01} \delta_j^{01} = c.$$

Consider now agent  $j$ 's indifference condition between forming zero and one link: Her critical type:  $\delta_j^{01}$  satisfies:

$$\delta_j^{01} (\delta_i^{01} (1 - \delta_k^{01}) + 3\delta_i^{01} \delta_k^{01} + 2(1 - \delta_i^{12} - \delta_k^{01})) = c.$$

Together, these imply

$$(1 - \delta_i^{12})(\delta_i^{01} + 2(1 - \delta_k^{01})) = 0.$$

This requires that either  $(1 - \delta_i^{12}) = 0$ , which contradicts that agent  $i$  places positive weight on two links, or  $(\delta_i^{01} + 2(1 - \delta_k^{01})) = 0$ , which implies that either  $\delta_i^{01}$  is negative, an impossibility, or that  $\delta_i^{01} = 0$  and  $\delta_k^{01} = 1$ . The former would give agent  $i$  negative utility for  $\theta = 0$ , and the latter contradicts the hypothesis that agent  $k$  puts positive weight on forming a single link. Therefore, agent  $i$  could not have placed positive weight on forming a single link. Note that this argument holds whether or not agents  $j$  and  $k$  were putting positive weight on forming two links or not; therefore, by symmetry, no agent can put positive weight on forming two links.

**Step 3.** Next, I show that when each agent places positive weight on forming a single link, the cutoffs they use are the same: The equilibrium must be symmetric. Consider the indifference cutoffs between forming 0 links and forming 1. The equilibrium cutoffs satisfy the following system of equations

$$\begin{aligned} \delta_i^{01} (\delta_j^{01} + 2\delta_k^{01} (1 - \delta_j^{01})) &= c, \\ \delta_j^{01} (\delta_k^{01} + 2\delta_i^{01} (1 - \delta_k^{01})) &= c, \\ \delta_k^{01} (\delta_i^{01} + 2\delta_j^{01} (1 - \delta_i^{01})) &= c. \end{aligned}$$

This system simplifies to the following system:

$$\begin{aligned}\delta_j^{01}(\delta_i^{01} + \delta_k^{01}) &= 2\delta_i^{01}\delta_k^{01}, \\ \delta_k^{01}(\delta_i^{01} + \delta_j^{01}) &= 2\delta_i^{01}\delta_j^{01}.\end{aligned}$$

Solving, eliminate  $\delta_i^{01}$  and see that

$$\delta_j^{01} = \delta_k^{01} \doteq \delta,$$

which immediately yields that  $\delta_i^{01} = \delta$ . ■

This shows that the symmetric circle equilibrium is the only architecture where agents do not mix and each places positive weight on forming a single link.

**Lemma 3** *The only equilibrium in which two agents place positive weight on forming a single link, when mixing is not allowed, is the periphery-sponsorship equilibrium.*

**Proof.** Without loss of generality, let agents  $i$  and  $j$  be the those placing positive weight on forming a single link. All 3 agents may or may not be placing positive weight on 2 links. I will show that the single links agents  $i$  and  $j$  are forming must both be to agent  $k$ . The other possibilities are that one they link to each other, or they forming a line ending at agent  $k$ .

**Step 1.** The first case, when agents  $i$  and  $j$  are linking each other, I eliminate because then both agents would prefer to deviate and link to agent  $k$ . To see this, consider agent  $i$ 's expected utility from linking  $j$  in this case:

$$\theta_i - c + (1 - \delta_j^{12}\delta_k^{02})\theta_i.$$

Her expected utility from linking  $k$  is

$$\theta_i - c + (1 - \delta_j^{01}\delta_k^{02})\theta_i.$$

Since  $\delta_j^{01} < \delta_j^{12}$ , her expected utility from linking  $k$  is strictly greater than her expected utility from linking  $j$ .

In the second case, suppose without loss of generality that  $i$  forms a single link to  $j$ , who forms a single link to  $k$ . I first show that it cannot be that all three agents place positive weight on forming two links. The argument is similar to those above. If they placed positive

weight on two links, then the indifference condition between one link and two links, for agents  $i$  and  $j$ , yields

$$\begin{aligned}\delta_i^{12} \delta_j^{01} \delta_k^{02} &= c, \\ \delta_j^{12} \delta_i^{01} \delta_k^{02} &= c.\end{aligned}$$

The indifference condition between zero and one link for  $i$  and  $j$  yields

$$\begin{aligned}(2\delta_j^{12} - \delta_j^{01})\delta_k^{02}\delta_i^{01} &= c, \\ (\delta_i^{01} - \delta_i^{12} - \delta_j^{01} + 2\delta_j^{12}\delta_k^{02}\delta_j^{01}) &= c.\end{aligned}$$

The indifference condition between 0 and 2 for agent  $k$  yields

$$\delta_k^{02}\delta_i^{01}\delta_j^{01} = c.$$

In the solution of this system of equations, however,  $\delta_i^{01} = \delta_i^{12}$ , which contradicts our hypothesis that agent  $i$  forms 1 link with positive probability.

**Step 2.** Suppose that only agents  $i$  and  $j$  place positive weight on two links. The indifference conditions for the critical types of agents  $i$  and  $j$  yield the following system of equations

$$\begin{aligned}(2\delta_j^{12} - \delta_j^{01})\delta_i^{01} &= c, \\ \delta_i^{12}\delta_j^{01} &= c, \\ \delta_j^{12}\delta_i^{01} &= c, \\ (\delta_i^{01} - \delta_i^{12} - \delta_j^{01} + 2\delta_j^{12})\delta_j^{01} &= c.\end{aligned}$$

which again implies that  $\delta_i^{01} = \delta_i^{12}$ . Therefore, no agent can place positive weight on forming two links. But then agent  $k$  strictly prefers to link to  $i$  and complete the circle if her  $\theta_k$  value exceeds

$$\frac{c}{\delta_i^{01} + 2\delta_j^{01} - 2\delta_i^{01}\delta_j^{01}},$$

which contradicts the hypothesis that only  $i$  and  $j$  form a single link.

Therefore, the links formed by agents  $i$  and  $j$  must both be to agent  $k$ . Again, suppose that agents  $i$  and  $j$  place positive weight on forming two links. The indifference conditions

of the criticals types, given by:

$$(\delta_j^{01}\delta_k^{02} - 2\delta_j^{12}\delta_k^{02})\delta_i^{01} = -c(\delta_j^{12}\delta_k^{02})\delta_i^{12} = c(\delta_i^{01}\delta_k^{02} - 2\delta_i^{12}\delta_k^{02})\delta_j^{01} = -c(\delta_i^{12}\delta_k^{02})\delta_j^{12} = c$$

imply that  $\delta_i^{01} = \delta_i^{12}$ , contradicting the hypothesis that agent  $i$  places positive weight on one link. Therefore  $i$  and  $j$  must be forming only a single link.

If agent  $k$  places positive weight on two links, then agents  $i$  and  $j$  will place zero weight on one link, which again contradicts the hypothesis. This follows from the indifference conditions of the critical types for each agent:

$$\begin{aligned} (2 - \delta_j^{01})\delta_k^{02}\delta_i^{01} &= c \\ (2 - \delta_i^{01})\delta_k^{02}\delta_j^{01} &= c \\ (\delta_i^{01} + \delta_j^{01})\delta_k^{02} &= 2c \end{aligned}$$

If  $k$  does not form two links, she forms no links at all. Since agent  $i$  and  $j$  form only a single link to  $k$ , the equilibrium is the periphery-sponsorship equilibrium.

■

Next, I eliminate the case where only one agent places positive weight on forming one link.

**Lemma 4** *There is no equilibrium where only one agent places positive weight on forming one link.*

**Proof.** Suppose not. Without loss of generality, let agent  $i$  link to agent  $j$ . Suppose also that at least one of  $\delta_k^{02}$  and  $\delta_j^{02}$  is strictly less than one. The indifference conditions for agent  $i$  imply

$$\delta_i^{01} = \delta_i^{12} = \frac{c}{\delta_k^{02}\delta_j^{02}},$$

which contradicts the hypothesis that agent  $i$  places positive weight on forming one link. If both  $\delta_j^{02}$  and  $\delta_k^{02} = 1$ , then whichever agent  $i$  does not form a link to receives zero utility, and would therefore prefer to deviate and link to  $i$  or  $j$  if her value is sufficiently large. Hence, this cannot be an equilibrium.

■

Finally, if any agent places positive weight on forming two links, without mixing, then it must be a center-sponsorship equilibrium.

**Lemma 5** *Suppose that no agent puts positive weight on forming one link, and agents do not mix. Then the only equilibrium with this property is a center sponsored equilibrium with one, two or three agents placing positive weight on forming two links.*

**Proof.** First, it cannot be that no agent places positive weight on forming two links in equilibrium; in this case the network would be empty, and agents with value above  $c$  would prefer to deviate and form two links. Therefore, the only profiles to check are those with one, two or three agents placing positive weight on two links.

All of these profiles are equilibria. The indifference conditions for agent  $i$ , for example, show that if agents  $j$  and agent  $k$  place zero weight on one link, then  $i$  will never place weight on 1 link in equilibrium, and will have the following cutoff between zero and two links:

$$\delta_i^{02} = \frac{c}{\delta_j^{02} \delta_k^{02}}.$$

If  $\delta_j^{02} = \delta_k^{02} = 1$ , that is, if  $j$  and  $k$  form no links, then  $i$  links to both if her value is above  $c$ . This is a one agent center-sponsorship equilibrium.

If  $\delta_j^{02} \delta_k^{02} = c$ , then  $i$  forms no links. This can occur if  $\delta_j^{02} = \delta_k^{02} = \sqrt{c}$ . This is a two agent center-sponsorship equilibrium.

If  $\delta_j^{02} \delta_k^{02} < c$  and both are less than 1, then there is a three agent center sponsorship equilibrium, where

$$\delta_i^{02} = \delta_j^{02} \delta_k^{02} = \sqrt[3]{c}.$$

■

This characterizes every possible equilibrium in the case with three agents, and completes the proof of proposition 3.

## Proof of Proposition 4

To show that this is an equilibrium, I will show that there are no profitable deviations; (1) insiders must not wish to link to anyone when their value is below their cutoff  $\theta_I$ , and must wish to link to every insider when their value is above their cutoff. (2) Insiders must also not wish to link to outsiders, no matter what other links they are forming, or what their value is. (3) Outsiders must not wish to form links to anyone when their value is below their

cutoff  $\theta_O$ , and (4) outsiders must wish to form only a single link to a random insider when their value is above their cutoff, and no other links.

Consider insiders first. The incremental value of a link to another insider, when not forming any other links, is

$$F(\theta_I)^{|I|-1}(\theta + \frac{n - |I|}{|I|}(1 - F(\theta_O))\theta) - c$$

That is, with probability  $F(\theta_I)^{|I|-1}$  the insiders are not already fully connected, and in that event, one additional link yields the value of an additional connection an insider,  $\theta$ , and potential a connected to outsiders connected to that insider. This has cost  $c$ . But by the definition of  $\theta_I$ , if  $\theta < \theta_I$ , this expression is negative. If  $\theta > \theta_I$ , from the definition of  $\theta_I$ ,

$$\begin{aligned} F(\theta_I)^{|I|-1} \frac{n - |I|}{|I|} (1 - F(\theta_O)) - (1 - F(\theta_I)^{|I|-1}) ((|I| - 1)\theta + (n - 1)(1 - F(\theta_O))\theta) \\ > (|I| - 1)(\theta - c) + (n - |I|)(1 - F(\theta_O)), \end{aligned}$$

Rearranging terms yields  $F(\theta_I)^{|I|-1}(\theta + \frac{n - |I|}{|I|}(1 - F(\theta_O))\theta) - c < 0$ .

A link to another insider is not profitable when an insider's value is below  $\theta_I$  and she is forming no other links. If she is forming links to outsiders, this incremental value is even lower; she will already be receiving some of the value in the second term of the interior expression above, so she is even less inclined to form links to other insiders.

The incremental value of a link to an outsider, when agent  $i$  is not herself already connected to every insider is

$$F(\theta_O)\theta + (1 - F(\theta_O))(F(\theta_I)^{|I|-1})(2\theta + \frac{n - |I| - 1}{|I|}(1 - F(\theta_O))\theta) - c.$$

That is, if that outsider is not already linking to the insiders, she will receive  $\theta$ . If that outsider is forming a link, then there is an additional benefit from linking to her; you get the value of the insider she is linking to, and any other outsiders that may be linking to that insider. But by our hypothesis on  $c$ , this expression is also negative. If agent  $i$  is already connecting the insiders herself, the incremental value of a link to an outsider is

$$F(\theta_O)\theta - c;$$

Only in the event that this outsider is not herself linking to some insider is there any incre-

mental value in linking to her. This is also negative by the hypothesis on  $c$ .

An outsider must not wish to form any links when her value is below her cutoff  $\theta_O$ . The incremental value of a link to an insider is

$$F(\theta_I)^{|I|}(\theta + \frac{n - |I| - 1}{I}(1 - F(\theta_O))\theta) + (1 - F(\theta_I)^{|I|})(|I|\theta + (n - |I| - 1)(1 - F(\theta_O))\theta) - c,$$

and by the definition of the outsider's cutoff, this is negative if  $\theta < \theta_O$ , and positive if  $\theta > \theta_O$ . If she is already forming a link to an insider, the incremental value of an additional link is

$$F(\theta_I)^{|I|}(\theta + \frac{n - |I| - 1}{|I|}(1 - F(\theta_O))\theta) - c;$$

Only if the set of insiders in not already connected is there any additional value in another link to an insider. This is negative by our hypothesis on  $c$ .

The incremental value to an outsider of a link to another outsider, when the agent is not already linking to the insiders, is

$$\begin{aligned} & \theta + (1 - F(\theta_O))(\theta + F(\theta_I)^{|I|}((|I| - 1)\theta + \\ & (n - |I| - 2)(1 - F(\theta_O))\theta) + (1 - F(\theta_I)^{|I|})\frac{n - |I| - 2}{|I|}(1 - F(\theta_O))\theta) - c. \end{aligned}$$

This link has value if the outsider selected is not already linked to any insiders, or if she is, then only if the set of insiders is not being connected any any insider, and the outsider selected is not already linking to same insider as this outsider is herself connected to. This is clearly less that simply connecting to the insider directly; In the event that this link to an outsider has value, a link to an additional insider has greater value. The incremental value when already linked to the inside is

$$F(\theta_O)\theta + (1 - F(\theta_O))F(\theta_I)^{|I|}\frac{|I| - 1}{|I|}\theta - c,$$

which is negative by the hypothesis on  $c$ . Thus, this profile of strategies is a Bayesian Nash Equilibrium.

## Proof of Corollary 1

Suppose the other agents follow  $s(\theta)$ , and form 0 links if  $\theta < \theta_I$ , and  $n - 1$  links if  $\theta \geq \theta_I$ . Consider the payoff to agent  $i$  of forming  $s$  links. In the event that someone else has formed  $n - 1$  links, agent  $i$ 's payoff is  $(n - 1)\theta_i - sc$ . This event occurs with probability  $(1 - F(\theta_I)^{n-1})$ . If all other agents form zero links, then her payoff is  $s\theta_i - sc$ , and this occurs with probability  $F(\theta_I)^{n-1}$ . Her expected payoff is therefore

$$(1 - F(\theta_I)^{n-1})(n - 1)\theta_i + F(\theta_I)^{n-1}s\theta_i - sc.$$

Only the last two terms depend on  $s$ , so the agent must maximize only

$$sF(\theta_I)^{n-1}\theta_i - sc = s(F(\theta_I)^{n-1}\theta_i - c).$$

This is positive if

$$\theta_i > \frac{c}{F(\theta_I)^{n-1}},$$

in which case, the maximizing choice of  $s$  is  $n - 1$ , the largest number of links possible. If it is negative, the maximizing choice of  $s$  is 0. Therefore, agent  $i$ 's best response is a cutoff rule, with cutoff  $\frac{c}{F(\theta_I)^{n-1}}$ . Any agent with type  $\theta < \frac{c}{F(\theta_I)^{n-1}}$  will form 0 links, and any agent with type  $\theta > \frac{c}{F(\theta_I)^{n-1}}$  will form  $n - 1$  links. Recall that

$$F(\theta_I)^{n-1}\theta_I = c \longrightarrow F(\theta_I)^{n-1} = \frac{c}{\theta_I},$$

so that the cutoff rule used by agent  $i$  is

$$\frac{c}{F(\theta_I)^{n-1}} = \frac{c}{\frac{c}{\theta_I}} = \theta_I.$$

Thus, it is a best response for agent  $i$  to use the same cutoff rule being used by the other  $n - 1$  agents,  $\theta_I$ , and  $\{s, \dots, s\}$  is a symmetric Bayesian Nash equilibrium, as was to be shown.

## Proof of Corollary 2

First, consider agent  $k$ , the center of the star. The payoff to agent  $k$  of forming a link is

$$\theta_k F(\theta_O^{n-1}) - c.$$

This is the probability that the agent is not already forming a link to her, minus the cost of the link  $c$ . This is positive if  $\theta_k > \frac{c}{F(\theta_O^{n-1})}$ . From the definition of  $\theta_O^{n-1}$ ,

$$F(\theta_O^{n-1}) = \frac{(n-1)\theta_O^{n-1} - c}{(n-2)\theta_O^{n-1}},$$

and by hypothesis

$$c \geq \frac{(n-1)\theta_O^{n-1}}{(n-2)\theta_O^{n-1} + 1}.$$

Substituting,

$$\frac{c}{F(\theta_O^{n-1})} > 1.$$

But since  $\theta_k \in [0, 1]$ , it is impossible that  $\theta_k > 1$ . Thus forming a link is never profitable for agent  $k$ , so she forms none, exactly as called for in strategy profile  $b$ .

For agents  $i \neq k$ , first examine the incentive to form a single link to agent  $k$ . The expected utility from doing so is

$$\theta_i(1 + (n-2)(1 - F(\theta_O^{n-1}))) - c, \tag{A.2}$$

and this is positive only if  $\theta_i > \theta_O^{n-1}$ . The value of each link in addition to a link to the center agent  $k$  is

$$\theta_i F(\theta_O^{n-1}) - c.$$

The argument above shows that this is negative. Therefore a single link to agent  $k$  dominates forming zero links, and forming a link to  $k$  and any number of links to other agents besides  $k$ . Finally, forming any number of links without a link to agent  $k$  is dominated by a single link to  $k$ . The expected utility of forming  $l$  links to agents  $j \neq k$  is

$$l\theta_i + (1 - F(\theta_O^{n-1}))^l(1 + (n-2-l)(1 - F(\theta_O^{n-1})))\theta_i - lc. \tag{A.3}$$

The first term is the directly utility from  $l$  connections, the second term is the expected utility from an indirect link to  $k$ , and the third term is the cost of link. The difference

(A.2)-(A.3)

$$(A.2) - (A.3) = c(l-1) + \theta_i((n-1-l)(1-F(\theta_O^{n-1})F(\theta_O^{n-1})^l + F(\theta_O^{n-1})(l-F(\theta_O^{n-1})l)) > 0.$$

Therefore, the utility of a single link to  $k$  dominates all other actions, if and only if  $\theta_i > \theta_O^{n-1}$ . Agent  $i$  therefore follows precisely the equilibrium strategy  $s_O$ .

## Proof of Proposition 5

The proof of the proposition follows exactly the argument in the proof of proposition 1. An agent's expected utility from contributing  $b_i$  in the contribution game is

$$(1 - F(\theta_I)^{n-1})\theta_i + F(\theta_I)^{n-1}\mathbf{1}_{\{b_i > k\}}\theta_i b_i.$$

Her best response is to contribute  $K$  if her value  $\theta_i > \theta_I$ , and contribute 0 if her values is below  $\theta_I$ .

## Proof of Proposition 6

The proof of the second part of the theorem follows the argument above, with the only difference being that she can only choose to indicate willingness to pay or not; she does not have all the alternative actions available in the network formation game. Her expected utility is

$$\mathbf{1}_{\{d_i=1\}}\left(\theta_i - \frac{K}{(1 + (n-1)(1 - F(\theta_O^n)))}\right).$$

Her best response is to signal willingness to pay if her value  $\theta_i > \theta_O^n$ , and to not signal willingness to pay if  $\theta_i > \theta_O^n$ .

## Proof of Proposition 7

First, note that under the partial order  $a < a'$  if  $\mu(a) < \mu(a')$ , the functions  $h$  and  $g$  are order preserving. Next calculate  $g(S_I(\theta, h(c, n)))$ .

$$g(s_I(\theta_i)) = \begin{cases} \mu(\{0\}^{n-1})c = 0, & \text{if } \theta_i \leq \theta_I; \\ \mu(\{1\}^{n-1})c = (n-1)c & \theta_i > \theta_I, \end{cases}$$

and applying  $h$ ,  $(h_2(n) - 1)h_1(c) = K$ . Furthermore, under the image of  $h$ , the cutoff  $\theta_I$  in the center-sponsorship equilibrium is the same as the cutoff in the contribution game. Therefore,  $g(s_I(\theta_i, h(\theta, c, n))) = t(\theta)$ , which is an equilibrium of the contribution game.

## Proof of Proposition 8

For  $h$  and  $g$  to be order preserving, consider a different order on  $A$ . Let  $a < a'$  if and only if  $a_{ik} < a'_{ik}$ . Then  $h$  and  $g$  are order preserving. Then  $g(s_O(\theta_i, h(\theta, c, n))) = r(\theta)$ , which is an equilibrium of the voluntary cost sharing mechanism.

### **Proof of proposition 9.**

If costs are 0, it is clearly dominant to link to all of your peers. Likewise, if costs are sufficiently large, it is dominant to form no links. All that remains is to show that these are the only possible configurations an agent's single link choices may following in equilibrium.

There are four possibilities; zero agents form a single link, one agents does, two agents do, or all three agents do. We can immediate eliminate the possibility only one agent does. To see this, let agent  $i$  link to agent  $j$  in equilibrium, and neither either  $j$  nor  $k$  form a single link in equilibrium. That at some cost  $c$ , agent  $j$  must be indifferent between forming two links and forming none, since equilibria are cutoff rules. But this can only be if the benefit the first and second link are equal; otherwise for a slightly lower cost, she will strictly prefer to form one link, a contradiction. But these expected values cannot be equal, because agent  $j$  has a higher probability of being already connected to agent  $i$  through  $i$ 's single link, and so  $j$  will strictly prefer to link to  $k$  than to  $i$ . Thus, this cannot be an equilibrium.

The case where no agents form a single link is the center-sponsorship equilibrium. The only equilibrium in which only two agents form a single link is the periphery-sponsorship equilibrium. To see that no other configuration in which two agents form a single link may be an equilibrium, consider the alternative possibilities; either two agents form a single link

to one another, while the third does not, or else two agents form links in a line, ending at the third agent, who does not form a link. In the first case, the third agent would wish to form a single link to the other two for an intermediate range of costs; the marginal benefit of a single link to one of them is sufficiently high. In the second case, the third agent in the line would wish to form a single link to complete the circle. In both cases the private cost at which agents wish to form two links is lower than the private cost at which it is better to form one link, rather than zero; that is, these agents will deviate from the putative equilibrium. The only possibility is the periphery-sponsorship equilibrium.

The case where all three form a single link is either the circle equilibrium, the full mixing equilibrium or the hybrid equilibrium. To see that no other configuration can be an equilibrium when all three agents form a single link, consider the alternative possibilities. We first eliminate the possibilities that only two agents mix. We next show that if no agents mix, it must be the circle equilibrium, if one agent mixes it must be the hybrid equilibrium, and if all three agents mix it must be the full mixing equilibrium.

It cannot be that two agents mix in forming their single link, because to do so they must be indifferent between linking to an agent who is mixing, and one who is not. Because these two agents will be connected to her with different probabilities, she will strictly prefer to link to the agent she is less likely to be connected to, and thus cannot mix.

If no agents mix, then the equilibrium architecture must be the circle, if all three agents form a single link with positive probability. For, suppose not: The only other possibility is that two agents form their single link to one another, and the third links to one of them. This cannot be an equilibrium; the agent who is not linked by the third agent will strictly prefer to deviate and link to the third agent, rather than link to the third agents "target." This is because she is already likely to be linked to the third agent's target, and less likely to be linked to the third agent, and she will prefer to link to the agent she is less likely to be connected to.

If one agent mixes, it must be the hybrid equilibrium. The only other possibility is that the agents being linked by the mixer do not form their links to one another, but either form both to the mixer, or form a line ending at the mixer. The former case cannot be an equilibrium. To see this, note that the agents linking the mixer will actually prefer to link to the other insider, rather than the mixer. This is because the mixer is already linking to her with positive probability, and the other insider is not. The latter cannot be an equilibrium either, because the mixer will not be indifferent between linking the two insiders; she will prefer to link the agent who is not linking her.

Finally, if all three agents mix, it must be the full mixing equilibrium. We need only show that all three agents must form their single link to the other agents with equal probability, and that they all use the same equilibrium cutoff. This will be implied by the the mixing indifference condition; in order for any agent to be willing to mix, she must be indifferent between linking to either of the other two. If all three agents mix, then they must all be using the same strategy, and linking to one another with equal probability. ■

### Proof of proposition 11.

It is clear that the expected value of the public good each agent expects to receive is higher the more links she forms. By the concavity of the utility function, this implies she will invest less. ■

### Proof of proposition 12.

To show that for the center-sponsorship equilibrium there is a critical  $k$  above which  $y_i^{\text{Low}}$  is 0, we write out an agent's optimization problem, given optimal play by his two peers, and the indifference condition that characterizes  $c_I$ :

$$\begin{aligned} \max_y U_{\text{Low}}(y, y_{\text{low}}, c) \\ \text{s.t.} \\ U'_{\text{Low}}(y_{\text{low}}, y, c) = 0 \\ U'_{\text{High}}(y_{\text{high}}, y, c) = 0 \\ U'_{\text{Low}}(y_{\text{low}}, y, c_I) = U'_{\text{High}}(y_{\text{high}}, y, c_I) \\ y \geq 0 \end{aligned}$$

This states that given if  $y_i^{\text{Low}}$  and  $y_i^{\text{High}}$  are chosen optimally by an agent's peers, their first order conditions will be satisfied with equality, and  $c_I$  will be defined by their indifference condition. At the critical  $k$ ,  $y_i^{\text{Low}} = 0$  will satisfy the low agent's first order condition exactly. We will check that when  $y_i^{\text{Low}} = 0$  is the optimal choice for other agents when their cost of linking is low, then  $y=0$  is the best response. The agent's lagrangian is given by

$$\mathcal{L} = U_{\text{Low}}(y) - \lambda_1 y - \lambda_2 U'_{\text{Low}}(y_i^{\text{Low}}, y, c) - \lambda_3 U'_{\text{High}}(y_i^{\text{High}}, y, c) - \lambda_4 (U'_{\text{Low}}(y_i^{\text{Low}}, y, c_I) - U'_{\text{High}}(y_i^{\text{High}}, y, c_I))$$

The taking derivatives with respect to the agent's choice,  $y$ , the first order condition is

$$U'_{\text{Low}}(y, 0, c) - \lambda_1 - \lambda_2 U''_{\text{Low}}(0, y, c) - \lambda_3 U''_{\text{High}}(y_i^{\text{High}}, y, c) - \lambda_4 (U''_{\text{High}}(y_i^{\text{High}}, y, c_I) - U''_{\text{Low}}(0, y, c_I)) = 0$$

We suppose that when  $y_{\text{low}} = 0$ ,  $y = 0$  is the optimal choice; that is,  $y = 0$  satisfies this equation. But note that

$$\begin{aligned} U''_{\text{Low}}(0, 0, c_I) &> 0, \\ U''_{\text{High}}(0, 0, c_I) &> 0. \end{aligned}$$

From the presumed optimality of the other agent's choices. Thus the first order condition of the problem reduces to

$$-\lambda_1 - \lambda_2 U''_{\text{Low}}(0, 0, c) - \lambda_3 U''_{\text{High}}(0, 0, c) - \lambda_4 (U''_{\text{High}}(y_i^{\text{High}}, y, c_I) - U''_{\text{Low}}(0, y, c_I)) = 0$$

Which implies that

$$\lambda_1 = -\lambda_2 U''_{\text{Low}}(0, 0, c) - \lambda_3 U''_{\text{High}}(0, 0, c) - \lambda_4 (U''_{\text{High}}(y_i^{\text{High}}, y, c_I) - U''_{\text{Low}}(0, y, c_I)) > 0$$

The last inequality follows because  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  are positive, since we suppose that each agent's first order condition holds exactly, and  $U''$  is strictly negative, due to the concavity assumptions of  $u$ . Therefore, the nonnegativity constraint on  $y$  binds, and we have verified that  $y = y_i^{\text{Low}} = 0$  is an equilibrium at this critical  $k$ . Due to the concavity assumptions on  $U$ , this is a global optimum. This method extends to the calculation of  $k^*$  is every other equilibrium, because it will always be the case that every other investment level is positive at the critical  $k^*$ . ■

### Proof of Proposition 10.

This is apparent from the fact that  $y_i^{\text{High}} > y_i^{\text{Mid}} > y_i^{\text{Low}}$ , and inspection of the indifference conditions that determine the linking cutoffs. ■

### Proof of Proposition 13.

To start, consider the periphery sponsorship, where the center agent invests  $y_{\text{Center}}^{\text{High}}$  when forming no links, and the outside agents invest  $y_i^{\text{High}}$  when forming no links. The cutoffs used by the outsiders are given by  $c^{\text{Low}}$  and  $c^{\text{High}}$ , while the single cutoff used by the insider

is given by  $c_I$ . To see that  $y_i^{\text{Mid}}$  is always positive, suppose not. If this is the case, then certainly  $y_i^{\text{Low}} = 0$  for every agent. Consider the first order condition of the outsider who is forming no links and of the outsider connecting to the insider.

$$F(c^{\text{Low}})(1 - F(c_I))u'(y + y_{\text{Center}}^{\text{High}}) + \quad (\text{A.4})$$

$$((1 - F(c^{\text{Low}}))(1 - F(c_I)) + F(c_I)F(c^{\text{High}}))u'(y) + (1 - F(c^{\text{High}}))F(c_I)u'(y + yh) - k, \quad (\text{A.5})$$

$$(1 - F(c_I))u'(y + y_{\text{Center}}^{\text{High}}) + F(c_I)F(c^{\text{High}})u'(y) + (1 - F(c^{\text{High}}))F(c_I)u'(y + yh) - k. \quad (\text{A.6})$$

The only difference between the outsider forming a link to the center at the agent who does not do so, is a higher probability of receiving the investment undertaken by the center. This has an additional  $(1 - F(c^{\text{Low}}))(1 - F(c_I))$  for the agent who links to the center. The first order condition of the insider who forms no links is given by

$$2F(c^{\text{Low}})(1 - F(c^{\text{High}}))u'(y + y_i^{\text{High}}) + (1 - 2F(c^{\text{Low}})(1 - f3))u'(y) - k \quad (\text{A.7})$$

These are all linear combinations of the marginal utility of investment when connected to one or more of an agent's peers. Because  $y_i^{\text{High}}$  and  $y_{\text{Center}}^{\text{High}}$  are positive, equations A.7 and A.5 are equal to zero at the choices  $y_i^{\text{High}}$  and  $y_{\text{Center}}^{\text{High}}$ :

$$F(c^{\text{Low}})(1 - F(c_I))u'(y_{\text{Center}}^{\text{High}} + y_i^{\text{High}}) + \quad (\text{A.8})$$

$$(F(c^{\text{Low}})F(c_I) + (1 - F(c^{\text{Low}}))(1 - F(c_I)))u'(yh) + (1 - F(c^{\text{High}}))F(c_I)u'(2yh) = k \quad (\text{A.9})$$

$$2F(c^{\text{Low}})(1 - F(c^{\text{High}}))u'(y_{\text{Center}}^{\text{High}} + y_i^{\text{High}}) + \quad (\text{A.10})$$

$$(1 - 2F(c^{\text{Low}})(1 - F(c^{\text{High}})))u'(ych) = k \quad (\text{A.11})$$

We can solve for  $u'(y_i^{\text{High}} + y_{\text{Center}}^{\text{High}})$  in each equation, and set them equal, to find

$$\frac{k - (1 - 2F(c^{\text{Low}})(1 - F(c^{\text{High}})))u'(y_{\text{Center}}^{\text{High}})}{2F(c^{\text{Low}})(1 - F(c^{\text{High}}))} = \frac{k - (F(c_I)F(c^{\text{High}}) + (1 - F(c^{\text{Low}}))(1 - F(c_I)))u'(y_i^{\text{High}}) - (1 - F(c^{\text{High}}))F(c_I)u'(2y_i^{\text{High}})}{F(c^{\text{Low}})(1 - F(c_I))}$$

Isolating  $k$ , we substitute this into the marginal utility of the agent linking the insider. Evaluating this at  $y = 0$ , we should have that marginal utility is negative. Otherwise, the agent will wish to make a positive investment. With this substitution, equation A.6 can be rewritten as

$$\begin{aligned}
& F(c^{\text{High}})F(c_I)u'(0) + (1 - F(c^{\text{High}}))F(c_I)u'(y_i^{\text{High}}) \\
& + \frac{2F(c_I)F(c^{\text{High}})^2 + 4F(c^{\text{Low}})F(c^{\text{High}}) - 4F(c^{\text{Low}})F(c_I)F(c^{\text{High}})}{1 - 2F(c^{\text{High}}) + F(c_I)}u'(y_{\text{Center}}^{\text{High}}) \\
& + \frac{2F(c_I)F(c^{\text{High}}) - 4F(c^{\text{High}}) - F(c_I)^2 - 4F(c^{\text{Low}}) + 4F(c^{\text{Low}})F(c_I) - 3F(c_I) + 4}{1 - 2F(c^{\text{High}}) + F(c_I)}u'(y_{\text{Center}}^{\text{High}}) \\
& - \left( \frac{2(1 - F(c^{\text{High}}))^2 F(c_I)}{(1 - 2F(c^{\text{High}}) + F(c_I))} \right)u'(2y_i^{\text{High}}).
\end{aligned}$$

Notice that the marginal utility in the negative is smaller in magnitude than the marginal utilities in the positive terms. Furthermore, the weight on the positives is to be greater than the weight on the negative term. Therefore, marginal utility at 0 is positive for the agent linking the center, and this cannot be an equilibrium.

This argument can be extended to every equilibrium structure, because when agents linking the center do not link, differences between the networks disappear; in each equilibrium, marginal utility will have the above form, and the same argument may be applied. ■

## REFERENCES

- [1] LA Adamic and E Adar. Friends and neighbors on the web. *Social Networks*, 25(3): 211–230, July 2003.
- [2] J.T. Addison and P. Portugal. Job search methods and outcomes. *Oxford Economic Papers*, 54(3):505, 2002. ISSN 0030-7653.
- [3] Jeremy Alden. Holding two jobs: An examination of moonlighting. In Stuart Henry, editor, *Informal Institutions: Alternative Networks in the Corporate State*, pages 43–57. New York: St. Martin’s Press., 1981.
- [4] M. Allingham and A. Sandmo. Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1:733–743, 1972.
- [5] P. Amaral and E. Quintin. A competitive model of the informal sector. *Journal of Monetary Economics*, 53(7):1541–1553, 2006.
- [6] A.R. Antunes and T. Cavalcanti. Start up costs, limited enforcement, and the hidden economy. *European Economic Review*, 51:203–224, 2007.
- [7] D.R. Baker. A structural analysis of the social work journal network: 1985–1986. *Journal of Social Service Research*, 15:153–168, 1992.
- [8] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1230, September 2000.
- [9] A. Banerjee and N. Andrew. Information, the dual economy, and development. *Review of Economic Studies*, 65:631–653, 1998.
- [10] Paul Allen Beck, Russell J. Dalton, Steven Greene, and Robert Huckfeldt. The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices. *The American Political Science Review*, 96(1):57–73, March 2002. ISSN 00030554. URL <http://www.jstor.org/stable/3117810>.
- [11] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives*, 12(3):151–170, 1998. ISSN 08953309. URL <http://www.jstor.org/stable/2647037>.

- [12] David Blau. An empirical analysis of employed and unemployed job search behavior. *Industrial and Labor Relations Review*, 45:738–52, 1992.
- [13] David Blau and Philip Robins. Job search outcomes for employed and unemployed. *Journal of Political Economy*, 98(3):637–55, 1992.
- [14] Stephen P. Borgatti and Martin G. Everett. Models of core/periphery structures. *Social Networks*, 21(4):375–395, October 2000. doi: 10.1016/S0378-8733(99)00019-2.
- [15] Steven Bortnick and Michelle Ports. Job search methods and results: Tracking the unemployed. *Monthly Labor Review*, 115(12):35–40, 1992.
- [16] Thomas Bradshaw. Jobseeking methods used by unemployed workers. *Monthly Labor Review*, 96(2):35–40, 1973.
- [17] Yann Bramoullé and Rachel Kranton. Public goods in networks. *Journal of Economic Theory*, 135(1):478–494, July 2007.
- [18] F. Busato and B. Chiarini. Market and underground activities in a two-sector dynamic equilibrium model. *Economic Theory*, 23:831–861, 2004.
- [19] Antoni Calvó-Armengol and Matthew Jackson. The effects of social networks on employment and inequality. *American Economics Review*, 94:426–54, 2004.
- [20] In-Koo Cho. Competitive equilibrium in a radial network. *The RAND Journal of Economics*, 34(3):438–460, 2003.
- [21] J. Choi and M. Thum. Corruption and the shadow economy. *International Economic Review*, 46(3):817–836, 2005.
- [22] H. Cremer and F. Gahvari. Tax evasion and optimal commodity taxation. *Journal of Public Economics*, 50:261 – 175, 1993.
- [23] E. et al. Dabla-Norris. What causes firms to hide output? the determinants of informality. *Journal of Development Economics*, 85:1–27, 2008.
- [24] S. Dessy and S. Pallage. Taxes, inequality and the size of the informal sector. *Journal of Development Economics*, 70:225–233, 2003.
- [25] James Elliott. Social isolation and labor market isolation: network and neighborhood effects on less-educated urban workers. *Sociology Quarterly*, 40(2):199–216, 1999.
- [26] M. Fafchamps and S. Lund. Risk-sharing networks in rural Philippines. *Journal of Development Economics*, 71(2):261–287, 2003. ISSN 0304-3878.
- [27] Armin Falk and Michael Kosfeld. It’s all about connections: Evidence on network formation. <http://econpapers.repec.org/paper/izaizadps/dp777.htm>, May 2003. URL <http://econpapers.repec.org/paper/izaizadps/dp777.htm>.

- [28] Lawrence F. Feick and Linda L. Price. The market maven: A diffuser of marketplace information. *The Journal of Marketing*, 51(1):83–97, 1987. ISSN 00222429. URL <http://www.jstor.org/stable/1251146>.
- [29] et al. Friedman. Dodging the grabbing hand: the determinants of unofficial activity in 69 countries. *Journal of Public Economics*, 76:459–493, 2000.
- [30] M. Fugazza and J.F. Jacques. Labor market institutions, taxation and the underground economy. *Journal of Public Economics*, 88:395–418, 2003.
- [31] A. Galeotti and S. Goyal. The Law of the Few. *American Economic Review*, 100(4):1468, 2010.
- [32] Joseph Gaughan and Louis Ferman. Toward an understanding of the informal economy. *The Annals of the American Academy of Political and Social Science*, 493(1):15–25, 1987.
- [33] Klarita Gerxhani. The informal sector in developed and less developed countries: A literature survey. *Public Choice*, 120:267–300, 09 2004.
- [34] J.K. Goeree, A. Riedl, and A. Ule. In search of stars: network formation among heterogeneous agents. *Games and Economic Behavior*, 67(2):445–466, 2009. ISSN 0899-8256.
- [35] R. Gordon and W. Li. Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, 93(7-8):855–866, 2009. ISSN 0047-2727.
- [36] Mark Granovetter. *Getting a job: A study of contacts and careers*. University of Chicago Press, second edition edition, 1995.
- [37] Mark S. Granovetter. *Getting a Job: A Study of Contacts and Careers*. Harvard University Press, 1974.
- [38] P Gregg and J Wadsworth. It Takes Two: Employment Polarisation in the OECD. Technical report, Centre for Economic Performance, LSE, September 1996.
- [39] Hendricks, Ken, Piccione, Michele, and Tan, Guofu. The economics of hubs: The case of monopoly. *The Review of Economic Studies*, 62(1):83–99, jan 1995.
- [40] Daniel A. Hojman and Adam Szeidl. Core and periphery in networks. *Journal of Economic Theory*, 139(1):295–309, March 2008.
- [41] Harry J. Holzer. Search method use by unemployed youth. *Journal of Labor Economics*, 6(1):1–20, 1988.
- [42] Yannis Ioannides and Linda Loury. Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, 43:1056–1093, December 2004.

- [43] Matthew Jackson. The economics of social networks. In Richard Blundell, Whitney Newey, , and Torsten Persson, editors, *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, volume 1, pages 223–233. Cambridge University Press, 2006.
- [44] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, September 2008. ISBN 0691134405.
- [45] M.O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996. ISSN 0022-0531.
- [46] S. Johnson, D. Kaufmann, , and P. Zoido-Lobaton. Regulatory discretion and the unofficial economy. *The American Economic Review*, 2(88):387392, 1998.
- [47] A.F. JORGE and L. Valadão. O setor informal no Brasil: metodologia e principais resultados. *Revista da ABET. Rio de Janeiro: IE/UFRJ, ano, 2*, 2002.
- [48] Elihu Katz, Paul Felix Lazarsfeld, and Elmo Roper. *Personal Influence*. Free Pr., 2006. ISBN 1412805074, 9781412805070.
- [49] Erich Kirchler. *The economic psychology of tax behaviour*. Cambridge University Press, 2007. ISBN 0521876745, 9780521876742.
- [50] Michael Kosfeld. Economic networks in the laboratory: A survey. *Review of Network Economics*, 3(1):19–41, 2004.
- [51] R.E. Kranton and D.F. Minehart. A theory of buyer-seller networks. *American Economic Review*, 91(3):485–508, 2001. ISSN 0002-8282.
- [52] Paul R. Krugman. *The Self-Organizing Economy*. Blackwell Publishers, 1996. ISBN 1557866996.
- [53] B. Fortin Lemieux, T. and P. Frechette. The effect of taxes on labour supply in the underground economy. *American Economic Review*, 84(1):231–254, 1994.
- [54] J. et. al. Losby. Informal economy literature review. Technical report, The Aspen Institute. Washington, DC., 2002.
- [55] M. McBride. Position-specific information in social networks: Are you connected? *Mathematical Social Sciences*, 56(2):283–295, 2008. ISSN 0165-4896.
- [56] Beth Mintz and Michael Schwartz. Interlocking directorates and interest group formation. *American Sociological Review*, 46(6):851–869, December 1981. ISSN 00031224. URL <http://www.jstor.org/stable/2095083>.
- [57] Mark S. Mizruchi. What do interlocks do? an analysis, critique, and assessment of research on interlocking directorates, November 2003.

- [58] James D. Montgomery. Social networks and Labor-Market outcomes: Toward an economic analysis. *The American Economic Review*, 81(5):1408–1418, December 1991. ISSN 00028282.
- [59] Kaivan Munshi. Networks in the modern economy: Mexican migrants in the u.s. labor market. *Quarterly Journal of Economics*, 118(2):549–599, May 2003.
- [60] S. Pedersen. *The shadow economy in Germany, Great Britain and Scandinavia: a measurement based on questionnaire surveys*. Rockwool Foundation Research Unit, 2003. ISBN 8790199324.
- [61] Michelle Ports. Trends in job search methods:1920-1992. *Monthly Labor Review*, pages 63–67, October 1993.
- [62] Cross R., Parker A., Prusak L., and Borgatti S.P. Knowing what we know: - supporting knowledge creation and sharing in social networks. *Organizational Dynamics*, 30:100–120, November 2001. doi: 10.1016/S0090-2616(01)00046-8.
- [63] Shanna Ratner. The informal economy in rural community economic development. Available from Department of Agricultural Economics, University of Kentucky Lexington, Kentucky 40546 TVA Rural Studies Program Contractor paper 00-03, 2000.
- [64] Albert Rees. Information networks in labor markets. *American Economics Review*, 56(2):559 – 566, May 1966.
- [65] A. Sandmo. The theory of tax evasion: a retrospective view. *National Tax Journal*, 58(4):643–663, 2005.
- [66] Andrea Galeotti, Sanjeev Goyal, and Jurjen Kamphorst. Network formation with heterogeneous players. Economics Discussion Papers 562, University of Essex, Department of Economics, Sep 2003.
- [67] F. Schneider and D. Enste. Shadow economies: Size, cause and consequences. *Journal of Economics Literature*, 38:77–114, 2000.
- [68] J. Slemrod and Yitzhaki S. Tax avoidance, evasion and administration. In: *J. Auerbach and M. Feldstein (eds.), Handbook of Public Economics.*, 3(1):231–254, 2002.
- [69] Giorgio Topa. Social interactions, local spillovers and unemployment. *The Review of Economic Studies*, 68(2):261–295, April 2001.
- [70] Thomas W. Valente. *Network Models of the Diffusion of Innovations*. Hampton Press, 1995. ISBN 1881303225.

## AUTHOR'S BIOGRAPHY

Dennis O'Dea was born in 1980 and attended the University of Wisconsin Madison, where he received a bachelor's degree in economics and mathematics. He then attend the University of Illinois at Urbana - Champaign, receiving his master's degree in 2007, and his Ph.D. in 2010. Since 2009, he has been a lecturer in economics in the Department of Economics at the University of Illinois at Urbana - Champaign.