Imperfect Monitoring in Communication Networks[1]

Michael McBride
University of California, Irvine

3151 Social Science Plaza, Irvine, CA 92697-5100, USA
949-824-7417 (tel), 949-824-2182 (fax)
mcbride@uci.edu

Revision: September 2004

Abstract

Individuals in many social networks imperfectly monitor other individuals' network relationships. This paper shows that, in a model of a communication network, imperfect monitoring leads to the existence of many inefficient equilibria. Reasonable restrictions on actions or on beliefs about others' actions can, however, eliminate many of these inefficient equilibria even with imperfect monitoring. Star networks, known to be efficient in many settings, are shown to have desirable monitoring characteristics. More generally, this paper provides a formal framework in which to study incorrect perceptions as an equilibrium phenomenon in social networks.

JEL Nos.: A14, C72, D20, D80.

Keywords: social networks, uncertainty, information, efficiency, star networks.

[1] I thank Ben Polak, David Pearce, Stephen Morris, Amil Dasgupta, Sanjeev Goyal, Chris Udry, Dirk Bergeman, Venkatesh Bala, Hongbin Cai, Bill Zame, Matthew Jackson, participants of the Topics in Game Theory group at Yale University, 2003 SWET Conference attendees, UC Irvine colloquium participants, and anonymous referees for comments on earlier drafts of this work.

1 Introduction

Friendships and other interpersonal relationships play important roles in the transmission of information, from news of job openings (Granovetter [21] [22]), to details about growing a particular crop (Conley and Udry [9]), to tips about welfare programs (Bertrand, Luttmer, and Mullainathan [3]). Moreover, these *social networks* evolve as individuals form, benefit from, and sever ties with others even though they, in the language of game theory, *imperfectly monitor* others' relationships in the network. For example, Friedkin [16] finds that an individual's "horizon of observability" is extremely restricted, usually comprising her direct network contacts and her contacts' contacts, and others have found how limited horizons lead network participants to maintain incorrect perceptions of other participants' network relationships (Kumbasar, Romney, and Batchelder [26], Bondonio [4], and Casciaro [6]).

Economists to date have used game theory to examine the formation and persistence of social networks, yet their work assumes that individuals, in equilibrium, completely observe (i.e., *perfectly* monitor) the entire network's structure. As such, the impact imperfect monitoring has on network formation outcomes is little understood, and many questions remain unanswered. How does observation of others' ties affect an individual's strategic decision to form ties? Does limited observation lead individuals to form different network relationships than they would otherwise? Are the differences economically meaningful? How much observation must individuals have to ensure the formation of an efficient network?

This paper presents the first systematic study of imperfect monitoring in endogenously forming social networks. In a simple model of a communication network, I mimic the observation present in actual networks by assuming that each individual only observes those network ties that are within x links from her in the network. For example, if $x = 1$ then a player only observes her direct network ties, if $x = 2$ then she observes her direct ties and the direct ties of her network neighbors, and so on. To find the network equilibria, I must account for players' observational limitations, however, the Nash Equilibrium concept is not appropriate since it imposes perfect monitoring. Instead, I use the *Conjectural Equilibrium* concept which is designed for games with imperfect monitoring and allows me to precisely model each individual's observation.[2] I can then compare the Conjectural Equilibria under various levels of x to see how the set of equilibrium networks changes as observation changes.

[2]See Battigalli, Gilli, and Molinari [2] and Gilli [20] for discussion of Conjectural Equilibrium.

This approach reveals how inefficient equilibria arise from imperfect monitoring. Perfect monitoring or $x \geq \frac{n}{2}$ (n is the number of players) is sufficient for efficient equilibria. However, as x decreases from $\frac{n}{2}$, equilibria can be over-connected in that individual may maintain costly ties to what are not observed to be parts of the network already accessed through other ties. Moreover, as x decreases from $\frac{n-2}{4}$, equilibria can also be under-connected when individuals cannot observe which areas of the network are currently not accessed. These inefficiencies arise from imperfect monitoring in two ways. First, players can incorrectly believe with high probability that the network is efficient and thus not make utility improving changes, even if the network is efficient. Second, players can correctly believe the network is not efficient but at the same time assign positive probability to many different networks being possible, thereby not knowing which move is a utility improving one. In each case, the imperfect monitoring provides no information to contradict players' beliefs, and so no player has an incentive to change the network.

Since the social costs from these inefficiencies can be significant, I examine sufficient conditions for efficiency. I find that reasonable restrictions on individuals' relationships and on their beliefs about others' relationships can eliminate many inefficient equilibria—even under very small x. Assuming players have common knowledge of rationality eliminates under-connections for any x. The *flow identification* condition, an original concept that allows players to identify the marginal benefit of each link without observing the entire network, eliminates over-connections for any x. Finally, the strictness refinement, which rules out any equilibria where at least one individual has multiple best responses, eliminates both over- and under-connections when $x \geq 2$ and results in only efficient, center-sponsored star equilibrium networks. This last finding deserves especial note. While previous work shows that star networks have special efficiency and stability properties, this paper shows that stars also have special informational properties. Because in a star network every connected individual is two or fewer links away, each individual with $x \geq 2$ can observe, in equilibrium, who else is connected to the network. I show how this knowledge yields efficiency. Hence, the strictness refinement ensures efficient equilibria even under the severe imperfect monitoring akin to that reported by Friedkin [16].

This paper fits into a growing literature that uses game theory to study endogenous network formation.[3] One segment of this literature examines broad classes of "allocation

[3] Recent examples of the growing literature by economists on networks include Dutta and Mutswami [14],

rules" (i.e., how network structures map to players' utilities) to find general features of networks that meet certain desirable properties, such as a coincidence of stability and efficiency. Because the set of allocation rules is very large, a second category of research looks at particular network settings and their naturally appropriate allocation rules. My paper is in this second category since it considers a specific network model. Within this category, my work is most closely related to Bala and Goyal [1]; I add imperfect monitoring to their communication network game. The closest work in the first category is by Jackson and Wolinsky [24], who study the tension between stability and efficiency. My paper considers a different tension—that between efficiency and imperfect monitoring—in a specific network context. My paper also makes a methodological contribution by being the first to use a non-Nash equilibrium concept in a network formation game. The network setting is a natural one for non-Nash concepts, not only because people imperfectly monitor others' actions in actual social networks, but also because the network provides immediate structure to the modeling of players' information about others' actions. In this way, my work provides social network researchers of other disciplines with a formal game theory framework within which to study incorrect network perceptions as an equilibrium phenomenon.[4]

The paper proceeds as follows. Section 2 introduces the basic model, and Section 3 examines it under perfect monitoring. Section 4 introduces imperfect monitoring and its implications for network formation. Section 5 examines further restrictions on players' beliefs and actions. Section 6 briefly discusses other variations, and Section 7 concludes.

2 The Basic Model

Suppose each player $i \in N = \{1, ..., n\}$, $3 \leq n < \infty$, knows an informative "fact" worth v to each player. A fact might be weather news, information about lucrative investment opportunities, descriptions of new productive techniques, etc. Although each individual automatically knows her own fact, the only way to learn another's fact is by communicating with that person directly or indirectly through one or more other individuals. Direct com-

Dutta and Jackson [12], Curranini and Morelli [10], and Kranton and Minehart [25]. Dutta and Jackson [13] compile much recent work, *Review of Economic Design* (2000) contains a recent symposium, and Jackson [23] provides a recent review of issues and results.

[4]My formal framework could be used by researchers of social perceptions (Kumbasar *et al.* [26], Bondonio [4], Casciaro [6]), the evolution of social relationships (e.g., Carley [5], Zeggelink [29], Doreian and Stokman [11]), and their implications for other social phenomena, such as mass movements (e.g., Chwe [7], [8]).

munication between any two players occurs if one or both players initiate a communication tie to the other, and initiating a link costs $c < v$ to the initiator. This cost captures the time, effort, or money expended or invested to form the communication link.

Each i chooses with whom to initiate ties. Let $s_i = (s_{i1}, ..., s_{ii-1}, s_{ii+1}, ..., s_{in})$ denote i's strategy, where $s_{ij} = 1$ signifies that i initiates a link to j, and $s_{ij} = 0$ signifies that i does not initiate a link to j. Let S_i denote i's set of possible link initiation profiles, and denote $S = S_1 \times ... \times S_n$. Strategy profile $s = (s_1, ..., s_n) \in S$ can be represented as a *graph* or network structure where nodes represent players and links represent communication ties. Player i learns j's fact only if there is a *path* that connects them. Formally, there is a path between i and j if either one of the following is true: $\max\{s_{ij}, s_{ji}\} = 1$ (direct communication), or there exist players $j_1, ..., j_m$ distinct from each other and from i and j such that $\max\{s_{ij_1}, s_{j_1 i}\} = \max\{s_{j_1 j_2}, s_{j_2 j_1}\} = \cdots = \max\{s_{j_m j}, s_{j j_m}\} = 1$ (indirect communication).[5]

A *component* $N_i \subseteq N$ is a subset of s with a path between any two players in the component and with no path between any player in the component and any player out of the component. Denote n_i the number of players in N_i. Let I_i denote the set of individuals to whom i initiates a link, $I_i = \{j \in N | j \neq i, s_{ij} = 1\}$, and (abusing notation) also the number of individuals in that set.

Each player has the following utility function $u_i(s_i, s_{-i}) = n_i v - I_i c$. Note that facts are transmitted clearly through the network (i.e., no flow decay) so that the value of j's fact to i is the same whether i and j are directly or indirectly connected. Figure 1(a) illustrates one possible s with $n = 6$. Player 1 has utility $u_1 = 4v - 2c$ since her component $N_1 = \{1, 4, 5, 6\}$ has four players $n_1 = 4$ and she initiates two direct[6] links $I_1 = \{5, 6\}$, both of which she initiated as represented by the dot[7] on her side of each direct link. Note that 6's utility is $u_6 = 4v$ even though she did not initiate links since $n_6 = 4$.

[5] Even though information flows both ways through a link, these are *directed* graphs because $s_{ij} = 1$ is distinct from $s_{ji} = 1$ in the assignment of link costs. A *non-directed* graph is one wherein $s_{ij} = 1$ and $s_{ji} = 1$ have identitical interpretations. See Wasserman and Faust [28] for a discussion of directed graphs.

[6] I use the term "direct link" to denote a link that connects i and j directly ($\max\{s_{ij}, s_{ji}\} = 1$) and not through other players, and this term is to be distinguished by the term "directed graph" which refers to qualitative features of links (see footnote 5).

[7] Directed networks are usually denoted by arrows, however, I use dots instead of arrows in this model because arrows can be misleading. The network is directed because of the costs and not the flow of facts, and an arrow might give the impression that facts only flow in the direction indicated by an arrow. The dot is to distinguish the directed nature of the graph from the flows, and it has been used in earlier work in this manner (e.g., Bala and Goyal [1]).

I analyze this model when players simultaneously choose link initiations. Before making decisions, players commonly know n, v, c, and all other assumptions stated above. After decisions are made, each player knows her resulting utility u_i and her set of initiations I_i, and each player observes some subset of all chosen actions (described later). I restrict attention to the pure equilibria of this well defined network game $\langle S_i, u_i \rangle_{i \in N}$.

Certain structural features play prominent roles in this paper. A network can have *redundant* links that form a *cycle*. In Figure 1(a), there is a direct redundant link, i.e., a 2-player cycle, between 4 and 5 because setting $s_{45} = 0$ while leaving $s_{54} = 1$ would not decrease N_i for any i. There is also a 3-player cycle consisting of 1, 5, and 6 since 1 receives 5's fact from 5 directly or through 6. Another feature is network connectedness. A network is *connected* if $N_i = N$ for all i, while a *disconnected* network is such that $N_i \subset N$ for all i. Figure 1(a) is disconnected with redundant links, Figure 1(b) is connected with redundant links, and Figure 1(c) is disconnected but without redundant links. Connected networks that have no redundant links, like Figures 1(d) and 1(e), are called *minimally connected* and have nice efficiency properties.

Proposition 1: *The set of efficient networks is the set of minimally connected networks. (Bala and Goyal [1])*

Since $v > c$, any i with $N_i \subset N$ is strictly better off without making anyone worse off by initiating a link to $j \notin N_i$, and it follows that an efficient network must be connected. Since redundant links are clearly inefficient, only minimally connected networks are potentially efficient. Since each minimally connected network yields total utility $n^2 v - (n-1)\, c$, they are all efficient in the sense of both sum of utilities and Pareto optimality.[8]

There are many advantages to using this model for my study. It captures many elements of actual social networks, such as the transmission of valuable information informally through individuals. Simultaneous link choices capture the notion that these networks arise from decisions made by individuals without formal coordinating devices. Since communication networks often span large distances, it is a meaningful setting in which to explore imperfect monitoring. Moreover, previous work studies this model under the assumption of perfect monitoring (e.g., Bala and Goyal [1], Galeotti, Goyal, and Kamphorst [19]), so it is a useful

[8] Both notions of utility have been used in networks. In general, the Pareto concept is appropriate when side payments are not allowed (Jackson [23]).

benchmark. Finally, because it is non-cooperative, I can use concepts already developed in non-cooperative game theory for the study of imperfect monitoring.

3 Perfect Monitoring

Perfect monitoring (full observation) implies that, after players simultaneously make their link decisions, each player sees the entire resulting network structure. The resulting network s^* is an equilibrium if, after fully observing s^*, no player i knows of a deviation from s_i^* that makes her strictly better off. Thus, the equilibrium concept for perfect monitoring should require each player to choose a best response to what all others actually do (i.e., to what all others are observed to do), and this is exactly what the Nash Equilibrium concept requires.

> **Definition:** *A (pure)* ***Nash Equilibrium*** *of game* $\langle S_i, u_i \rangle_{i \in N}$ *is a profile of strategies* $\{s_i^*\}_{i \in N}$ *such that for each* $i \in N$, $u_i\left(s_i^*, s_{-i}^*\right) \geq u_i\left(s_i', s_{-i}^*\right)\ \forall s_i' \in S_i$.

The set of network equilibria under perfect monitoring must thus be the set of Nash Equilibria, which can be completely characterized.

> **Proposition 2:** *Under perfect monitoring, the set of Nash Equilibria is the set of minimally connected networks. (Bala and Goyal [1])*

The logic is straightforward (see Bala and Goyal [1] for a formal proof). Suppose simultaneous link choices with $n = 6$ result in Figure 1(a). This network is not an equilibrium for two reasons. First, since i observes s, she knows exactly who is in N_i and who is not in N_i. Since $v > c$, she is strictly better off initiating a link to any $j \notin N_i$, and it follows then that any equilibrium must be connected. Second, since i observes s, she observes any redundant links, which she will not maintain in equilibrium. Thus, an equilibrium must be minimally connected. That any minimally connected network is an equilibrium follows because any non-redundant link must be a best response since the lowest marginal benefit it can provide, v, must be greater than the marginal cost c.

The potential role of coordination is apparent since there is no guarantee that a minimally connected network will result from simultaneously chosen links. Although we might think that formal coordination is required for efficiency, Bala and Goyal [1] describe simple dynamics in which players' actions (in a repeated game set-up) will converge to minimally connected networks without any formal coordination between players. Recent experimental work supports the notion that efficient networks in this game can be reached without formal coordination.[9] Thus, we can put aside coordination issues for the moment without the fear that we are missing an essential factor driving the equilibria. More importantly, since any full information equilibrium is efficient, if we find inefficient equilibrium under imperfect monitoring then it must be due to the change in informational structure.

4 Imperfect Monitoring

4.1 Methodological Issues

The possibility of imperfect monitoring raises three methodological issues. First, if players do not observe the links of all other players, then what subset of links do they observe? I assume that each individual observes all links within geodesic distance $x \geq 1$.[10] The area enclosed by dotted lines in Figure 2(a) captures what player 1 observes in the given network when $x = 1$. She observes only her links with 5 and 6 and no other links. When $x = 2$, as in Figure 2(b), she additionally observes the links between 4 and 5 and between 5 and 6. Figures 2(c) and 2(d) capture what she observes with $x = 3$, and $x = 4$ respectively. Because 7 and 8 are not in her component, she will never observe them under any finite x. Clearly, holding the structure fixed and increasing x results in each player observing weakly more actions. Also notice that at low x, one player will generally observe some part of the network that is also observed by another, but there might also be parts of the network one player observes that another does not, and also that neither observes.

This setting, which I call *x-link observation*, mimics to some extent the empirical finding that individuals have limited horizons of observability in that they are more likely to correctly perceive individuals that are closer to them in their networks (Friedkin [16], Kumbasar, Romney, and Batchelder [26], Bondonio [4], Casciaro [6]). There is also a strategic reason to consider this setting. After decisions result in a network outcome, player i could tell j of some player k unobserved by j in hopes of getting j to link with k, which could be to i's,

[9] Falk and Kosfeld [15] present results from experiments testing Bala and Goyal's [1] predictions. Equilibria in the 2-way flow model (used in my paper) are not reached as often as in the 1-way flow model, a finding that likely arises from the combination of notions of fairness and the asymmetry in payoffs of the 2-way flow model's strict equilibria.

[10] The $x \geq 1$ restriction is a natural one. A player should be aware of her own direct relationships whether or not she initiated the relationships.

but not j's, advantage. Although some "cheap talk" might entail credible claims, I do not study the bounds of cheap talk credibility here and leave such work to future researchers. I do note, however, that players might have difficulties establishing the credibility of claims, which could leave them to "believe only what they see," as is the case in x-link observation.

This discussion raises a second question: what is the appropriate equilibrium concept if players do not fully observe others' actions? A Nash Equilibrium assumes that players commonly know (in equilibrium) the actual profile of all players' actions, and a concept that imposes such knowledge is not appropriate if players cannot observe all actions. I instead use the non-Nash concept called *Conjectural Equilibrium* that is designed for games with imperfect monitoring. Such a game is defined as follows:

Definition: *An* ***imperfect-monitoring game*** *is a combination* $\langle S_i, u_i, m_i \rangle_{i \in N}$.

After actions have been chosen and s generated, player i receives a message (signal) m_i that reveals some subset of s. In general, m_i will depend on the actions chosen, so let i's signal be a function $m_i(s)$ (or $m_i(s_i, s_{-i})$). In a perfect monitoring environment, after players' choices result in s, m_i reveals s to each i, i.e., $m_i(s) = s\ \forall i$. With imperfect monitoring in the form of x-link observation, $m_i(s)$ reveals to i that part of s within x links of i in i's component N_i in that s. Let $L_i(y)$ be the set of $j \in N_i$ that are exactly y links away along some path from i to j, where $L_i(0) = \{i\}$ and $L_i(\infty) = \{j | j \notin N_i\}$. We can now formally describe x-link observation by the following assumption.

Assumption: *Consider a network game with imperfect monitoring characterized by x-link observation. After players make simultaneous choices resulting in s, player i's message $m_i(s)$ is the following:*

$$m_i(s) = \left\{ \begin{array}{c} s_j \text{ for all } j \in \{L_i(0) \cup L_i(1) \cup \cdots \cup L_i(x-1)\} \text{ and} \\ s_{kj} \text{ for all } k \text{ and } j \text{ such that } k \in L_i(x) \text{ and } j \in L_i(x-1) \end{array} \right\}.$$

The first part in the bracket is all links made by any player $x-1$ links or closer to i. The second part further adds those players x links away who link to someone $x-1$ links away. Note that links initiated by someone $x+1$ links away will not be observed, nor will links initiated by someone x links away to someone $x+1$ links away.

To define a Conjectural Equilibrium, let $\pi_i(s)$ be a probability distribution over all $s \in S$, and interpret π_i as i's beliefs.

Definition: *A **Conjectural Equilibrium** is a profile of actions and beliefs* $(s_i^*, \pi_i^*)_{i \in N}$ *such that for each* $i \in N$*:*

(i) $\sum_{s \in S} \pi_i^* (s)\, u_i (s_i^*, s_{-i}) \geq \sum_{s \in S} \pi_i^* (s)\, u_i (s_i', s_{-i})$*, for all* $s_i' \in S_i$*;*

(ii) for any $s \in S$ *with* $\pi_i^* (s) > 0$*, it must be that* $m_i (s^*) = m_i (s)$*.*

Condition (i) states that in equilibrium each player's action s_i^* must be a best response given her conjectured beliefs π_i^*. Condition (ii) states that for any player that assigns non-zero probability to a certain state of the world s, the signal received by that player in that state of the world must equal that player's signal in the true state of the world s^*. In other words, a player's beliefs must not contradict the information about the state of the world in the player's message.[11]

For the network game, a $(s_i^*, \pi_i^*)_{i \in N}$ combination is an equilibrium if, after making simultaneous link decisions and after receiving signals, no player has an incentive to change beliefs or actions. As is standard in game theory, I assume players commonly know the game set-up, which includes knowledge of v, c, utility functions, and x (i.e., the signal functions). Because players act to maximize utility, it is also commonly assumed that players know their payoffs. For this reason, I make the added restriction that π_i^* must also account for i's equilibrium utility u_i^* in the sense that π_i^* cannot assign non-zero probability to a state s' state would yield her utility different than u_i^*. I impose this restriction because it captures the idea that an equilibrium network should be one in which a player will not want to change her decision even after *ex post* link observation and utility realization. Note that restriction is implied by full information but not by imperfect monitoring, so my making this restriction explicit is new to the network literature.

An immediate implication of this restriction and commonly knowledge of the game set-up is the following:

Remark 1: *Consider an equilibrium* $(s_i^*, \pi_i^*)_{i \in N}$*, where* n_i^* *is the number of players in* i*'s component in* s^**. For any state* $s' \in S$ *such that* $\pi_i^* (s') > 0$*, it is necessary that the number of players in* i*'s component in* s' *must also equal* n_i^**.*

[11]See Battigalli, Gilli, and Molinari [2] and Gilli [20] for complete descriptions of imperfect monitoring games and the Conjectural Equilibrium concept. I note that Fudenburg and Levine's [17] Self-confirming Equilibrium is a Conjectural Equilibrium with the restriction that i's signal is the strategies that others play at all information sets that are reached with positive probability, so that i's beliefs are correct along the equilibrium path but not necessarily correct off that path. Also note that the Conjectural Equilibrium concept does not impose rationalizability, although it can be added (Rubinstein and Wolinsky [27]).

The result follows from the utility function $u_i^* = n_i^* v - I_i c$. With v, I_i, and c known by i, any s' that yields utility different than u_i^* would do so only because i's component in s' differed in size from n_i, and i knows s' cannot be the true state since she knows u_i^*. It is thus necessary that any state assigned non-zero probability has n_i^* in i's component.[12]

It is well known that the Conjectural Equilibrium concept potentially allows for many possible equilibria because it places so few restrictions on beliefs. A third methodological question arises: if a player does not observe certain players' links, then what should that player believe about those players' links? Instead of making arbitrary assumptions about what players should believe, I use a different approach. In essence, I will pick a network structure and look for beliefs that make the actions-beliefs combination an equilibrium.

One shortcoming of this approach is that it might allow for a very large set of equilibria, which would imply weak predictive power for this equilibrium concept. However, this is also an advantage since I can find all possible equilibrium structures without making additional restrictions about what players should believe about unobserved actions. Also, since the only restriction on beliefs is that they be consistent with revealed information (e.g., conjectural equilibrium may be thought of as a necessary condition for equilibrium), it allows me to find the set of equilibrium networks that could arise under any network dynamics and beliefs updating procedures. Moreover, if I am able to find a small set of equilibrium networks even under these weak conditions, then the results will be that much more striking. Although I consider restrictions later (Section 5), the concept is thus, for now, less restrictive than other concepts in a meaningful way.[13]

4.2 Equilibria with x-link Observation

The definition of Conjectural Equilibrium implies the following.

[12] Instead of restricting π_i^* to account for knowledge of u_i^* and the game set-up, I could alternatively obtain the result stated in Remark 1 by making n_i^* part of i's signal. I chose to not include it in i's signal so that the signals refer only to i's observation of others' links and not the observation of her utility.

[13] For example, consider the difference between a Conjectural Equilibrium and a Bayesian Equilibrium. The latter assumes that players begin with a common prior over possible states of the world, and that they use all available information to update their beliefs by Bayes's Rule. The former does not place any restrictions on prior beliefs, and this matches the notion of limited observation in a dynamic network setting. Since players' information would depend on their limited observation, it is likely that players will not have common beliefs about the state of the world. Furthermore, not restricting players to use Bayes's Rule can allow for differences in beliefs updating. As such, the Conjectural Equilibrium allows us to find all equilibrium networks that can arise under any network dynamics and any beliefs updating procedures.

Remark 2: *Any equilibrium under perfect monitoring is also an equilibrium under x-link observation for any x. Moreover, any equilibrium under x-link observation, $x \geq 1$, is also an equilibrium under link observation for all $x' \leq x$.*

A Nash Equilibrium is a Conjectural Equilibrium with the restriction that $\pi_i^* (s^*) = 1$, $\forall i \in N$. Since x-link observation puts fewer restrictions on equilibrium beliefs than does Nash equilibrium, the restrictions on beliefs in a Nash Equilibrium still satisfy the weaker restrictions (in a nested sense) in a Conjectural Equilibrium with a finite x. Essentially, the "correct beliefs" in a Nash Equilibrium will sustain a network as an equilibrium under limited observation if players happen upon those correct beliefs. The second part of Remark 2 follows from similar logic. Fix $x > 1$ and $x' < x$. x'-link observation places fewer restrictions on the beliefs than does x-link observation, so that even incorrect beliefs under the larger, x-link, observation must still satisfy the weaker restrictions under the lower, x'-link, observation.

A nice feature of this result is that the efficient Nash networks are still equilibria under imperfect monitoring. Unfortunately, we have no explanation of how players happen upon the correct beliefs when they have limited observation. Although players' beliefs must correctly reflect what is within their observational range, they might not accurately reflect what is outside their observational range.

To see this, first note that a cycle of size y is observed by any i in the cycle if $y \leq 2x$, and it is not observed by any i in the cycle if $y > 2x$. It is $2x$ because i sees out distance x along two paths, and she sees the cycle if she sees one player in both paths. 1's cycle in Figure 2(a) is size 3, but is not observed since $3 > 2x = 2$. It is observed in Figure 2(b) where $3 \leq 2\,(2) = 4$. Not observing a cycle implies that the cycle can exist in equilibrium.

Proposition 3: *If $x \geq \frac{n}{2}$ then any equilibrium network must be minimal, but if $x < \frac{n}{2}$ then there always exist equilibrium networks with one or more cycles of size $y = \{2x+1, 2x+2, ..., n\}$.*

Proof: *($x \geq \frac{n}{2}$ implies any equilibrium is minimal)* Suppose an equilibrium $(s_i^*, \pi_i^*)_{i \in N}$ where s^* has a y-member cycle, where player i is a member of the cycle, and without loss of generality, $s_{i,i+1} = 1$. Consider i in the cycle. From above, if $x \geq \frac{y}{2}$, then player i observes the cycle and π_i^* must assign probability zero to any state without this y-member cycle. According to these beliefs, i is strictly better off by setting $s_{i,i+1} = 0$. Thus, $(s_i^*, \pi_i^*)_{i \in N}$ cannot be an equilibrium.

Since $x \geq \frac{n}{2}$ implies that any n-player or smaller cycle would be observed, it follows that no equilibrium can have cycles of size n or smaller. Since n is the largest possible cycle, it follows that no equilibrium can have cycles.

(Existence of cycle when $x < \frac{n}{2}$ and $y \geq 2x+1$) From above, we know that a cycle is not observed if $x < \frac{y}{2} \Rightarrow y > 2x$. Because x and y are integers, a cycle will not be observed if $y \geq 2x + 1$.

Consider a connected equilibrium $(s_i^*, \pi_i^*)_{i \in N}$ that has exactly one cycle with $y \in \{2x + 1, ..., n\}$ members. As above, label the players of the cycled component 1, ..., y, and, without loss of generality, consider player i, $s_{i,i+1}^* = 1$. Denote player z_i the player who is exactly x links away from i on the path from i to z through $i+1$. By x-link observation, i does not observe the link between z_i and $z_i + 1$ because $z_i + 1$ is $x + 1$ links away through z_i and because z_i is at least $x + 1$ links away on the second path from i to z_i through $i - 1$.

Consider a connected network structure s_i' which is equivalent to s^* in every respect except that s_i' does not have the link between z_i and $z_i + 1$. Because s^* has only one cycle, s' must be minimal. Let $\pi_i^* (s') = 1$ and $\pi_i^* (s'') = 0$ for all $s'' \in S$, $s \neq s'$. Notice that $m_i (s') = m_i (s^*)$ and that s_i^* is a best response given π_i^* (i believes N_i is connected by Remark 1 and thus will not initiate a new link, and i believes s' is minimal and thus will not remove any links).

Construct π_i^* in that manner for each member of the cycle. For any i not in the cycle, set $\pi_i^* (s^*) = 1$. As constructed, no player's beliefs contradict her message and each player's strategy is a best response to her beliefs. Thus, for any network s^* such that $y \in \{2x + 1, ..., n\}$, we can find beliefs π_i^* that sustain s^* as an equilibrium and that are not contradicted by players' messages. □

Figure 3(a) with $x = 1$, for example, is an equilibrium, if π_1^* assigns probability 1 to Figure 3(b) being the true state and if the other players have similarly constructed beliefs. The key is that the signal 1 receives in each figure is the same and that her beliefs lead her to believe she is currently playing a best response. In general, cycles can exist in varying sizes so long as they are not too small relative to x, and the number of equilibrium cycles increases as x decreases or as n increases. Figure 3(c), which has overlapping 3-, 4-, and

5-player cycles, is an equilibrium if with $n = 5$ if we drop x to 1. Figure 3(d), which has two cycles, is an equilibrium when holding $x = 2$ but increasing n to 9.

Inefficiencies can also arise from under-connections. Figure 4(a) with $x = 2$ can be an equilibrium if π_1^* assigns probability $\frac{1}{7}$ to each of the networks in Figure 4(b), if the other players have similar beliefs, and if $\frac{6}{7}v < c < v$. 1's action is a best response since initiating a link to any player not in her observational range yields a net expected benefit of $\frac{6}{7}v - c < 0$. I will later discuss why individuals should not expect others to be isolated, but for now this example illustrates what is necessary and sufficient for a disconnected equilibrium. First, there must be a $j \in N_i$ who is not observed by i. Second, the expected benefits of making a link to someone not observed must be sufficiently low, so c must be sufficiently high.

> **Proposition 4:** *If $x > \frac{n-4}{4}$, then any equilibrium network is connected. However, if $x \leq \frac{n-4}{4}$ and c sufficiently high, then there exist disconnected network equilibria, and each player in a disconnected equilibrium must be in a component with at least $2x + 2$ players.*

I work through the proof here in the text and then discuss the results.

Proof of Proposition 4. *Part I—Necessity.* Define $d(i, j)$ to be the shortest geodesic distance between i and j in some s, and let $d_i^* \in \max_{j \in N_i} \{d(i, j)\}$ be the distance between i and that player in her component who is farthest from i. Now define $b \in N_i$ to be the "best observer" in N_i, i.e., b is the member of N_i whose d_i^* is the smallest. Clearly, if $x \geq d_b^*$ then b observes all n_i players in N_i, while if $x < d_b^*$ then no player observes all n_i members of N_i. Finally, note that the largest d_b^* can be is $\frac{n_i}{2}$ when n_i is even and $\frac{n_i - 1}{2}$ when n_i is odd, which occurs when N_i is a line component. For example, $b = 1$ and $d_b^* = 2$ in the line network in with $n_i = 5$ in Figure 3(b), and $b = 1$ and $d_b^* = 3$ in the line network in Figure 1(d) with $n = 6$.

1. *$n_i \geq 2x + 2$ is necessary for a disconnected equilibrium.* Consider disconnected equilibrium $(s_i^*, \pi_i^*)_{i \in N}$, and suppose $x \geq d_b^*$. Since $x \geq d_b^*$, b observes n_i players in N_i. By Remark 1, $\pi_b^*(s') = 0$ for any state s' in which any $j \notin N_i$ is a member of N_i. According to π_b^*, player b believes that linking to some $j \notin N_i$ would not be a redundant with probability one, thus yielding a net utility increase of at least $v - c > 0$. Because b believes she is strictly better off making this link, $(s_i^*, \pi_i^*)_{i \in N}$ cannot be an equilibrium. Thus, $x < d_b^*$ is necessary for a disconnected equilibrium.

When n_i is even, $d_b^* \geq \frac{n_i}{2}$ (from above), so when n_i is even, it follows that $x < \frac{n_i}{2}$ is necessary for a disconnected equilibrium. Since n_i is an integer, the necessary condition becomes $x \leq \frac{n_i-2}{2}$ for even n_i. For n_i odd, $x < \frac{n_i-1}{2}$ is necessary by similar logic, which becomes $x \leq \frac{n_i-3}{2}$. These two weak inequalities become $x \leq \frac{n_i-2}{2}$ for generic n_i, which yields $n_i \geq 2x+2$ as a necessary condition for any disconnected equilibrium component.

2. $x \leq \frac{n-4}{4}$ *is necessary for a disconnected equilibrium.* Since no player can be isolated in equilibrium, it follows that any disconnected equilibrium s^* will have at least two components, each of size at least $2x+2$. Hence, in any disconnected equilibrium it is necessary that $n \geq 2(2x+2) \Rightarrow x \leq \frac{n-4}{4}$, and any equilibrium with $x > \frac{n-4}{4}$ must be connected.

Part II—Sufficiency. To show the existence of a disconnected network equilibrium when $x \leq \frac{n-4}{4}$, I construct a $(s_i^*, \pi_i^*)_{i \in N}$ combination and show that it is an equilibrium when c is sufficiently high. From Remark 2, an equilibrium network under $x = \frac{n-4}{4}$ will also be an equilibrium network under any other $x < \frac{n-4}{4}$. Thus, we need only show existence when x equals the maximum integer less than or equal to $\frac{n-4}{4}$.

1. *Construct structure.* If n is even, partition N into two sets of players N_1 and N_n, such that $N_1 = \left\{1, ..., \frac{n}{2}\right\}$ and $N_n = \left\{\frac{n}{2}+1, ..., n\right\}$. Then construct s^*_{even} so that N_1 and N_n comprise two separate line components such that

$$s^*_{even} = \left\{ \begin{array}{c} s^*_{12} = s^*_{23} = ... = s^*_{\frac{n}{2}-1,\frac{n}{2}} = 1 \\ s^*_{\frac{n}{2}+1,\frac{n}{2}+2} = s^*_{\frac{n}{2}+2,\frac{n}{2}+3} ... = s^*_{n-1,n} = 1 \\ \text{all other } s^*_{ij} = 0 \text{ for all } i,j \in N \end{array} \right\}.$$

Note that $n_1 = n_2 = \frac{n}{2}$. If n is odd, partition N into two sets of players N_1 and N_n, such that $N_1 = \left\{1, ..., \frac{n-1}{2}\right\}$ and $N_n = \left\{\frac{n-1}{2}+1, ..., n\right\}$. Then construct s^*_{odd} so that N_1 and N_n comprise two separate line components similar to when n is even, the difference being $n_1 = \frac{n-1}{2}$ and $n_n = \frac{n+1}{2}$.

2. $x \leq \frac{n-4}{4}$ *is sufficient for* $K_i > 0$ *for all* i. Consider s^* constructed immediately above (either s^*_{even} or s^*_{odd}, whichever is appropriate given n). For $i \in N_i$, denote O_i to be set of players observed by i with o_i equal to the number of players in that set; let $\overline{O}_i = N/O_i$ be the compliment of O_i, with $\overline{o}_i$ the number in that set, and let $K_i = n_i - o_i$ be the number of players in n_i that are not observed.

Given the constructed s^* for some n, the best observer b will be the middle agent in the line network. Note that b will not observe all members of her component with $x \leq \frac{n-4}{4}$.

Since $b \in N_b$ observes o_b members out of n_b, and since b knows there are n_b members of N_b (Result 1), then b knows that there are exactly $K_b > 0$ players in N_b that she does not observe. Moreover, since b is the best observer, it must be true that $K_i > 0$ for each $i \in N_b$.

3. *Construct* π_i^*. Consider the following state s': (1) s' is exactly equal to s^* within i's observational range; (2) exactly K_i players from $\overline{O}_i$ each initiate a link to some player $j \in N_i$ who is exactly x links from i; (3) all other players in $\overline{O}_i$ are isolated. By x-link observation, the links between the K_i players and j are not observed. Let $\widehat{S}_i$ denote the set of all such possible s' structures, and note that there are $\binom{\overline{o}_i}{K_i}$ of them in $\widehat{S}_i$. Let

$$\pi_i^*(s') = \begin{cases} \frac{1}{\binom{\overline{o}_i}{K_i}} \text{ for each } s' \in \widehat{S}_i \\ 0 \text{ for any other } s' \text{ in } S \end{cases}$$

As constructed, these beliefs sum to 1. Also note that, according to π_i^*,

$$\begin{aligned} \Pr\left[j \in N_i | j \in \overline{O}_i\right] &= \frac{\binom{\overline{o}_i - 1}{K_i - 1}}{\binom{\overline{o}_i}{K_i}} \quad \forall j \in \overline{O}_i \text{ and} \\ \Pr\left[j \notin N_i | j \in \overline{O}_i\right] &= 1 - \frac{\binom{\overline{o}_i - 1}{K_i - 1}}{\binom{\overline{o}_i}{K_i}} \quad \forall j \in \overline{O}_i. \end{aligned}$$

The first probability is the probability that linking to j is redundant, and the second probability is that of making a non-redundant link.

4. s_i^* *is* i*'s best response given* π_i^* *if* c *is sufficiently high.* For s_i^* to a best response for i, i must be no better off removing or initiating links. According to π_i^*, none of i's links are redundant, so i believes she is strictly worse off removing any links. Consider adding links. Clearly, i believes she is strictly worse off by adding a link to any $i \in O_b$ because such a link would be redundant, so consider adding a link to $j \in \overline{O}_j$. Given π_i^*, i's expected net utility gain from initiating new links to $y \leq \overline{o}_i$ members of $\overline{O}_i$ is

$$\sum_{t=1}^{y} \left(\left(1 - \frac{\binom{\overline{o}_i - 1}{K_i - 1}}{\binom{\overline{o}_i}{K_i}} \right) v - c \right)$$

which is less than 0 if

$$c > \left(1 - \frac{\binom{\overline{o}_i - 1}{K_i - 1}}{\binom{\overline{o}_i}{K_i}} \right) v \equiv c_i^*.$$

Thus, if $c > c_i^*$, then i will not initiate any new links. Notice that $c_i^* < v$. Hence, if c is sufficiently high then s_i^* is a best response for i given π_i^*.

5. $(s_i^*, \pi_i^*)_{i \in N}$ *is an equilibrium.* We can construct π_i^* as above for every i in both N_1 and N_n. Define $c^* = \max_{i \in N} \{c_i^*\}$. If $c > c^*$, then s_i^* will be each i's best response. Moreover, each π_i^* was constructed so that $m_i(s^*) = m_i(s')$ for all $s' \in \widehat{S}_i$. Thus, we have met all conditions for $(s_i^*, \pi_i^*)_{i \in N}$ to be an equilibrium. □

Propositions 3 and 4 consider over- and under-connections separately, but equilibria can be both over- and under-connected. Combining Propositions 3 and 4 we see that $x \geq \frac{n}{2}$ is sufficient for efficiency; that over-connected equilibria exist when $x \in \left(\frac{n-4}{4}, \frac{n}{2}\right)$; and that equilibria that are simultaneously over- and under-connected exist when $x \leq \frac{n-4}{4}$. Limited observation thus threatens the efficient functioning of networks. Moreover, the inefficiencies get worse as x decreases, as seen by the relative efficiency losses for over- and under-connections. For example, the sum of utilities in Figure 5 increase as we move from (a) to (d): $72v - 12c$, $72v - 10c$, $144v - 13c$, $144v - 11c$. The worst inefficiencies arise from under-connections in this model, which follows from v being greater than c and from the positive externalities generated by links.

Imperfect monitoring leads to inefficient equilibria because it allows players to maintain incorrect beliefs about the network. First, players can be sure (or nearly sure) but wrong in assigning probability 1 (or close to 1) to states that are efficient but not the true state. This occurs in Figure 3(a), where each player incorrectly assigns probability 1 to a minimally connected network, when in truth it is connected but not minimal. Because players believe the network is efficient, they have no incentive to identify and eliminate inefficiencies. In essence, the imperfect monitoring does not prevent misplaced overconfidence. Second, players can correctly believe that the network is not efficient but have beliefs spread over a range of possible states so that any single action is more likely to do harm than good. This occurs in Figure 4(a), where each player knows the network is disconnected, but no player knows exactly who is and is not outside of her own component. If a player was more sure of the state, even incorrectly sure, then the network would not be an equilibrium because the player would risk initiating a new link. In this instance, the imperfect monitoring essentially leads to underconfidence or too much uncertainty in beliefs.

Note that these inefficiencies can arise more generally in various network formation models. If imperfect monitoring makes it difficult for players to accurately distinguish the true state from other states then players can sustain incorrect beliefs and not take actions to remove inefficiencies. If the true state is nearly efficient, then Pareto improvements may be missed due to players' misplaced overconfidence. If the true state is far from efficient but no player has enough confidence that any single given action is Pareto improving, then the wide uncertainty can lead to very inefficient equilibria.

5 Sufficient Restrictions for Efficiency

Increasing x eliminates inefficient equilibria, but might other restrictions eliminate inefficiencies while holding x fixed? First consider restricting players' actions by ruling out equilibria in which at least one player has multiple best responses, i.e., consider only strict equilibria.[14] Much prior work uses this restriction because it greatly refines the set of equilibria (e.g., Bala and Goyal [1], Galeotti, Goyal, and Kamphorst [19]). It also has a pertinent dynamic interpretation. Network models are often motivated by dynamic network formation, and strict equilibria are absorbing states under many dynamic processes.[15] We thus might expect to see a higher proportion of strict equilibria than weak equilibria in actual social networks.

Bala and Goyal [1] show that in this model with full information the set of strict equilibria is the set of connected, center-sponsored stars, each of which is efficient. A center-sponsored star is a component such as Figure 1(e) in which all links are initiated by a single player called the center. Full information is not necessary for this result, however.

Proposition 5: *If $x \geq 2$, then the set of strict network equilibria is the set of connected, center-sponsored stars.*

Proof: *(N_i that is not a center-sponsored star cannot be a strict equilibrium component)* Consider component N_i in a strict equilibrium network s^*, where N_i is not a center-sponsored star. By the definition of a center-sponsored star, N_i must have at least two players $i, j \in N_i$ who initiate links.

If $n_i = 2$, then it must be true that $s_{ij} = s_{ji} = 1$, and the direct redundant link can be identified and profitably removed under full observation, thus con-

[14] A strict Nash Equilibrium would have $u_i\left(s_i^*, s_{-i}^*\right) > u_i\left(s_i', s_{-i}^*\right)$ for all $s_i \in S_i$, for all i. A strict Conjectural Equilibrium would similarly have ">" instead of "≥" in condition (i) of the definition.

[15] Consider a repeated game of network formation with full information where links are formed in each period. Suppose players are in a non-strict equilibrium in period t. If, given the structure in t, a player has multiple best responses and randomizes among them in period $t+1$, then the resulting structure in period $t+1$ will differ from that in period t. In a strict equilibrium, no player makes these changes.

tradicting the equilibrium. Now suppose $n_i > 2$. Since $i, j \in N_i$, there must be a path between them. If they are directly connected and their only links are to each other, then there must be a direct redundant link, which, again, cannot be an equilibrium. If they are directly connected and there is no direct redundant link between them, then, without loss of generality, it must be that $s^*_{ij} = 1$ and $s^*_{jk} = 1$ for some $k \in N_i$, $k \neq i$. With $x \geq 2$, i identifies the link between j and k and knows she can set $s_{ij} = 0$ and $s_{ik} = 1$ without a decrease in utility. This violates the definition of a strict equilibrium. Now suppose that i and j are not directly connected. Then there must be a player k in the path between i and j. Without loss of generality, let $s^*_{ik} = 1$. Then there must be some player l (who can be j) such that $\max\{s^*_{kl}, s^*_{lk}\} = 1$. With $x \geq 2$, i can observe the link between k and l and set $s_{ik} = 0$ and $s_{il} = 1$ without a drop in utility, which violates the definition of strict equilibrium. All possible cases have been considered, so N_i cannot be in a equilibrium component.

(A disconnected network cannot be a strict equilibrium) Suppose disconnected equilibrium (s^*, π^*). There must be a component $N_i \subset N$ in s^* that is a center-sponsored star. Let i be the star's center. Since N_i is a star, i observes all $j \in N_i$ with $x \geq 2$. i's beliefs must thus correctly reflect exactly which players are not in N_i. According to these beliefs, i is strictly better off connecting to any $j \notin N_i$, so s^* cannot be an equilibrium.

(Sufficiency) Follows from Proposition 2 since a minimally connected network combined with correct beliefs for everyone is a strict equilibrium. $\square$

The argument relies on a link initiator being able to observe her neighbor's links to know a link switching opportunity. If one link initiator observes a neighbor's link, which requires only with $x \geq 2$, then she can switch her link from her neighbor to her neighbor's neighbor and be no worse off. Only the center-sponsored star is immune to the this link switching.

The striking aspect of this result is that only 2-link observation is sufficient for efficiency. How do we interpret this? Since strict equilibria are absorbing states of many dynamic processes, one implication is that we can find explicit network dynamics that always converge to an efficient equilibrium even when players have very low link observation. Another striking feature is that star networks have special monitoring properties. At least one individual, the star's center, knows the component's true structure, and this fact ensures a connected equilibrium. And since the peripheral players are at most two links from one another, they know that the network is connected and thus will not initiate any new links which would be redundant. These informational properties are enough to ensure efficiency.

The strictness condition restricts players' actions, but we can also restrict players' beliefs. The common knowledge of rationality assumption restricts players to rationalize their beliefs.[16] If i knows j's utility function and that j seeks to maximize her payoff, then i must believe j will choose a best response to her beliefs π_j. Moreover, π_j must correspond, in equilibrium, to j's message m_j and must assume that other players are also playing best responses, and so on. This assumption has immediate implications for connectedness.

Proposition 6: *Any rationalizable equilibrium is connected for any x.*

In using Figure 4(a) to illustrate the existence of disconnected networks, player 1 believed that any player not in her component was isolated. However, no rational player will be isolated, so if players are commonly known to be rational then no player, in equilibrium, can assign non-zero probability to a network in which another player is isolated. This will raise 1's expected marginal value of making a non-redundant link, and thus increase the likelihood she initiates a new link. Proposition 6 follows from extending this basic logic.

Proof of Proposition 6. Denote $\Pr[j \notin N_i|\pi_i] \equiv \sum_{s\in S}[\pi_i(s) I(j \notin N_i|s)]$, where $I(\cdot)$ is the indicator function that takes value 1 when the argument is true. Further define $Z_i \equiv \{j \in N| \Pr[j \notin N_i|\pi_i] > 0\}$, z_i to be the number of players in that set, and l to be the member of Z_i for whom the probability of not being in N_i is the highest:

$$l \in \left\{k \in X_i | \max_{j\in X_i} \Pr[j \notin N_i|\pi_i]\right\}.$$

1. $\Pr[l \notin N_i|\pi_i^*] \geq \frac{n-n_i}{z_i}$ *in any disconnected equilibrium.* Consider equilibrium $(s_i^*, \pi_i^*)_{i\in N}$. By definition

$$\begin{aligned}\sum_{j\in Z_i} \Pr[j \notin N_i|\pi_i^*] &\equiv \sum_{j\in Z_i}\sum_{s\in S}[\pi_i^*(s) I(j \notin N_i|s)] \\ &= \sum_{s\in S}\sum_{j\in Z_i}[\pi_i^*(s) I(j \notin N_i|s)].\end{aligned}$$

[16] See Fudenberg and Tirole [18] for exented discussion of common knowledge of rationality and rationalizable strategies.

By Remark 1, i knows that her component has exactly n_i members. As such, in any equilibrium, each state s for which $\pi_i^*(s) > 0$ must have exactly $n - n_i$ players not in N_i. It follows that

$$\begin{aligned}\sum_{s\in S}\sum_{j\in Z_i}\left[\pi_i^*(s)\, I\left(j \notin N_i | s\right)\right] &= \sum_{s\in S}\left[\pi_i^*(s)\left(n - n_i\right)\right] \\ &= \left(\sum_{s\in S}\pi_i^*(s)\right)\left(n - n_i\right) = n - n_i.\end{aligned}$$

Now suppose $\Pr\left[l \notin N_i | \pi_i^*\right] < \frac{n-n_i}{z_i}$. Then, since l has the highest probability of not being in N_i, it must be true that $\sum_{j\in Z_i}\Pr\left[j \notin N_i | \pi_i^*\right] < \left(\frac{n-n_i}{z_i}\right) z_i$. But this violates the fact that the probabilities must sum to $n - n_i$ as shown. Thus, it must be true that $\Pr\left[l \notin N_i | \pi_i^*\right] \geq \frac{n-n_i}{z_i}$ in any disconnected equilibrium.

2. *In a disconnected equilibrium, i must believe with non-zero probability that there exists $n_j < n_i$.* Suppose a disconnected equilibrium $(s_i^*, \pi_i^*)_{i\in N}$ with distinct components N_i and N_j, and suppose $\pi_i^*(s) = 0$ for any $s \in S$ in which $n_i < n_j$. Consider i's choice to link to l. Since $\Pr\left[l \notin N_i | \pi_i^*\right] \geq \frac{n-n_i}{z_i}$, and since π_i^* says $n_j \geq n_i$, the marginal value of a non-redundant link is weakly greater than $n_i v - c$. Thus, i's believed net gain from linking to l is at least $\frac{n-n_i}{z_i} n_i v - c$.

With $n \geq n_i + n_j$, $n_j \geq n_i$, and $n \geq z_i$, it follows that $\frac{n-n_i}{z_i} \geq \frac{1}{2}$. Combining this with $n_i \geq 2x+2$ from Proposition 4 yields $\frac{n-n_i}{z_i} n_i v - c \geq \frac{1}{2}(2x+2)v - c$, which is strictly positive for all $x \geq 1$ since $v > c$. Thus, according to π_i^*, i must believe she is strictly better off in expectation by linking to l, which contradicts the assumption of equilibrium. Hence, it is necessary that i must believe with non-zero probability that there exists a component with fewer players than her own component.

3. *A rationalizable network must be connected.* Consider disconnected equilibrium $(s_i^*, \pi_i^*)_{i\in N}$. From above, there must exist some state s' with component of size $n_j < n_i$ such that $\pi_i^*(s') > 0$. However, to rationalize s', player $j \in N_j$ in state s' must assign non-zero probability to the existence of a component $n_k < n_j$, say in some state s'', and this state can also only be rationalized by j if k rationalizes the existence of an even strictly smaller component, and so on, until some player z in some state s''' must rationalize a state of the world in which a player is isolated, which cannot be rationalized. Hence, i cannot rationalize the existence of a network with a component strictly smaller than her own.

If player i cannot rationalize the existence of any component strictly smaller than her own, and any disconnected network in which i believes any component is bigger than her own will lead her to form connections, then the only possible rationalizable networks are connected. The proof has not used any specific x, so this holds for all x. $\square$

That rationalizability ensures connected equilibria is significant since the largest inefficiencies arise from under-connections. However, connected but non-minimal networks can be rationalized if players do not observe the cycles, so rationalizability does not ensure an efficient network. I now propose a restriction, called *flow identification*, that is new to the literature and that ensures no over-connections.

> **Definition:** *A network game $\langle S_i, u_i, m_i \rangle_{i \in N}$ has* ***flow identification*** *when each player knows for each of her links what her final utility would be if all of her links but that particular link were removed, and each player knows for any two of her links what her final utility would be if all of her links but those two were removed.*

The name signifies that players identify, in essence, the marginal "utility flow" of each link. In state s with $i, j \in N_i$, let $u_{i|j}$ be i's utility if all i's links (whether initiated by i or by j) were removed except her link with j, holding all other links fixed. Let $u_{i|jk}$ be i's utility if all i's links were removed except her links with j and k (again, holding all else fixed). Under flow identification after the simultaneous link choices are made, each i knows $u_{i|j}$ and $u_{i|jk}$ for all $j, k \in N_i$. It follows that since i knows I_i, v, and c, i will thus know $n_{i|j}$ and $n_{i|jk}$, which are the corresponding component sizes.

Knowing both $n_{i|j}$ and $n_{i|jk}$ has an important implication: $n_{i|j} < n_{i|jk}$ signifies that the link with k is not redundant, but if $n_{i|j} = n_{i|jk}$ then the link with k is redundant.[17] This is the case in Figure 5(a) where flow identification reveals to player 1 that $n_{1|2} = n_{1|6} = n_{1|26} = 6$. Equilibrium beliefs must reflect this fact, which means that any network with cycles cannot be an equilibrium. Proposition 7 summarizes this result.

> **Proposition 7:** *Any flow identification equilibrium has no cycles for any x.*

[17] Note that $n_{i|j} \leq n_{i|jk}$ since the addition of the k link can only increase i's component.

Why would flow identification exist? One possibility is that flow identification follows from a more primitive assumption about the nature of the good that flows through the network. For example, communication networks transmit information, and if 1 learns the same information from two of her neighbors, she may infer the presence of a cycle. While this example does not relate directly to knowledge of utility as defined by flow identification, it has a similar implication in that it results in redundant links being identified.

Since flow identification brings minimality and common knowledge of rationality brings connectedness, the combination of the two ensures an efficient equilibrium. Remarkably, we ensure an efficient outcome by restricting beliefs and without increasing observational capabilities. Furthermore, flow identification and common knowledge of rationality approximate some features of actual networks. Overlaps in communicated facts will signal to individuals that a link could be redundant, and flow identification can be present in dynamic network environments with the temporary elimination and re-addition of individual links. As for believing other individuals are rational, I note that informal social networks often connect individuals like friends, family, and co-workers with whom there is some degree of interaction, and this suggests that individuals are aware to some degree of other network participants' motives, incentives, and thought-processes.

The main implication of these results is that potential inefficiencies arising from limited observation are less likely to arise if the network has link-switching, qualitatively distinctive information flows, and knowledge of other players' rationality. Each of these restrictions works in a different way to eliminate the potential for inefficiencies. Common knowledge of rationality allows players to identify some networks as not being consistent with rational behavior, thereby allowing them to assign probability 0 to states they otherwise could not have ruled out. Flow identification directly reveals more information about the state, thus allowing players to better distinguish one state from another. The strictness refinement works quite differently. Since center-sponsored stars are the only strictly stable components, any equilibrium must have the center of the star know the component with certainty (assuming $x \geq 2$), thereby leading the center to know of utility improving changes.

6 Other Considerations

Two variations deserve brief mention. If $v < c$ then the equilibria can change dramatically, although much of the intuition remains the same. The empty network (in which all players are isolated) is an equilibrium under both full observation or any x-link observation, and, if $v < \frac{c}{n-1}$, it is the only equilibrium. The empty network is also rationalizable since players can rationalize being isolated. Furthermore, while center-sponsored stars are immune to switching, they cannot be equilibria when $v < c$ because each link's marginal utility is less than its cost. Thus, the only strict equilibrium is empty.

A second variation is *flow decay* whereby facts lose value as they travel through the network. Suppose the utility function is now $u_i = \delta^{d(i,j)} v - I_i c$, where $\delta \in [0, 1]$ is the decay parameter and $d(i, j)$ is the shortest path distance in links between i and j. Over-connections that arise inefficiently in Sections 3, 4, and 5 with no decay ($\delta = 1$) can now be efficient with flow decay ($\delta < 1$) since players would rather get facts from closer to the source. In fact, if δ is sufficiently low then the set of equilibria is the same no matter what level of monitoring is present. If $v - c > \delta v \Rightarrow c < (1 - \delta) v$ then i prefers initiating a new direct link to j (holding all other links fixed) even if j is only two links away (and, of course, if j is more than two links away). Thus, no matter her other links, she always prefers a directly link to j. Her best response is to set $s_{ij} = 1$ for any j where $s_{ji} = 0$ and $s_{ij} = 0$ otherwise. Hence, for any $x \geq 1$, the set of equilibria is the set of "complete" networks (without direct redundant links), and these are efficient so that even extreme imperfect monitoring does not affect efficiency. The results are more complicated with δ close to but strictly less than 1. I note that the strictness condition can lose refining power. A player will switch links if it brings her closer to other players, and she will know of such opportunities under perfect monitoring, but under imperfect monitoring, she might be unsure of which switch is a proper one. Star networks will no longer be the only strict equilibria.

7 Conclusion

This paper shows that limited observation of others' network relationships can lead to inefficient outcomes, and that these inefficiencies can be economically significant. Nonetheless, certain restrictions on players' actions or beliefs—ones that can arise naturally in some network settings—will eliminate some of these inefficiencies. The main implication of this finding is that many actual communication networks can disseminate information reasonably well despite limited observation. Situations with frequent link switching, common knowledge of motives and incentives, flow identification, and highly valued facts are more likely to avoid inefficiencies generated by imperfect monitoring. Settings with large flow decay should also avoid these inefficiencies.

Future research has many avenues to consider. The coincidence of flow decay and imperfect monitoring deserves further study. Researchers should also look for mechanisms that can overcome inefficiencies associated with limited information. Since incorrect beliefs can cause inefficient outcomes, researchers should also examine the source of individuals' beliefs about parts of the network they cannot observe. Other work should study the sources of persistent instabilities in networks. My work demonstrates that link switching is one type of instability that can possibly aid in the formation of efficient networks, and empirical work can seek to test this prediction. Previous empirical work does not directly address the role limited observation plays in network formation, but such work will ultimately lend greater insights into the evolution of economically significant social relationships.

8 References

1. V. Bala, S. Goyal, A non-cooperative model of network formation, Econometrica 68 (2000), 1181-1230.
2. P. Battigalli, M. Gilli, C. Molinari, Learning and convergence to equilibrium in repeated strategic interactions: an introductory survey, Ricerche Econ. 46 (1992), 335-378.
3. M Bertrand, E. Luttmer, S. Mullainathan, Network effects and welfare cultures, Quar. J. Econ. 115 (2000), 1019-1055.
4. D Bondonio, Predictors of accuracy in perceiving informal social networks, Soc. Networks 20 (1998), 301-330.
5. K. Carley, A theory of group stability, Am. Soc. Rev. 56 (1991), 331-354.
6. T. Casciaro, Seeing things clearly: social structure, personality, and accuracy in social network perception, Soc. Networks 20 (1998), 331-351.
7. M. Chwe, Structure and strategy in collective action, Am. J. Soc., 105 (1999), 128-156.
8. M. Chwe, Communication and coordination in social networks, Rev. Econ. Stud. 67 (2000), 1-16.
9. T. Conley, C. Udry, Social learning through social networks: the adoption of new agricultural activities in Ghana, Am. J. Agr. Econ. 83 (2001), 668-673.

10. S. Currarini, M. Morelli, Network formation with sequential demands, Rev. Econ. Design 5 (2000), 229-250.

11. P. Doreian, F. Stokman, Eds., Evolution of social networks, Gordon and Breach Publishers, Amsterdam, 1997.

12. B. Dutta, M. Jackson, The stability and efficiency of directed communication networks, Rev. Econ. Design 5 (2000), 251-272.

13. B. Dutta, M. Jackson, Eds., Models of the strategic formation of networks and groups, Springer-Verlag, Heidelberg, 2003.

14. B. Dutta, S. Mutuswami, Stable networks, J. Econ. Theory 76 (1997), 322-344.

15. A. Falk, M. Kosfeld, It's all about connections: evidence on network formation, IZA Discussion Paper No. 777, 2003.

16. N. Friedkin, Horizons of observability and the limits of informal control in organizations, Soc. Forces 62 (1983), 54-77.

17. D. Fudenberg, D. Levine, Self-confirming equilibrium, Econometrica 61 (1993), 523-545.

18. D. Fudenberg, J. Tirole, Game theory, MIT Press, Cambridge, 1996.

19. A Galeotti, S. Goyal, J. Kamphorst, Network formation with heterogeneous players, Tinbergen Institute and Department of Economics, Erasmus University Rotterdam, 2004.

20. M. Gilli, On non-Nash equilibria, Games Econ. Beh. 27 (1999), 184-203.

21. M. Granovetter, The strength of weak ties, Am. J. of Soc. 78 (1973), 1360-1380.

22. M. Granovetter, The strength of weak ties: a network theory revisited," Soc. Theory 1 (1983), 201-233.

23. M. Jackson, A survey of models of network formation: stability and efficiency, in: G. Demange and M. Wooders (Eds.), Group Formation in Economic Networks: Networks, Clubs, and Coalitions, Cambridge University Press, Cambridge, forthcoming.

24. M. Jackson, A. Wolinsky, A strategic model of social and economic networks, J. Econ. Theory 71 (1996), 44-74.

25. R. Kranton, D. Minehart, A theory of buyer-seller networks, Am. Econ. Rev. 91 (2001), 485-508.

26. E Kumbasar, A.K. Romney, W. Batchelder, Systemic biases in social perception, Am. J. Soc. 100 (1994), 477-505.

27. A Rubinstein, A. Wolinsky, Rationalizable conjectural equilibrium: between nash and rationalizability, Games Econ. Beh. 6 (1994), 299-311.

28. S. Wasserman, K. Faust, Social network analysis: methods and applications, Cambridge Univ. Press, Cambridge, 1994.

29. E. Zeggelink, Evolving friendship networks: an individual-oriented approach implementing similarity, Soc. Networks 17 (1995), 83-110.

Figure 1

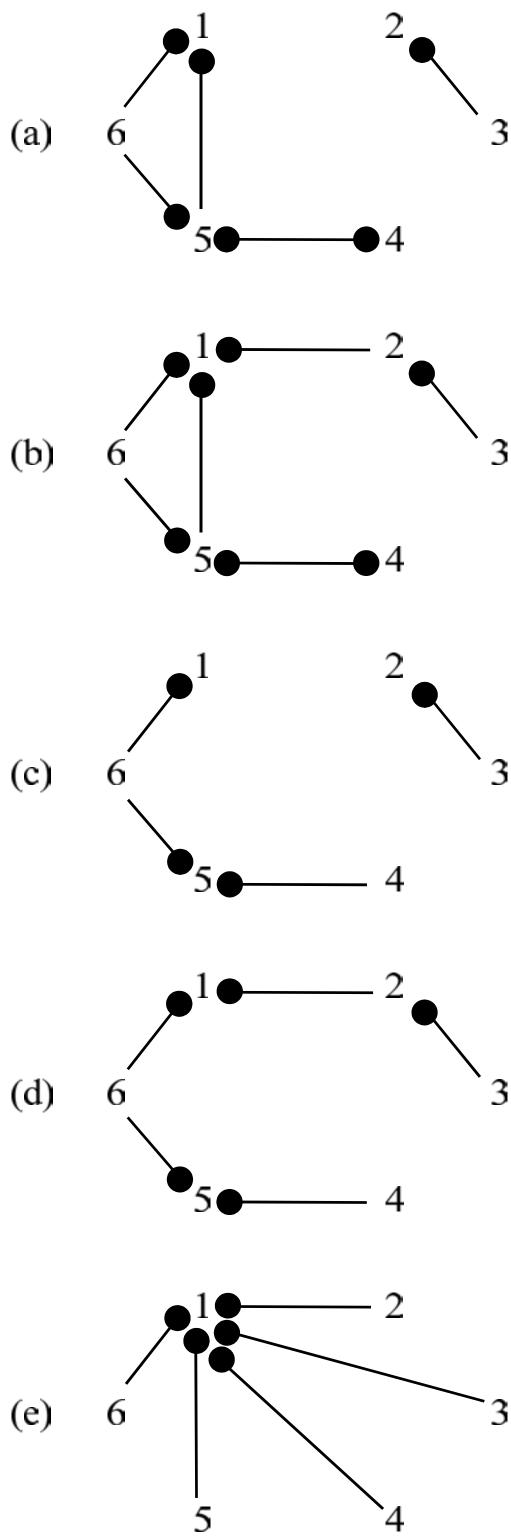


Figure 2

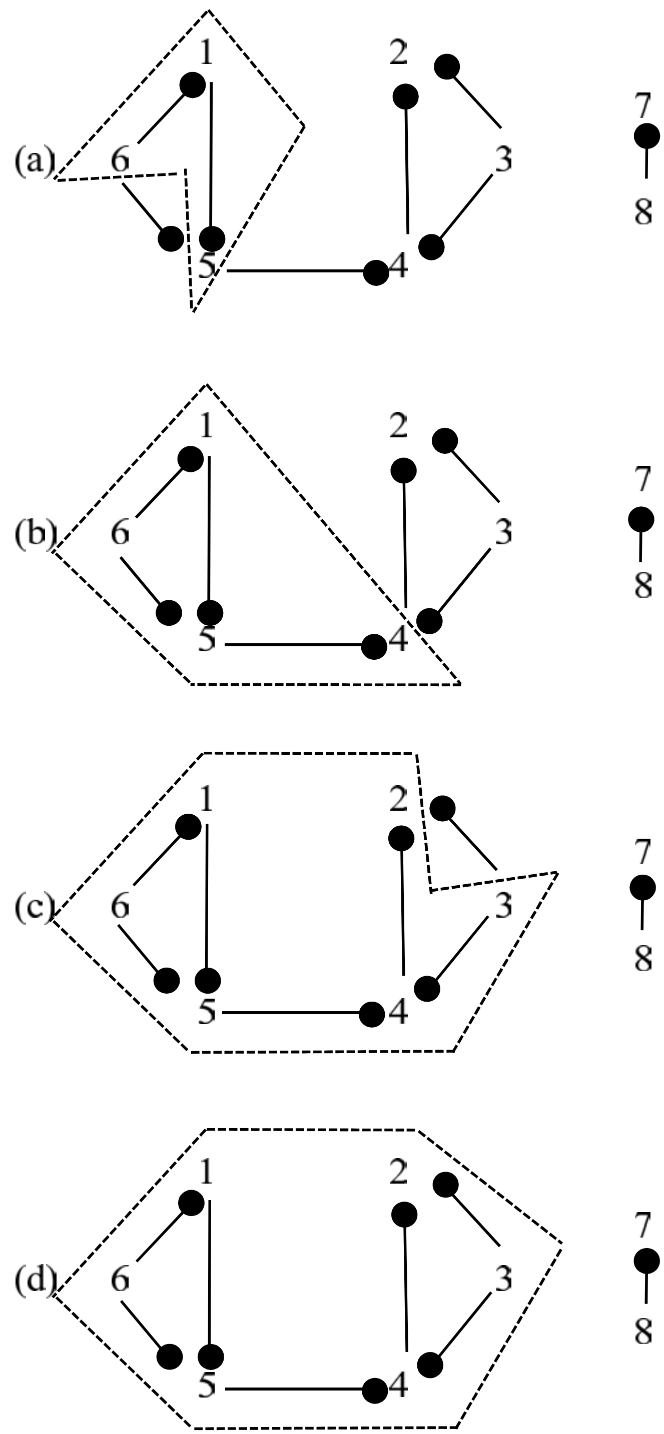
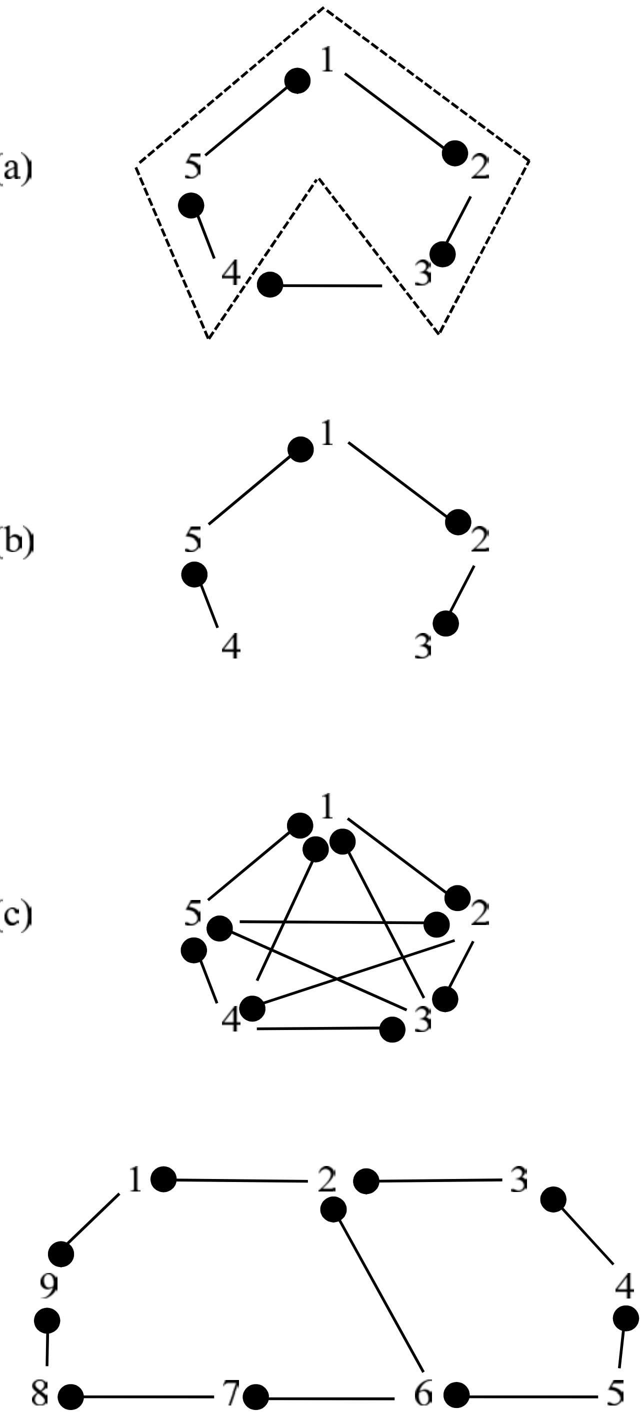


Figure 3



(a)

(b)

(c)

(d)

Figure 4

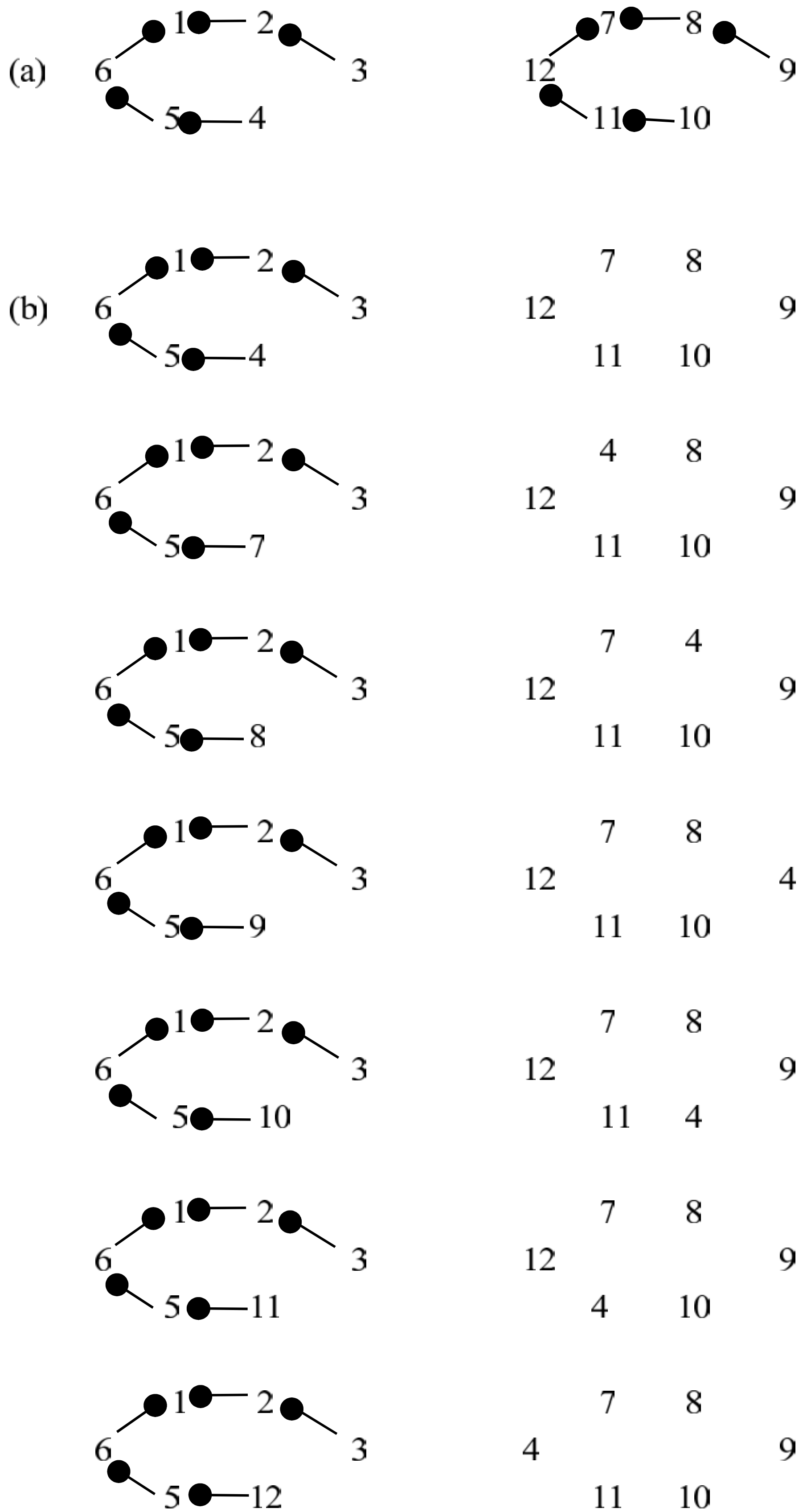
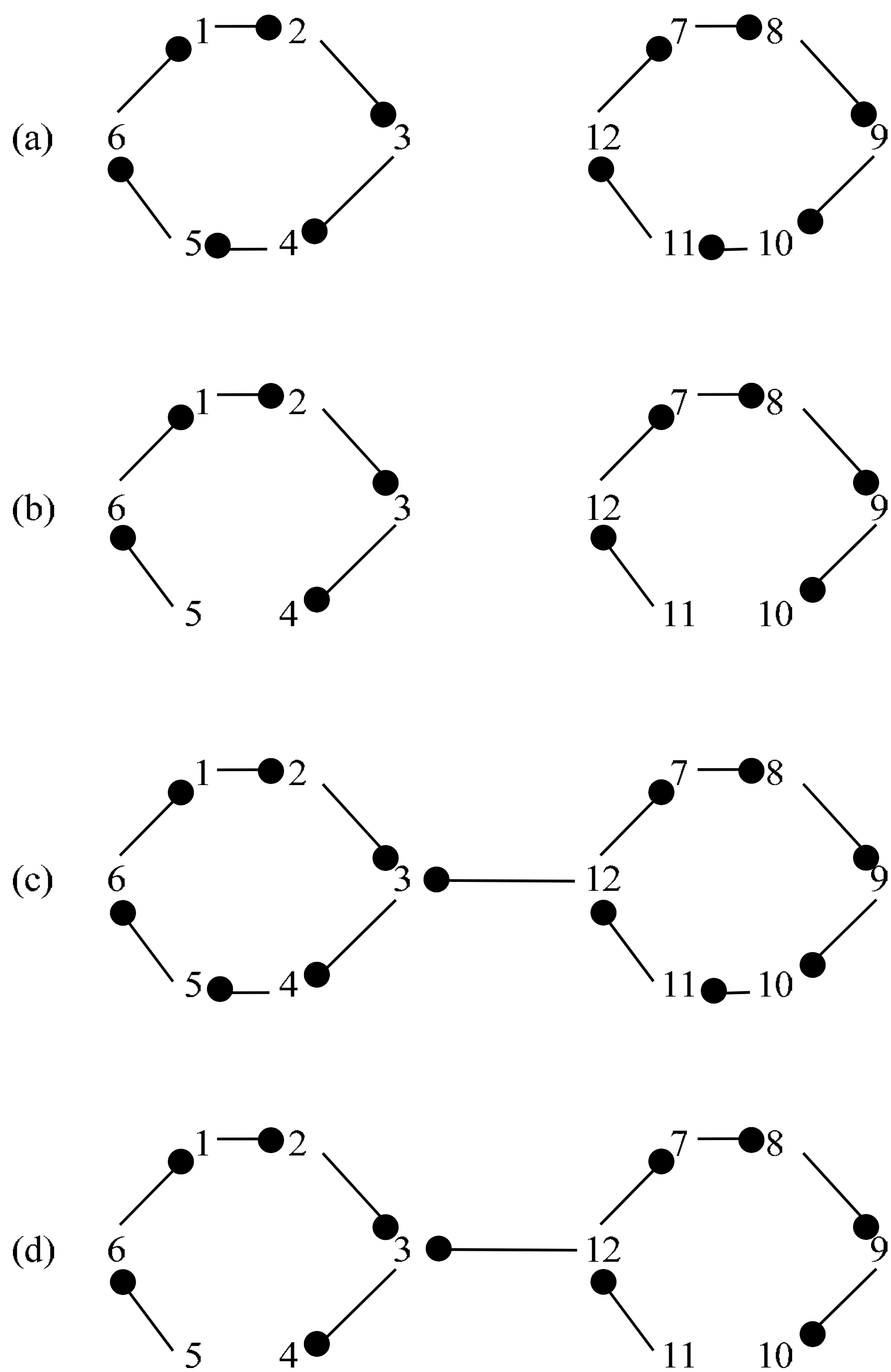


Figure 5


(a)
1
2
3
4
5
6
7
8
9
10
11
12
(b)
1
2
3
4
5
6
7
8
9
10
11
12
(c)
1
2
3
4
5
6
7
8
9
10
11
12
(d)
1
2
3
4
5
6
7
8
9
10
11
12