Position-specific Information in Social Networks: Are You Connected?

Michael McBride*
Department of Economics
University of California, Irvine

This version: July 2007

Abstract

Individuals in social networks often imperfectly monitor others' network relationships and have incomplete information about the value of forming new relationships. This paper formally examines these informational limitations in a simple model of network formation. Although incomplete information and imperfect monitoring each lead to the existence of inefficient equilibria that would not exist if participants had full information, each generates a different type of inefficiency. These inefficiencies increase in number and scope as information becomes more localized. Thus, my results suggest that actual social networks will be structured inefficiently in general.

JEL Classifications: A14, C72, D20, D80.

Keywords: limited horizons, observation, communication, connections.

Social networks of relationships underlie many economic and social activities—from the trade of goods and services in non-centralized markets [Kranton and Minehart (2001)] to the spread of information about new productive techniques [Conley and Udry (2001)] and job openings [Calvó-Armengol and Jackson (2004)]. Yet, all of these social networks have a common feature that has been largely ignored in theoretical work: individuals initiate and maintain social ties with very incorrect beliefs about the structure of the network.[1] For example, in a recent edited volume of sociologists' models of network formation, Stokman

*3151 Social Science Plaza, Irvine, CA, 92697-5100, mcbride@uci.edu. I thank Ben Polak, David Pearce, Stephen Morris, Amil Dasgupta, Sanjeev Goyal, Chris Udry, Dirk Bergeman, Venkatesh Bala, Hongbin Cai, Bill Zame, Matthew Jackson, Stergios Skaperdas, participants of the Topics in Game Theory group at Yale University, attendees at the 2003 South West Economic Theory (SWET) Conference and the 2007 Public Economic Theory Conference, UC Irvine colloquium participants, and anonymous referees for comments and support.

[1] This feature has been empirically studied in various sociological research. For example, see Laumann (1969), Kumbasar, Romney, and Batchelder (1994), Bondonio (1998), and Casciaro (1998).

and Doreian (1997) list "actors optimize based on **local information** only" as one of five "guiding principles" they recommend for network formation modeling, yet they acknowledge that "None of the contributions in this volume incorporates this idea systematically" (244-248, their emphasis).[2] At the other end of the spectrum are economists' game theoretic models, which usually assume that individuals have full and global information.[3]

This paper examines the role local information plays in network formation by focusing on two distinct types of informational limitations. First, an individual usually does not observe exactly who is connected to whom in the network. Second, she is not always certain about the value of certain connections. In game theoretic terminology, we refer to these limitations as *imperfect monitoring* and *incomplete information*, respectively, and each is likely to affect the type of network formation outcome. Suppose Ann learns of a job opening from her friend Bill, who learned of the job from his cousin Chris, who just happened to learn of the job from his neighbor Diane. If Ann does not observe Chris interact with Diane (imperfect monitoring), then she might not know the origin of the job news. She also might not know that Diane knows of other good job openings (incomplete information). Ann could be better off by connecting directly to Diane but not know it because of her limited information. As a result, limited information leads the persistence of an inefficient network.

Will limited information lead to the persistence of networks that differ from those that would arise under full information? If so, why, and are the differences economically meaningful? Does imperfect monitoring affect network formation differently than incomplete information? This paper examines these questions using a systematic game theoretic approach. I use as a starting point a model of network formation first introduced by Bala and Goyal (2000) and later extended by Galeotti, Goyal, and Kamphorst (2006). In this model, each individual has a "fact" that can be communicated through the network, and each player derives utility from another's fact only if she learns it through the network.

I introduce two new tools to study position-specific information in networks. First, I introduce the formal notion of *x/y-link observation*: an individual monitors all social ties

[2] Local information modelling is their second principle. The other four are: one, indivuals are instrumental in their network participation; three, networks evolve by agents acting in parallel; four, keep models simple; and, five, models should be empirically testable.

[3] See Dutta and Jackson (2003) for a collection of many of the important contributions to this field.

that are within x links and observes the types of all individuals that are within y links from her in the network. Second, I introduce the *Generalized Conjectural Equilibrium* (GCE) concept. Although the Nash Equilibrium (NE) concept can be used to examine this network formation game, it also restricts each player to have correct beliefs about other individuals' types and links in equilibrium. This restriction is a strong one for a setting in which each individual's information depends on her unique position in the network. Other extensions and refinements of the NE concept make similar restrictions, e.g., the Bayesian NE concept restricts individuals to have correct beliefs about the probability distribution over player types. The GCE concept relaxes this restriction and instead allows each individual's equilibrium beliefs to be incorrect—as we observe in the real world. The only restriction on beliefs is that they be consistent with the individual's limited information.

My first main result concerns the general effect of position-specific information on equilibrium networks. Network equilibria under full information in this model are either empty (no links at all) or minimally connected (each player accesses every other's fact with no over-connections) and are often efficient. However, I present a complete characterization of GCE networks to show that as position-specific information becomes more limited (as x and y decrease), the number of inefficient equilibria increases. The underlying reason is intuitive: observing less of one's network prevents one from identifying, and thus eliminating, network inefficiencies. In fact, in the least information setting, $x = 1$ and $y = 0$, virtually any network that makes each individual no worse off than being isolated can be an equilibrium. The "virtually any" includes numerous networks that are inefficient due to over-connections, under-connections, or both.

A second finding is that the two informational limitations—imperfect monitoring and incomplete information—have fundamentally different equilibrium implications. Perfect monitoring allows individuals to identify, and thus remove, over-connections because it allows an individual to observe when she is paying for costly links that are unnecessary. Notice, however, that perfect monitoring does not necessarily reveal the potential value of not yet formed links, so it does not imply that new, utility enhancing links will be formed. Complete information, on the other hand, reveals the value of being linked to certain individuals, so it allows individuals to identify (some) new link formations that are utility increasing. It does not, conversely, identify when current links are not necessary. Thus, the ability to monitor actions generally allows individuals to remove over-connections but does not compel individuals to overcome under-connections, while knowledge of types allows individuals to identify under-connections, but does not allow individuals to identify over-connections. Efficiency may require both types of information to be present.

Overall, my analysis provides an equilibrium-based theoretical understanding for the empirical finding that individuals maintain incorrect perceptions of their networks. A limited "horizon of observability" (low x and y) prevents an individual from identifying errors in her beliefs about the network, and if those incorrect beliefs are not contradicted by her limited information, then those incorrect beliefs can persist. Because this result has less to do with the specific strategic nature of the particular network game examined here than with the idea of persisting incorrect beliefs, it applies quite generally to various network settings. Thus, real-life networks in which individuals have very limited observation will likely be inefficient. Moreover, attempts to overcome the effects of perfect monitoring and incomplete information should take into account the different effect each type of uncertainty has on network efficiency.

This paper is the first game theoretic study of both imperfect monitoring and incomplete information in social networks. McBride (2006a) is the first paper to relax the full information assumption using a non-Nash Equilibrium concept. Using this same basic model, that paper shows how inefficiencies arising from imperfect monitoring (but not incomplete information) can be overcome by placing certain natural restrictions on players' actions and beliefs about others' actions. This paper differs by further adding incomplete information and by introducing a new issue—how different types of uncertainty lead to different types of inefficiencies in network equilibria. McBride (2006b)[4] adapts the GCE concept introduced herein to a mutual consent network formation setting.[5]

[4]Although McBride (2006b) was published before this paper, this paper was actually written first. McBride (2006b) cites and builds upon this paper.

[5]McBride (2006a), this paper, and McBride (2006b) should properly be understood as the first, second, and third papers, respectively, in a series of papers on limited information in social networks.

1 The Basic Model

Each player $i \in N = \{1, ..., n\}$ knows a "fact" that is valued by all other players. The fact might be investment information, valued news, insights into a new productive technology, etc. Let v_i denote the value of i's fact to all players, which can also be thought of as i's type. At the beginning of the game, each i is assigned her value $v_i \in [0, \infty)$ from an unknown distribution F_v. Let $v = (v_1, ..., v_n)$ be the realized profile of values, and let V denote the set of possible profiles.

Player i learns j's fact either through a direct bilateral tie with j or indirectly through a path of other players' direct ties. The tie exists if one or both players initiate a communication link, and initiating a link costs $c < \infty$ to the initiator. The cost captures the time, effort, or money invested to form and maintain the link. Each i must choose with whom to initiate links. Player i's strategy is thus a vector $s_i = (s_{i1}, ..., s_{ii-1}, s_{ii+1}, ..., s_{in})$, 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. Denote S_i the set of i's possible link decisions, and let $S = S_1 \times ... \times S_n$. Letting $s = (s_1, ..., s_n) \in S$ be one possible profile of link choices, s thus implies a network structure or *graph*. A node in the graph represents an individual player, and the bilateral ties between nodes represent communication links.

Say that there is a *path* between i and j if they are directly linked, $\max\{s_{ij}, s_{ji}\} = 1$, or if they are indirectly linked, i.e., there exist players $j_1, ..., j_m$ distinct from each other and from i and j such that

$$\max\{s_{1j_1}, s_{j_11}\} = \max\{s_{j_1j_2}, s_{j_2j_1}\} = ... = \max\{s_{j_mj}, s_{jj_m}\} = 1.$$

Define a network *component* to be a subset of players such that there exists a path between any two players in the subset, and there is no path between a player in the subset and a player not in the subset. Given structure s, denote $N_i \subseteq N$ to be i's component. Further denote $I_i = \{j \in N | s_{ij=1}\}$, which is the set of i's link initiations, and let $|I_i|$ be the number of people in that set.

Each player has an identical utility function $u_i(s_i, s_{-i}|v) = \sum_{j \in N_i} v_j - |I_i| c$. Thus, the value of j's fact to i does not depend on how many links away j is in i's component so long as she is in i's component (see Section 5 for a consideration of *flow decay*). Also, the value of j's fact to i does not depend on how many other players are in i's component.

To summarize, the network game has the following timing:

1. Each player is privately assigned her type (value) v_i according to F_v, which results in the type vector v.

2. Each player simultaneously chooses s_i, which results in the network structure s.

3. Each player receives some information (described below) about s and v.

Figure 1(a) illustrates one possible s. An arc between two nodes represents a communication tie or link, and, following convention for this model, the dot indicates which side of the tie was the link initiator. For player 1, $N_1 = \{1, 4, 5, 6\}$, $I_1 = \{5, 6\}$, and $|I_i| = 2$. Note that players in the same component can have different utilities since they may each make a different number of link initiations, e.g., if $v_i = \overline{v}$ for all i, then $u_1 = 4\overline{v} - 2c$ but $u_4 = 4\overline{v} - c$.

One important feature of a network is its connectivity. Figure 1(b) adds $s_{12} = 1$ to 1(a). A network such as this in which all players are in one component (i.e., $N_i = N$ for all i) is called *connected*. Figure 1(a) is *disconnected* because $N_i \subset N$ for all i. A special case of a disconnected network is the *empty* network in which each player is *isolated*, i.e., $N_i = \{i\}$ for all i. A second important feature is whether or not a network has redundant link initiations called *cycles*. Figure 1(c) is identical to Figure 1(a) but with two redundant link initiations removed: $s_{15} = s_{45} = 0$. A network without redundant link initiations is called *minimal*. Figure 1(d) depicts a *minimally connected* network, which is special because any efficient (maximized sum of utilities) network that is non-empty must be minimally connected.

Proposition 1: *Fix $v \in V$. If $\sum_{i \in N} v_i > c$ then the set of efficient networks is the set of minimally connected networks. If $\sum_{i \in N} v_i < c$ then the unique efficient network is the empty network.*

Proof: Suppose a non-empty efficient network. Minimality follows directly since the distance between links does not matter so long as a path exists. Connectedness also follows: if an efficient network had a component N_i with $n_i < n$, then it must generate value $\sum_{i \in N_i} v_i$ greater than cost c; but adding a link between i and $j \notin N_i$ will generate value $\sum_{i \in N_i} v_i$ to j, which means net social utility must go up since the link costs only c. Thus, any non-empty efficient network must be minimally connected.

Given v, whether an efficient network is empty or minimally connected will depend on which generates the highest sum of utilities. Any minimally connected network generates $n \sum_{i \in N} v_i - (n-1)\, c$, and the empty network generates $\sum_{i \in N} v_i$. Comparing these gives the proposed condition.

There are many advantages to using this particular model. First, it captures many features of actual social networks. Valuable information is often communicated through informal networks, and these networks arise from the uncoordinated decisions of individuals to form or sever ties. Moreover, communication networks often extend across large geographical boundaries, so it is likely that an individual will not observe the communication ties between other individuals (imperfect monitoring) nor the specific value of another person's information (incomplete information). Thus, this model is rich enough to capture these informational considerations, but also simple enough to conduct formal analysis. Finally, previous work has studied this model under full information [Bala and Goyal (2000), Galeotti, Goyal, and Kamphorst (forthcoming)] and imperfect monitoring [McBride (2006a)], so it serves as a useful point for comparison.

These advantages noted, there is version of limited information that could be considered. Galeotti, Goyal, and Kamphorst (forthcoming) consider the full information model with heterogeneity in values and link formation costs, and they find that each type of heterogeneity has different implications for equilibrium structure. My focus on just value heterogeneity relates to the communication network story underlying my model. Because communication technology is fairly common among individuals, their link formation costs should be similar. It is heterogeneity in the value of information that distinguishes one individual from another. That said, future work should look at the heterogeneity in link cost.

2 The Full Information Nash Equilibrium Benchmark

In the full information case, each player fully observes s and v in period 3. Thus, no i will change her link decisions if, given v, her choice s_i is a best response to what the others actually do s_{-i}. This is exactly a Nash equilibrium (I examine only pure equilibria):

> **Definition 1:** *Fix v. A (pure)* ***Nash Equilibrium (NE)*** *of the network game is a strategy profile $(s_i^*)_{i \in N}$ such that for each $i \in N$, $u_i\left(s_i^*, s_{-i}^*|v\right) \geq u_i\left(s_i', s_{-i}^*|v\right)$ $\forall s_i' \in S_i$.*

It follows that the set of network equilibria under full information is the set of NE. An additional definition will help to characterize this set.

> **Definition 2:** *Fix (v, s). Say that component N_i in s has a* ***low-valued, link-receiving subcomponent (LLS)*** *if there exists an $i, j \in N$* with $s_{ij} = 1$, *such that (i) setting $s_{ij} = 0$ partitions N_i into two separate components N_i' and N_j', and (ii) $\sum_{k \in N_j'} v_k < c$.*

Intuitively, any i would remove a link to j if i knows j is in a LLS since the link provides marginal benefits less than the marginal cost c. We can now describe $E_{full}(v)$, the set of network structures that can be sustained as equilibria under full information given type profile v.

> **Proposition 2:** *Fix v. If $v_i < c$ for all i, then $E_{full}(v)$ contains the empty network and all LLS-free minimally connected networks. If $v_i > c$ for at least one i, then $E_{full}(v)$ contains all LLS-free minimally connected networks.*

The main logic is straightforward and follows from that used by Bala and Goyal (2000) in their examination of the symmetric types case. If the network is not minimal, then there must be a redundant link that could be removed to make someone better off. Thus, any network must be minimal. If the network is not empty but is disconnected, and if all individuals in their disconnected components prefer remaining in their component to removing all their links, then connecting separate components makes all parties strictly better off. The reason is that the value of any component must exceed c for an individual to want to remain in it, and this implies that connecting separate components will be in all parties' interests. Finally, if all types are low valued then the empty network is an equilibrium because initiating a link to an isolated players yields few benefits.

Although Proposition 2 completely characterizes the set of Nash networks, its description is not very detailed because the set of LLS-free minimally connected networks changes as v changes (all else constant). We can say more about $E_{full}(v)$ and obtain intuition about the model by making assumptions about v. In general, as the v_i's increase, the set of Nash networks increases. Technically speaking, $E_{full}(v) \supseteq E_{full}(v')$ if $v_i \geq v'_i$ for all i. The intuition follows from Proposition 2: if you pick some g with LLSs and then increase the values, then those LLSs might no longer be LLSs, which would mean they are sustainable as an equilibrium. In fact, the number of equilibria becomes extremely large as more players become high-valued ($v_i > c$) because a link-receiving subcomponent with high-valued player is high-valued and not a LLS.

Perhaps the easiest case to see this is the symmetric case in which $v_i = v_0$ for all i, first examined by Bala and Goyal (2000). If $v_0 < c$, then, by Proposition 2, the empty network is a Nash network. In fact, if v_0 is sufficiently small, then the empty network is the only Nash network. This occurs when the highest possible benefits to be gained through linking, $(n-1)\,v_0$, are less than the cost of forming even a single link, c. If v_0 increases just enough so that $v < c < (n-1)\,v_0$, then the periphery-sponsored star depicted in Figure 2(a) is a NE. In fact, the periphery-sponsored star could be the only connected NE. To see this, imagine $(n-1)\,v_0 = c + \varepsilon$, with $\varepsilon > 0$ small. The periphery-sponsored star in Figure 2(a) is a NE, but the network in Figure 2(b) is not. The subcomponent comprised of players 1, 2, 5, and 6 constitutes a LLS for player 4 because i's benefit from the link with 5 would yield $(n-2)\,v_0$, which with ε sufficiently small would be less than c. Once v_0 increases enough so that $(n-2)\,v_0 > c$, then the network in Figure 2(b) becomes a NE because the 1-2-5-6 subcomponent is no longer a LLS. Eventually, as v_0 increases even more to exceed c, the empty network is no longer a NE, but any minimally connected network is. Even the center-sponsored star depicted in Figure 2(c) in which one player pays for all links is now an equilibrium.

The following corollary, though not comprehensive, illustrates that this basic logic extends to the more general situation of asymmetric types.

Corollary 1: *Fix v.*

(a) If $\sum_{j \neq i} v_j < c$ for at least two i and $\sum_{j \neq i} v_j \geq c$ for all other i, then the empty network is the unique Nash network.

(b) If $\sum_{j \neq i} v_j < c$ for exactly one i and $\sum_{j \neq i} v_j \geq c$ for all other i, then the connected, periphery-sponsored star with i in the center is the only Nash network.

(c) If $v_i \geq c$ for all i, then $E_{full}(v)$ contains the set of all minimally connected networks.

Proving this corollary is straightforward using logic similar to that given when discussing the symmetric case. The key difference is that $\sum_{j \neq i} v_j$ is the largest possible benefit that i can receive in a network instead of $(n-1)\, v_0$. Yet, as in the symmetric case, if this sum is less than the cost of forming one link for all i, then there is no scenario in which any i will form any links, and the empty network is the only equilibrium network. If, on the other hand, $v_i \geq c$ for all i, then $\sum_{j \neq i} v_j$ is also greater than c, and any subcomponent must not be low-valued. Any minimally connected network would now be an equilibrium.

Also note that since any minimally connected network is efficient, any non-empty equilibrium is efficient, and these efficient equilibria always exist except when the v_i's are extremely low. Thus, the full information NE comprise a useful benchmark. If we have inefficient non-empty equilibria under incomplete information or imperfect monitoring, then it will be due to the change in information available to the players. Of course, for some v, the empty network might be both the only equilibrium and inefficient. This occurs when both $\sum_{i \in N} v_i > c$ but and $v_i < c$ for all i. Standard reasoning applies: an individual considers only her own marginal benefits of a link and not the social benefits, which due to the positive externalities will exceed her marginal benefits.

3 Incomplete Information and Imperfect Monitoring

3.1 Information and Equilibrium

To mimic the incomplete information and imperfect monitoring present in actual social networks, I introduce the notion of *x/y-link observation*.[6] Each i in period 3 will observe the ties that are within geodesic distance $x \geq 1$ and the values of all players within geodesic distance y. Assume $x \geq 1$ and $0 \leq y \leq x$ for now, although this will be relaxed later. These are natural restrictions because an individual should know with whom she has direct links at the least ($x \geq 1$), and she should only be able to observe another's type if she observes that type in the network ($y \leq x$). Moreover, observing another's links does not imply that that the type can be observed (y can be strictly less than x). In Figure 3(a), the dashed line encloses the links observed by player 1 when $x = 1$, and the boxes show the types 1 observes when $y = 0$. 1 observes her link initiation to 6, 5's link initiation to her, her own type, and nothing else. She does not observe 5 or 6's types.

Clearly, as x or y or both increase, each i observes weakly more of her network. In Figure 3(b) with $x = 2$ and $y = 2$, 1 observes all that she did when $x = 1$, but also 5 and 6's types, 5 and 6's direct links, and the types of the people on the other side of those direct links—player 4's type in this case. Also, i never observes any player not in her component if x and y are finite, e.g., 1 observes all links and types in her component but never observes the link between 7 and 8 when $4 \leq x, y < \infty$. Finally, note that x/y-link observation mimics to some extent what individuals observe in actual social networks. Individuals gain their information about the network through their own network interactions, and they are more likely to observe the parts of the network closer to themselves.[7]

Because x/y-link observation yields a network game with both incomplete observation and imperfect monitoring, we want an equilibrium concept that allows for these deviations from full information. The NE concept restricts players to commonly know v and s in equilibrium, and this restriction is a strong one to make, especially in a setting where information is limited and specific to an individual's position in the network. Nash refinements, as NE

[6] This x/y-link observation concept generalizes the x-link observation concept introduced by McBride (2006a), which formalizes imperfect monitoring, to further allow for incomplete information.

[7] See Friedkin (1983), Kumbasar, Romney, and Batchelder (1994), Bondonio (1998), and Casciaro (1998).

themselves, make the same restriction. Other variations on the NE concept also make a similar restriction. For example, the Bayesian NE concept for games of incomplete information assumes that players commonly know the prior probability distribution over types.

Instead of restricting players to have this convergence in beliefs (or in beliefs about priors), I generalize the *Conjectural Equilibrium* (CE) concept, which is designed for games with imperfect monitoring, to also allow for incomplete information. I will formally define the *Generalized Conjectural Equilibrium* concept after first defining a game of incomplete information and imperfect monitoring. I will then describe how I apply this concept in the network game.

Definition 3: *A* ***game of incomplete information and imperfect monitoring*** *is a combination*

$$\left\langle N, \Theta, A, \Pi, (u_i)_{i \in N}, (m_i)_{i \in N} \right\rangle,$$

where: N is a set of players; Θ is a set of states; A_i is the set of actions for $i \in N$ and $A = \times_{i \in N} A_i$; Π_i is i's set of probability distributions over $\{A, \Theta\}$ and $\Pi = \times_{i \in N} \Pi_{i \in N}$; $u_i : \{A, \Theta\} \to R$ is i's utility function; and $m_i : \{A, \Theta\} \to M_i$ is i's signal or message function with message space M_i.

The inclusion of signal functions distinguishes this from a standard game of incomplete information. The signal function is used in a games of imperfect monitoring to formalize what actions a player monitors or observes. Here, however, the signal function also formalizes the subset of types a player observes.

As explained by Gilli (1999) when discussing the CE concept, it is commonly assumed that an individual's signal includes, at a minimum, the individual's own payoff, which means that any states that yield the same signal must also yield the same utility (though the reverse need not be true). This assumption can be specifically motivated for the network formation game. If we think of an equilibrium network as one that may persist over time, then it should be one in which an individual will not want to change her links even after she knows the benefits she receives from network participation in any given period.

While a general game of imperfect monitoring and incomplete information need not have additional restrictions on signal functions, x/y-link observation does impose additional restrictions. Under x/y-link observation, the message m_i is that part of the network that is observed by i given x and y. Because i's observation is specific to her position in the network, each player will generally observe a different part of s and v, e.g., compare 1's observation depicted in Figure 3(a) with 2's observation in the same network depicted in 3(c). Moreover, different networks can give a player the same message, e.g., 1's message is identical in Figures 3(a) and 3(d), assuming they have the same v.

We can now define the equilibrium concept used herein.

Definition 4: *Fix $\theta \in \Theta$. A **Generalized Conjectural Equilibrium (GCE)** is a profile of actions and beliefs $(a_i^*, \pi_i^*)_{i \in N} \in \{A \times \Pi\}$ such that for each i:*

(i) $\sum_{(a'_{-i},\theta') \in \{A_{-i} \times \Theta\}} \pi_i^* \left(a_i^*, a'_{-i}, \theta'\right) u \left(a_i^* | a'_{-i}, \theta'\right) \geq \sum_{(a'_{-i},\theta') \in \{A \times \Theta\}} \pi_i^* \left(a''_i, a'_{-i}, \theta'\right) u \left(a''_i | a'_{-i}, \theta'\right)$ $\forall \ a''_i \in A_i$,

(ii) For any $(a', \theta') \in \{A \times \Theta\}$ s.t. $\pi_i^* (a', \theta') > 0$, it must be that $m_i (a', \theta') = m_i^* (a^*, \theta)$.

(iii) For any $(a', \theta') \in \{A \times \Theta\}$ s.t. $u_i (a' | \theta') \neq u_i (a^* | \theta)$, it must be that $\pi_i^* (a', \theta') = 0$.

Condition (i) states that in equilibrium each i's action a_i^* must be a best response given her conjectured beliefs π_i^*. Condition (ii) places a restriction on each i's beliefs: for any state of the world (a', θ') that i assigns non-zero probability, it must be true that the message i receives in that state equals the message i receives in the true state of the world (a^*, θ). In other words, a player's beliefs must not contradict her message. Condition (iii) restricts beliefs so that i cannot assign non-zero probability to any state of the world that her knowledge of her own utility tells her cannot be the true state of the world. This condition follows from the assumption mentioned above that signals reveal at least as much information as the payoff.

Note that a CE is a GCE of a game with complete information (e.g., $\Theta = \{\theta\}$ or θ publicly observed), so the GCE concept generalizes the CE concept to games of incomplete information. Also note that the GCE does not by itself impose common knowledge of rationality, a fact which, as will be seen below, has non-trivial implications for some of the results.[8]

It might seem counterintuitive that the GCE (and the earlier CE) concept depicts an equilibrium as resulting from simultaneous moves but post-move monitoring. However, we can understand and motivate the GCE concept in two ways similar to how we think of the NE concept. The first way is to recognize that, as a static stability concept, the GCE definition does not assume how beliefs about others' actions and types are formed, it only says that a particular profile of beliefs and actions constitutes an equilibrium if certain conditions (best response and beliefs consistent with signals) are met. The NE is similar in this regard because it does not itself claim how actions are formed but only says that under certain action profiles no player has an incentive to deviate. The second way is to motivate the GCE concept by thinking of a dynamic learning environment. This is exactly how the CE and another non-Nash concept, *Self-Confirming Equilibrium*, have been motivated in prior work (e.g., Gilli 1999 and Fudenburg and Levine 1993). The basic idea is to imagine that the static game is repeated over time, and players receive only limited information about other players' types and strategies as the game progresses. In this case, players could have incorrect beliefs that persist over time, play best responses to those beliefs, and have no incentive to change their beliefs if they never receive information to contradict those beliefs. An equilibrium could thus be achieved in which individual maintain incorrect beliefs. Similar learning stories are commonly used to motivate the NE concept, the difference being that players in a GCE are allowed to not fully learn all players' actions.

Given v, it will often be possible to have multiple beliefs profiles that, when combined with s^*, meet GCE conditions (i)-(iii). To see this, suppose $v_i = \overline{v} > c$ for all i, and that Figure 3(a) is the s formed in stage 2. Given $x = 1$ and $y = 0$ depicted, setting π_1^* to assign probability 1 to the (s, v) combination in Figure 3(e) will satisfy conditions (i)-(iii) for player

[8]See Battigalli, Gilli, and Molinari (1992) and Gilli (1999) for extended discussions of the Conjectural Equilibrium (CE) concept in games with imperfect monitoring. While I generalize the CE concept, it has also been further refined. Fudenburg and Levine's (1993) Self-confirming Equilibrium is a CE in which i's signal contains the strategies that all others play at all information sets on the equilibrium path. Rubinstein and Wolinsky's (1994) Rationalizable Conjectural Equilibrium further assumes common knowledge of rationality.

1, and so will π_1^* that assigns probability 1 to the (s, v) combination in Figure 3(f). We can find similar beliefs for the other players to meet conditions (i)-(iii) for them. Because different beliefs can often be combined with the same s^* to make an equilibrium, I will say that structure s^* is an equilibrium network if there exists a beliefs profile π^* such that the $(s_i^*, \pi_i^*)_{i \in N}$ combination is a GCE that also meets Condition (iii). I will find the set of all these equilibrium structures given x/y observation and type profile v, $E_{x/y}(v)$.

As will be seen, the primary disadvantage of using a non-Nash concept is that it places so few restrictions on equilibrium beliefs that it often leaves a very large set of equilibria. However, there are advantages to this approach. Because conditions (i)-(iii) place so few restrictions on equilibrium beliefs, they can be seen as the minimum necessary conditions for a network equilibrium. This also guards against making restrictions difficult to justify in the network setting with limited observation. The underlying question concerns what should i assume about the actions and types of individuals she does not observe. Restricting players to have common knowledge of F_v, for example, as assumed in a Bayesian Equilibrium, is not appropriate since it is difficult to justify when each player observes a different part of the network.

3.2 Characterization of Network Equilibria

An immediate implication of the GCE and x/y-link definitions is that any equilibrium structure under a certain level of information is also an equilibrium under less information.

Lemma 1: *Fix v.*

(a) $E_{full}(v) \subseteq E_{x/y}(v)$ for all $x \geq 1$ and y such that $0 \leq y \leq x$.

(b) $E_{x/y}(v) \subseteq E_{x'/y'}(v)$ for all x' and y' such that $1 \leq x' \leq x$ and $0 \leq y' \leq y$.

The key logic behind Lemma 1 is that larger observation places greater restrictions on the equilibrium beliefs, so if a network meets the stricter restrictions of higher observation it will meet the looser restrictions of lower observation. For part (a), it is immediately apparent from comparing Definitions 1 and 4 that a Nash Equilibrium s^* is an equilibrium under x/y-link observation when $\pi_i^*(v, s^*) = 1$ for all i since full information restricts each player's equilibrium beliefs to be correct. Similar logic yields (b). Fix s^*, x, and x' with $x' < x$, and hold y fixed. Whatever i observes under x' is also observed under x, but not the reverse. As such, any restrictions on π_i^* created by x'/y-link observation will also be restrictions on π_i^* created by x/y-link observation, but x/y-link observation will have even more restrictions. Thus, beliefs that can be sustained under x will also be sustained under x', so any $(s_i^*, \pi_i^*)_{i \in N} \in E_{x/y}(v)$ will also be in $E_{x'/y}(v)$. Of course, the additional restrictions under x and y will further refine the set of equilibria, so that there may be equilibria in $E_{x'/y}(v)$ not in $E_{x/y}(v)$.

Lemma 1 establishes that limited observation does not necessarily prevent a network from achieving what it could under full information, such as efficient, minimally connected networks. However, we cannot explain how individuals would happen to form accurate beliefs, particularly when they have very limited observation. When observation is limited, individuals may sustain incorrect beliefs in the form of assigning non-zero probability to a state s' not equal to the true state s^* in equilibrium. As observation decreases, many inefficient non-Nash equilibria will exist because players are less likely to observe whether a link initiation is worthwhile or not, i.e., a player is less likely to observe the presence of a cycle or a LLS.

First consider cycles. An equilibrium network cannot have cycles that are too small, but it can have cycles outside of players' observational ranges. Suppose $v = \{a, a, a, a, a, a, 2a, 2a\}$, $a < c < 2a$, and let s^* be the network in Figure 3(a). If π_1^* assigns probability 1 to the structure and v' depicted in Figure 3(e) where $e < c$, then conditions (i)-(iii) are met for her. This works because 1 believes that she is not initiating a link in a cycle and that any new links to any other player would be either redundant or to a low-valued player. Because 1 observes so little, her beliefs are not contradicted by what she observes nor by her utility. As x increases to 2 or higher, however, this cycle cannot exist in equilibrium. Since a player sees along both directions of a cycle, any cycle of size $2x$ or smaller will be observed, but any cycle $2x + 1$ or larger will not be observed.

LLSs may also exist in equilibrium. Clearly, an LLS of size $y - 1$ or smaller would be within i's observational range and thus could not exist in equilibrium. But conditions (ii) and (iii) may also "identify" other LLSs even if they are not explicitly observed. It turns out that two conditions must be met to not be "identifiable:" the player must have sufficient "unaccounted for utility" that could possibly (from the player's point of view) be attached to the LLS thus making the LLS not really a LLS, and there must be players whose values are not observed but who could be generating that unaccounted for utility.

Proving these claims requires new notation. First, some preliminary definitions. Let $N_i^{x/y}(s,v)$ be the subnetwork that comprises exactly what i's signal m_i reveals. For $j \in I_i$, let $P_{ij} = \{k \in N | \text{there is a path from } i \text{ to } k \text{ through } j\}$, and let $P_{ij}^{x/y} = \{k \in N_i^{x/y}(s,v) | \text{there is a path from } i \text{ to } k \text{ through } j\}$. In words, P_{ij} is the set of all players on the path from i to j, while $P_i^{x/y}$ is the set of all players on the path that are in $N_i^{x/y}$. Note that k is in P_{ij} but not in $P_{ij}^{x/y}$ if the path through j exists but is not observed. Also, let $d(i,j|s)$ be the shortest path distance between i and j in s, where $d(i,j|s) = \infty$ if $j \notin N_i$ and $d(i,i|s) = 0$. The key definitions are the following.

- Let $v_{ij+}(s,v)$ be the actual—but possibly not observed—marginal value of i's link to j:
$$v_{ij+}(s,v) = \begin{cases} 0, & \text{if } j \text{ and } i \text{ are in a cycle} \\ \sum_{k \in P_{ij}} v_k, & \text{otherwise} \end{cases}.$$

- Let $v_{ij+}^{x/y}(s,v)$ be the observed marginal value of i's link with j:
$$v_{ij+}^{x/y}(s,v) = \begin{cases} 0, & \text{if } j \text{ and } i \text{ are in a cycle of size } 2x \text{ or smaller} \\ \sum_{d(i,k) \leq y, k \in P_{ij}^{x/y}} v_k, & \text{otherwise} \end{cases}.$$

- Let $LLS_i^{x/y} \equiv \left\{ j \in I_i | v_{i,j+}^{x/y}(s,v) < c \right\}$ be the players to whom i initiates links that could be LLSs given her observation. Notice that $j \in LLS_i^{x/y}$ does not imply that j is indeed part of a LLS; it just implies that j *might* be part of a LLS.

- Let $\widetilde{v}_i^{x/y}(s,v) \equiv u_i(s) - \sum_{j \in N \text{ s.t. } d(i,j|s) \leq y} v_j$ be "i's utility that is received but unaccounted for given her observation."

- Finally, define $\widetilde{v}_{ij+}^{x/y}(s',v') = v_{ij+}(s',v') - v_{ij+}^{x/y}(s,v)$. Note that $v_{ij+}^{x/y}(s,v)$ is known by i, but since i might assign non-zero probability to some state (s',v'), $\widetilde{v}_{ij+}^{x/y}(s',v')$ captures what i would see as that part of the marginal value of the link with j that is unobserved in state (s',v'). Note that if the link is redundant in state (s',v') then $\widetilde{v}_{ij+}^{x/y}(s',v') = -v_{ij+}^{x/y}(s,v)$.

I demonstrate these notations with Figure 4(a), where the "b" next to 1 means $v_1 = b$, and so on, assuming $e < c < b$. In that network, $N_1^{2/1}$ would be depicted as Figure 4(b), since 1 does not observe the link between 3 and 4. $P_{12}(s,v) = \{2,3,4\}$ and $P_{12}^{2/1}(s,v) = \{2,4\}$. The distance $d(1,3|s) = 3$. The marginal value of 1's link to 2 is $v_{12+}(s,v) = 3e$, and the observed marginal value is $v_{12}^{2/1}(s,v) = e$. The set of 1's possible LLSs is $LLS_1^{2/1} = \{2\}$ since $v_{12}^{2/1}(s,v) = e < c < v_{16}^{2/1}(s,v) = b$. 1's total unaccounted for utility is $\widetilde{v}_1^{2/1}(s,v) = (3b+3e-2c) - (3b-e-2c) = 2e$.

Now we return to analyzing Figure 4(a). With $2 \in LLS_1^{2/1}$, are there beliefs π_1 such that keeping her link with 2 is a best response? For her current action to be a best response, there must exist at least one state s' given non-zero probability by π_i^* in which the link with 2 is neither an LLS or part of a cycle (otherwise 1's best response must involve removing the link), and this s' must have some player k or group of players $k_1, ..., k_z$, connected to 2 who makes 1's link to 2 yield a marginal benefit greater than c to player 1. But this is not possible if e is sufficiently small since there is only $\widetilde{v}_1^{2/1}(s,v) = 2e$ of unaccounted for utility. To meet GCE conditions (ii) and (iii), player 1 must have exactly $\widetilde{v}_1^{2/1}(s,v) = 2e$ of unaccounted for utility in any s' assigned non-zero probability, and even if all of that were gained solely through 1's link with 2, that would still make the link with 2 worth only $3e$ which is less than c if e is sufficiently small. Even though 1 cannot observe exactly who is connected to 2, she would still recognize that the link to 2 is to a LLS because there is not enough received but unobserved utility that could make the link to 2 worthwhile. Thus, for any π_1^* that meets GCE conditions (ii) and (iii), s_1^* would not be a best response.

How much is enough $\widetilde{v}_i^{x/y}(\cdot)$? By expected payoff maximization given beliefs π_i^*, to not remove any link to any $j \in LLS_i^{x/y}$ it is necessary that there is enough $\widetilde{v}_i^{x/y}(\cdot)$ to make the link with any $j \in LLS_i^{x/y}$ not be an LLS:

$$\sum_{(s',v') \in (S \times V)} \pi^*(s',v') \left(\widetilde{v}_{ij+}^{x/y}(s',v') + v_{ij+}^{x/y}(s,v) \right) \geq c.$$

Since $v_{ij+}^{x/y}(s,v)$ must be the same in any (s',v') assigned non-zero probability, this becomes

$$\sum_{(s',v') \in (S \times V)} \pi^*(s',v') \, \widetilde{v}_{ij+}^{x/y}(s',v') \geq c - v_{ij+}^{x/y}(s,v).$$

Because this inequality must hold for all $j \in LLS_i^{x/y}$, we can sum across those j's to get

$$\sum_{j \in LLS_i^{x/y}} \sum_{(s',v') \in (S,V)} \pi_i^* (s', v') \widetilde{v}_{ij+}^{x/y} (s', v') \geq \sum_{j \in LLS_i^{x/y}} \left(c - v_{ij+}^{x/y} (s, v) \right),$$

and since the left hand side cannot exceed the total amount of unaccounted for utility $\widetilde{v}_i^{x/y} (s^*, v)$, it follows that

$$\widetilde{v}_i^{x/y} (s^*, v) \geq \sum_{j \in LLS_i^{x/y}} \left(c - v_{ij+}^{x/y} (s, v) \right).$$

Thus, for i to maintain a link to an LLS, there must be sufficient unaccounted for utility.

In addition to having enough $\widetilde{v}_i^{x/y} (\cdot)$, there must also exist players to whom that unaccounted for utility can be attributed. For example, Figure 4(c) has $\widetilde{v}_1^{x/y} (s, v) = b + e$. While this, if all of this unaccounted for utility is added to 2, would make 1's link to 2 worthwhile, this is not possible in any π_1^* that satisfies GCE conditions (ii) and (iii) since 1 observes by $x = 2$ that 2 has no additional links. Thus, if $y < x$, there must exist a player with unobserved type but who is observed in $N_i^{x/y}$ to be $y + 1$ links away on the path through $j \in LLS_i^{x/y}$. By similar logic, if $x = y$, then there must be some j not observed in $N_i^{x/y}$.

These necessary conditions on cycles, unaccounted for utility, and unobserved players and types are also sufficient.

Proposition 3: *Fix v. $E_{x/y}(v)$ consists of all s^* such that*

(a) s^* has no cycles of size $\leq 2x$;

(b) $\widetilde{v}_i^{x/y} (s^*, v) \geq \sum_{j \in LLS_i^{x/y}} \left(c - v_{ij+}^{x/y} (s^*, v) \right)$ for any i with $\left| LLS_i^{x/y} \right| \geq 1$;

(c) for any i with $\left| LLS_i^{x/y} \right| \geq 1$ (i) if $x = y$ then there exists at least one $k \notin N_i^{x/y}$, but (ii) if $x > y$, then for each $j \in LLS_i^{x/y}$ there exists a player $k \in N_i^{x/y}$ exactly $y + 1$ links away from i through j.

Proof: *Necessity.* See discussion above.

Sufficiency. Fix v, and consider a state s^* that meets (a), (b). Also suppose that $x = y$ and (c-i) is satisfied. Consider i with $LLS_i^{x/y} = \{j_1, \ldots, j_z\}$, $z \geq 1$, and consider k from (c-i). Construct z different states $s^1, \ldots, s^z$ in the following manner: each has subnetwork $N_i^{x/y}$; state s^1 has k initiate a link to player l_1 who is x links away from i through j_1, state s^2 has k initiate to l_2 who is x links away from i through j_2, and so on; and any $j \notin \left\{N_i^{x/y} \cup \{k\}\right\}$ is isolated. Let v' have $v'_j = v_j$ as observed for all $j \in N_i^{x/y}$, $v'_k = \widetilde{v}_i^{x/y}(s^*, v)$, and set $v'_j = \overline{v} < c$ for all $j \notin \left\{N_i^{x/y} \cup \{k\}\right\}$. Notice that (b) is satisfied. Finally, set

$$\pi_i^*\left(s^t, v'\right) = \frac{c - v_{ij_t}^{x/y}}{\sum_{j \in LLS_i^{x/y}} \left(c - v_{ij_t}^{x/y}\right)}$$

for $t = 1, ..., z$. As constructed, $\sum_{(s',v') \in (S \times V)} \pi_i^*(s', v') = 1$, and GCE conditions (ii) and (iii) are met. To show that condition (i) is met, notice that, given π_i^*, the expected cost of removing the link with j_t, $t = 1, ..., z$, must exceed the expected benefit of removing the link:

$$\frac{c - v_{ij_t}^{x/y}}{\sum_{j \in LLS_i^{x/y}} \left(c - v_{ij_t}^{x/y}\right)} \widetilde{v}_i + v_{ij_t}^{x/y} \geq c.$$

With (b) satisfied, it must be true that

$$\frac{c - v_{ij_t}^{x/y}}{\sum_{j \in LLS_i^{x/y}} \left(c - v_{ij_t}^{x/y}\right)} \widetilde{v}_i + v_{ij_t}^{x/y} \geq \frac{c - v_{ij_t}^{x/y}}{\widetilde{v}_i} \widetilde{v}_i + v_{ij_t}^{x/y} = c - v_{ij_t}^{x/y} + v_{ij_t}^{x/y} = c,$$

so keeping the link is a best response. Adding links is not a best response since, according to π_i^*, all $j \notin N_i^{x/y} \cup \{k\}$ are isolated with $v_j < c$ as constructed. Thus, GCE condition (i) holds, which means all GCE conditions are met.

Now suppose $x > y$ and (c-ii) is satisfied. Consider i with $LLS_i^{x/y} = \{j_1, \ldots, j_z\}$, $z \geq 1$. Let state s' have $N_i^{x/y}(s^*, v)$ as does s, but make any $j \notin N_i^{x/y}$ isolated. By (a), s' is minimal. Set $v'_j = v_j$ for values observed given y; set $v'_{k_1} = c - v_{ij_1+}(s^*, v)$ for k_1 from (c-ii), set $v'_{k_2} = c - v_{ij_2+}(s^*, v)$ for k_2, and so on for $k_3, \ldots, k_{z-1}$; set $v'_{k_z} = \widetilde{v}_i^{x/y} - \sum_{t=1}^{z-1} v'_{k_t}$; and set $v'_j = 0$ for all others. Notice that (b) is satisfied and that i's expected payoff under (s', v') equals that under the true state (s^*, v). Finally, set $\pi_i^*(s', v') = 1$. As constructed, π_i^* meets GCE conditions (ii) and (iii). Moreover, given π_i^*, removing any link yields a weak decrease in expected payoff, while adding a link yields a strict decrease in expected payoff. With GCE condition (i) met, it is a GCE.

Finally, consider a player i with $\left|LLS_i^{x/y}\right| = 0$ in a state s^* that meets condition (a). Set $\pi_i^*(s', v) = 1$ where s' equals s^* except that for any cycle of size $2x+1$ or larger in s^*, a link that is $x+1$ away along the cycle is removed so that s' is minimal. As constructed, π_i^* meets GCE conditions (ii) and (iii), and given π_i^*, s_i^* is a best response, thus meeting GCE condition (i).

Notice how the set of GCE networks in Proposition 3 compares with the set described in Proposition 2. In particular, whereas any NE network that is connected must also be minimal, that is not necessarily true for a GCE in which players have incorrect beliefs. Proposition 3 establishes that the incorrect beliefs arising from limited information can lead to network outcomes in which there are too many links (cycles), too few links (disconnected or empty), or both (cycles in non-connected components). This difference arises because the GCE relaxes the NE's restriction that players' beliefs are correct in equilibrium. If players' have incorrect beliefs, then what they believe to be best responses might not in reality be best responses. In this sense, applying the GCE concept allows us to see how the incorrect beliefs resulting from limited information lead to outcomes different than those under full information.

The lowest observation setting, $x = 1$ and $y = 0$, puts Proposition 3 in perspective.

Corollary 2: *Fix v. $E_{1/0}(v)$ consists of all s^* without direct redundant links such that $u_i(s^*|v) \geq v_i$ for all i.*

Proof: Consider 1/0-link observation. Condition (a) of Proposition 3 now says that s^* can have any cycle with 3 or more players, so the only cycles not allowed are direct redundant links. Since $y = 0$ implies i observes only her own type, it follows that $LLS_i^{x/y} = I_i$, $\widetilde{v}_i(s^*, v) = u_i(s^*|v) - (v_i - c|I_i|)$, and $v_{ij+}(s^*, v) = 0$ for all $j \in LLS_i^{x/y}$ for each i. Condition (b) in Proposition 3 thus becomes

$$\begin{aligned} u_i(s^*|v) - (v_i - c|I_i|) &\geq c|I_i| \\ u_i(s^*|v) &\geq v_i. \end{aligned}$$

Now consider condition (c) of Proposition 3. Any s^* must have a player $y+1=1$ links away from i along the path through $j \in I_i$. This is true in any network since j herself is one link away from i. Putting all this together yields Corollary 3.

The $u_i(s^*|v) \geq v_i$ condition is naturally interpreted as a participation constraint. In other words, any network without direct redundant links that makes each i no worse off than being isolated can be an equilibrium under 1/0-link observation. This includes the empty network, non-minimal and disconnected networks, or minimally connected networks. Note that Corollary 2 implies that when $v_i > c$ for all i, every $s \in S$ without direct redundant links is in $E_{1/0}(v)$ since the participation constraint will is met in any such s. That is, *if all players are high-valued then literally any network (without directly redundant links) can be sustained as an equilibrium.* This "anything goes" type result occurs because each i's observation is so limited that we can construct beliefs that make her believe her component is minimal and that any other player not in her component is isolated and low-valued.

4 Incomplete Information vs. Imperfect Monitoring

This section examines the different impacts of incomplete information and imperfect monitoring by comparing two starkly different informational settings. The first setting is ∞/0-link observation where every i perfectly monitors s but only observes values according to y-link observation with $y = 0$. The second setting is 1/∞-link observation where every i knows v, but where they only only observe their own direct links. Although this second setting violates the assumption that $x \geq y$, and thus violates our intuitive sense about the type of information people would have in social networks, it serves as an important case to help us understand the different effects of the two types of limited information.

First consider perfect monitoring. Since perfect monitoring implies that any cycle would be identified and reflected in an individual's beliefs, it follows that any non-minimal structure cannot be an equilibrium. The empty network (which is minimal) is clearly an equilibrium since each i can have π_i^* assign probability 1 to the empty network with $v_j = \overline{v} < c$ for all $j \neq i$. Moreover, a minimally connected network with the participation constraint met will also be an equilibrium since π_i^* can assign probability 1 to a state where her link initiations yield a return higher than the link cost. Non-empty, disconnected networks that are minimal and meet the participation constraint are also equilibria. Even though i observes some component N_j, π_i^* may assign a low value to that component, thus preventing i from wanting to initiate a link to players in the other components. Because she has no direct observation of the value of N_j, this belief meets equilibrium conditions (ii) and (iii).

This last point, although true, ignores the more realistic possibility that i believes that j would only be in some component $N_j \neq N_i$ if her own participation constraint were met, thereby implying to i that the value of N_j exceeds c. Thus, if players attribute rationality to the other players, the disconnected, non-empty networks will not be equilibria, as i will link to $j \in N_j \neq N_i$. Notice that we only need mutual (not common) knowledge of rationality for this result. Proposition 4 summarizes:

> **Proposition 4:** *Fix v. $E_{\infty/0}(v)$ consists of all minimal s^* (including the empty network) such that $u_i(s^*|v) \geq v_i$ for all i. With mutual knowledge of rationality, $E_{\infty/0}(v)$ consists of the empty network and all minimally connected s^* such that $u_i(s^*|v) \geq v_i$ for all i.*

Comparing Proposition 4 with Propositions 2 and 3 reveals that perfect monitoring has an immediate implication about network architecture in that it implies minimality, but that it only affects connectedness if there is mutual knowledge of rationality since only then does it allows players to infer something about the value of other components. Its impact on connectedness is still limited, though, since the empty network is always an equilibrium even if all players are high-valued. Thus perfect monitoring brings us closer to an efficient network by ensuring minimality, but it cannot ensure connectedness.

The results differ dramatically when we have complete information but imperfect monitoring. Consider $1/\infty$-link observation where players commonly observe v but only observe s according to $x = 1$. While the empty network is an equilibrium if $v_i < c$ for all i, it is not an equilibrium when $v_i > c$ for at least one i. Any isolated player's beliefs must assign the high value to any high-valued player, so she could not be isolated in equilibrium. Thus, complete information, unlike perfect monitoring, has immediate implications about connectedness when there is a high-valued player.

Whether or not the equilibrium is connected will depend in a complicated way on a player's ability to discern, using her observation and knowledge of her utility, which players are in her component. Because of this, it is not possible to give a precise necessary and sufficient condition for connected equilibria for generic v. I instead examine two special cases.

The first scenario is one in which each i can identify from her utility alone which others are in her component.

> **Definition 5:** *Say that type profile v is **distinct** if $\sum_{j\in N'} v_j \neq \sum_{j\in N''} v_j$ for all subsets $N', N'' \subseteq N$.*

In essence, each player's fact is sufficiently different from another so that the sum of any combination of players will differ from the sum of any other combination of players. One example of distinct types with $n = 5$ is $v = \{1, 10, 100, 1000, 10000\}$. Notice that if numbers are picked at random from a continuous distribution, then any v profile is generically distinct. The following necessary condition follows immediately:

> **Proposition 5:** *Fix v, and suppose $v_i > c$ for at least one i. If v is distinct, then any $s^* \in E_{1/\infty}(v)$ must be connected, without direct links, and have $u_i(s^*|v) \geq v_i$ for any i.*

If v is distinct and perfectly known by each i, then any π_i^* that meets GCE conditions (ii) and (iii) must identify the players in i's component. If N_i does not include any j with $v_j > c$ (or if i is isolated) then i's current action cannot be a best response since according to π_i^*, initiating a link to that j would make i strictly better off. Of course, cycles outside of the observational range can still exist in equilibria. Thus, distinct types with at least one high-valued player is sufficient to have a connected network, but it cannot guarantee minimality, i.e., complete information here affects connectedness but not minimality.

Now consider a second scenario in which all players have the same type $v_i = \overline{v}$. Player i can now tell exactly how many individuals are in her component, although she may not be able to tell exactly which ones (aside from her direct links). This case with $\overline{v} > c$ and general x is the model studied by McBride (2006a), from which we get the following.

Proposition 6: [From McBride (2006a)] *Fix v with $v_i = \overline{v} > c$.*

(a) If $n < 8$ then any $s^* \in E_{1/\infty}$ is connected, but if $n \geq 8$ and c sufficiently close to v then $E_{1/\infty}$ also contains disconnected networks.

(b) With common knowledge of rationality, every $s^* \in E_{1/\infty}$ is connected for any $n > 1$.

I provide some intuition and an example and refer the reader to McBride (2006a) for the detailed proof. If N_i has three or fewer players, then there will be some player in N_i who not only knows the exact number of others in N_i but also knows the identities of all members. Since v is known, that player knows that she can link to any player not in N_i and be strictly better off. Thus, a component in a disconnected equilibrium must have at least four members. Since there must be at least two component to be disconnected, it follows that a disconnected equilibrium must have $n \geq 8$.

Suppose $n = 8$, and let s^* have two, four-player, non-star components, akin to Figure 5(a). By her utility, 1 knows that there is a player outside her observation that is in her component. Suppose π_1^* is chosen so that π_1^* assigns probability $\frac{1}{5}$ to each of the networks in Figures 5(b)-(f). These beliefs meet equilibrium conditions (ii) and (iii). Moreover, not removing links is clearly optimal, and not initiating any new links to any $j \in \{3, 5, 6, 7, 8\}$ is optimal if

$$c > \Pr[j \notin N_i](\overline{v}) + \Pr[j \in N_i](0) = \frac{4}{5}\overline{v}.$$

So, if c is sufficiently high then s_1^* is 1's best response given our constructed π_1^*. Constructing similar beliefs for the other players gives us a disconnected equilibrium to illustrate part (a).

Notice, however, that if player 1 believes the other players to be rational, then she should not believe any player $j \notin N_i$ to be isolated. The payoff to linking with that j is not $\overline{v}$ but is $2\overline{v}$ or more, thus making the RHS of the above inequality be at least $\frac{8}{5}\overline{v}$. Now the inequality can never hold, so we cannot rationalize a disconnected equilibrium. This logic applies more generally to any n.

We can draw some conclusions from Propositions 5 and 6. Neither perfect monitoring nor complete information are sufficient alone for connectedness or minimality, however each does have a different impact on network efficiency. Complete information of types v, unlike perfect monitoring, implies any equilibrium is non-empty if at least one player is high-valued. Whether or not it guarantees a connected equilibrium will depend, first, on whether players can identify their component members and, second, on whether players ascribe rational behavior to other players. When all players are low-valued, however, complete information has no implications for connectedness. Finally, complete information generally has no implication for minimality. Thus, loosely speaking, complete information and perfect monitoring each provide one aspect of efficiency, but in general neither alone is sufficient for efficiency even when every Nash Equilibrium is efficient.

An important implication of these findings is that developing a mechanism to overcome one type of information limitation but not another (e.g., imperfect monitoring but not incomplete information) may not necessarily be enough to ensure efficiency for any generic network. A mechanism designer will have to consider the specifics of the network setting.

5 Other Considerations

5.1 Strict Equilibria

Economists often focus on strict equilibria when looking for network equilibria in this model [Bala and Goyal (2000), Galeotti, Goyal, and Kamphorst (2004)]. The strictness restriction, which rules out Nash Equilibria in which any player has more than one best response, changes "$\geq$" in condition (i) to "$>$." A practical reason for this restriction is that it greatly refines the set of equilibria, thus simplifying the analysis. A theoretical reason to focus on strict equilibria is their nice dynamic properties in repeated game settings. Actual networks form and evolve over time, and since strict equilibria are absorbing states, we may expect a preponderance of strict equilibria to exist.

The strictness restriction matters when a player's observation reveals that she has another decision that yields the same utility. In practice, this is best thought of as "link-switching." Suppose the players back in Figure 2(b) have full information. If $v_1 > c$, then the figure is a Nash Equilibrium. But it is not a strict Nash equilibrium since player 3 has multiple best responses, e.g., she receives the same utility by switching her link from 4 to 5. This logic applies to any instance where one player initiates a link to another player that also has another link.

Bala and Goyal (2000) prove that the center-sponsored star, like that in Figure 2(c), is the only component structure immune to this link-switching in this model with full information. McBride (2006a) shows that this result extends to the complete information case of x/∞-link observation if $x \geq 2$ because to know of the link-switching opportunity a player only needs observe a neighbor's link. Such is the case in Figure 2(b), which is not in $E_{x/\infty}(v)$ for any $x \geq 2$. Of course, this logic will also apply with incomplete information.

Proposition 7: *If $x \geq 2$, then for any y under x/y-link observation, any strict equilibrium component must be a center-sponsored star.*

Proof: Suppose equilibrium s^* has component N_i that is not a center-sponsored star. Thus, there must exist an $i, j \in N_i$ who both initiate links. If $|N_i| = 2$ then it must be true that $s^*_{ij} = s^*_{ji} = 1$, which is a direct redundant link that cannot exist in an equilibrium with $x \geq 1$.

Now suppose $|N_i| > 2$. Again, there must be $i, j \in N_i$ who both initiate links. If $d(i,j) = 1$ and, without loss of generality, $s^*_{ij} = 1$, then i observes I_j. If $s^*_{ji} = 1$, then i's link is again observed to be redundant and cannot be an equilibrium. If $s^*_{ji} = 0$ then there must be some k such that $s^*_{jk} = 1$. Since i observes this link, i observes that she receives no decrease in utility by setting $s_{ij} = 0$ and $s_{ik} = 1$, so s^* is not a strict equilibrium. If $d(i,j) > 1$, then there must be a player k in the path between i and j, and, without loss of generality, let $s^*_{ik} = 1$. Now i observes that she can switch her link from k to any player directly connected to k on the path to j and receive the same utility, which implies s^* is not a strict equilibrium.

A few things are worth noting. First, center-sponsored stars are minimal, so in any strict equilibrium, each non-empty component is arranged efficiently even if the network as a whole is not efficient. Second, non-empty strict equilibria might not exist even when non-empty weak equilibria do. For example, if $v_i = c - \varepsilon$, $\varepsilon > 0$ small for all $i \in N$, then a periphery-sponsored star is a weak equilibrium (and efficient), but the only strict equilibrium is the empty network. Third, strict equilibria might not exist. If v is distinct, has multiple high- and low-valued players, and is commonly known ($1/\infty$-link observation), then any equilibrium must be connected. However, if $x \geq 2$, then it must also be a center-sponsored star. But this cannot be an equilibrium since an equilibrium cannot have a low-valued stem. This suggests that the strictness refinement, although very powerful and leads to minimal components, may refine too strongly. Finally, although any component is a center-sponsored star, strictness does not imply connectedness. In fact, with x and y finite, the empty network is always a strict equilibrium.

5.2 Decay

Flow decay captures the idea that the benefits of a link depend on the geodesic distance of that link. For example, information in a communication network is lost or altered as it passes through more people, so i's benefits from j are higher if j is closer to i in the network. One way to capture this idea is with the utility function $u_i = \sum_{j \in N_i} \delta^{d(i,j|s)-1} v_j - c\,|I_i|$, where $\delta \in [0,1]$ is the decay factor, and, again, $d\,(i,j|s)$ is the shortest path distance in links between i and j (and $d\,(i,j|s) = \infty$ if $j \notin N_i$).

Not surprisingly, under full information flow decay often reduces the maximum distance between any i and j in the same component of an equilibrium. This is because players want to be closer to one another to reduce decay. This, of course, will depend on the size of the decay. Sufficiently large decay (δ close to 0) will mean that i would rather form a direct link to j than benefit from any indirect link to j. Hence, flow decay can lead to redundant links being efficient in equilibrium. For example, if $\delta = 0$ and $v_i = \overline{v} > c$ for all i, then any Nash Equilibrium must have everyone directly connected to everyone else. Direct redundant links will not exist, but all other redundancies will exist.

These principles will still apply when flow decay exists coincidently with incomplete information and imperfect monitoring, however, the implications of these ideas weaken as x and y decrease. In fact, flow decay has a somewhat minor effect when $x = 1$ and $y = 0$.

Proposition 8: *Fix v, and suppose there is flow decay δ. $E_{1/0}(v)$ consists of all s^* without direct redundant links such that $u_i(s^*|v) \geq v_i$ for all i.*

Proof: *Necessity.* Follows directly since any redundant link would be within a player's observational range, and a player who does not meet the participation constraint is better off removing all link initiations.

Sufficiency. Consider a network that does not have redundant links and where the participation constraint is met. Let π_i^* assign probability 1 to s' where $N_i = N_i^{1/0}$ and all others are isolated, and probability 1 to v' where $v_i' = v_i$, $v_j = \frac{u_i^* - v_i}{|I_i|} + c$ with $\varepsilon > 0$ small for all $j \in I_i$, and $v_k = 0$ for all $k \notin N_i^{1/0}$. With such beliefs, equilibrium conditions (i)-(iii) are met for each i.

Note that the wording used to describe the set of equilibrium networks in Corollary 2 and Proposition 8 is identical. The only difference is that decay reduces the overall value of most components, thus making fewer networks meet the participation constraint. Thus, the set of equilibria with flow decay may be smaller than that without flow decay even though I describe them in similar manners. The idea is akin to that of link switching used to find strict equilibria except here it is link addition instead of switching. A player must observe the presence of another link and be able to tell if the marginal benefit of having a direct link to someone already indirectly connected is better than the marginal cost of the link. With $x = 1$ and $y = 0$, she is not aware of this opportunity. Also notice that increasing x and y leads to more efficient networks even if it means more cycles. This contrasts with the elimination of smaller cycles as x increases when there is no flow decay.

5.3 Mutual Consent Networks

While this paper assumes that a bilateral link is formed by only one person, maintaining a link in many social networks requires the mutual consent and effort of both sides of the tie. It is thus natural to ask how the results presented here compare to the mutual consent setting. McBride (2006b) addresses this question using the x/y-link observation introduced herein and a new stability concept designed for the study of mutual consent network formation.[9]

[9] This paper was actually written before McBride (2006b), but because McBride (2006b) was published first, I can refer to some of its results here.

That paper shows that the general impact of limited observation on network outcomes (i.e., that it expands the set of stable networks) applies in the mutual consent setting for the same reason it does here: relaxing the restriction that beliefs must be correct allows individuals to maintain many other possible actions in equilibrium. Thus, the general result is indeed a general one that applies across network settings.

However, McBride (2006b) also shows that, unlike the model studied in this paper, many of the new equilibria that arise due to limited observation can actually be more efficient that the most efficient ones under full observation. The reason is that there can exist a tension between stability and efficiency in the mutual consent setting, and if individuals can maintain incorrect beliefs in equilibrium, then some of those beliefs may lead them to choose actions that are socially beneficial though individually detrimental. Whether there are unilateral link formation settings in which this result arises is a topic for future research.

6 Conclusion

This paper shows how and why imperfect monitoring and incomplete information lead to the existence of inefficient equilibria. These inefficient equilibria increase in number and degree of inefficiency as observation becomes more limited. Since the empirical work suggests observation to be approximately $x = 2$ and $y = 2$ in many actual informal networks [Friedkin (1983)], my findings predict that many actual networks operate inefficiently. My findings also reveal that limiting players' information about others' ties has different implications than limiting information about players' types.

Although this paper provides the first formal, theoretical explanation for the presence of highly inefficient equilibria due to position-specific information, others have assumed as a starting point that social networks are inherently inefficient for informational reasons. For example, Cross, Nohria, and Parker (2002) take the inefficiency of networks as an implicit starting point in their article for management professionals "Six Myths about Informal Networks—and How to Overcome Them." One of the myths they suggest needs to be overcome is "We can't do much to aid informal network." Since informal networks have structured social interactions for millennia, it is likely that individuals have developed ways to overcome some of the network inefficiencies arising from limited observation. These may involve mechanisms that establish credibility about claims of one's own links. Discovering how individuals act to lessen the network inefficiencies predicted by my analysis constitutes a potentially fruitful avenue of research.

References

[1] **Bala, Venkatesh; Goyal, Sanjeev.** "A Non-cooperative Model of Network Formation." *Econometrica*, September 2000, 68(5), pp. 1181-1230.

[2] **Battigalli Pierpaolo; Gilli, Mario; Molinari, M. Christina.** "Learning and Convergence to Equilibrium in Repeated Strategic Interactions: an Introductory Survey." *Ricerche Economiche*, October-December 1992, 46(3-4), pp. 335-378.

[3] **Bondonio, Daniele.** "Predictors of Accuracy in Perceiving Informal Social Networks." *Social Networks*, October 1998, 20(4), pp. 301-330.

[4] **Calvó-Armengol, Antoni; Jackson, Matthew.** "The Effects of Social Networks on Employment and Inequality." *American Economic Review* , June 2004, 94(3), pp. 426-454.

[5] **Casciaro, Tiziana.** "Seeing Things Clearly: Social Structure, Personality, and Accuracy in Social Network Perception." *Social Networks*, October 1998, 20(4), pp. 331-351.

[6] **Conley, Timothy; Udry, Christopher.** "Social Learning through Social Networks: the Adoption of New Agricultural Activities in Ghana." *American Journal of Agricultural Economics*, August 2001, 83(3), pp. 668-673.

[7] **Cross, Rob; Nohria, Nitin; Parker, Andrew.** "Six Myths About Informal Networks—and How to Overcome Them." *MIT Sloan Management Review*, Spring 2002, 43(3), pp. 67-75.

[8] **Doreian, Patrick; Stokman, Frans N.; eds.** *Evolution of Social Networks.* Amsterdam: Gordon and Breach Publishers, 1997.

[9] **Dutta, Bhaskar; Jackson, Matthew; eds.** *Networks and Groups: Models of the Strategic Formation.* Heidelberg: Springer-Verlag, 2003.

[10] **Friedkin, Noah E.** "Horizons of Observability and Limits of Informal Control in Organizations." *Social Forces*, September 1983, 62(1), pp. 54-77.

[11] **Fudenburg, Drew; Levine, David.** "Self-confirming Equilibrium." *Econometrica*, May 1993, 61(3), pp. 523-545.

[12] **Galeotti, Andrea; Goyal, Sanjeev; Kamphorst, Jurjen.** "Network Formation with Heterogeneous Players." *Games and Economic Behavior*, February 2006, 54(2), 353-372.

[13] **Gilli, Mario.** "On non-Nash Equilibria." *Games and Economic Behavior*, May 1999, 27(2), pp. 184-203.

[14] **Granovetter, Mark.** "The Strength of Weak Ties." *American Journal of Sociology*, May 1973, 78(6), pp. 1360-1380.

[15] **Kranton, Rachel; Minehart, Deborah.** "A Theory of Buyer-seller Networks." *American Economic Review*, June 2001, 91(3), pp. 485-508.

[16] **Kumbasar, Ece; Romney, A. Kimball; Batchelder, William.** "Systemic Biases in Social Perception." *American Journal of Sociology*, September 1994, 100(2), pp. 477-505.

[17] **Laumann, Edward.** "Friends of Urban Men: An Assessment of Accuracy in Reporting Their Socioeconomic Attributes, Moral Choice, and Attitude Agreement." *Sociometry*, March 1969, 32(1), pp. 54-69.

[18] **McBride, Michael.** "Imperfect Monitoring in Communication Networks." *Journal of Economic Theory*, January 2006a, 126(1), pp. 97-119.

[19] **McBride, Michael.** "Limited Observation in Mutual Consent Networks." *Advances in Theoretical Economics*, 2006b, 6(1), Article 3.

[20] **Rubinstein, Ariel; Wolinsky, Asher.** "Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability." *Games and Economic Behavior*, March 1994, 6(2), pp. 299-311.

[21] **Stokman, Frans N.; Doreian, Patrick.** "Evolution of Social Networks: Processes and Principles," in P. Doreian and F.N. Stokman, eds., *Evolution of Social Networks.* Amsterdam: Gordon and Breach Publishers, 1997, pp. 233-250.

[22] **Wasserman, Stanley; Faust, Katherine.** *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press, 1994.

Figure 1

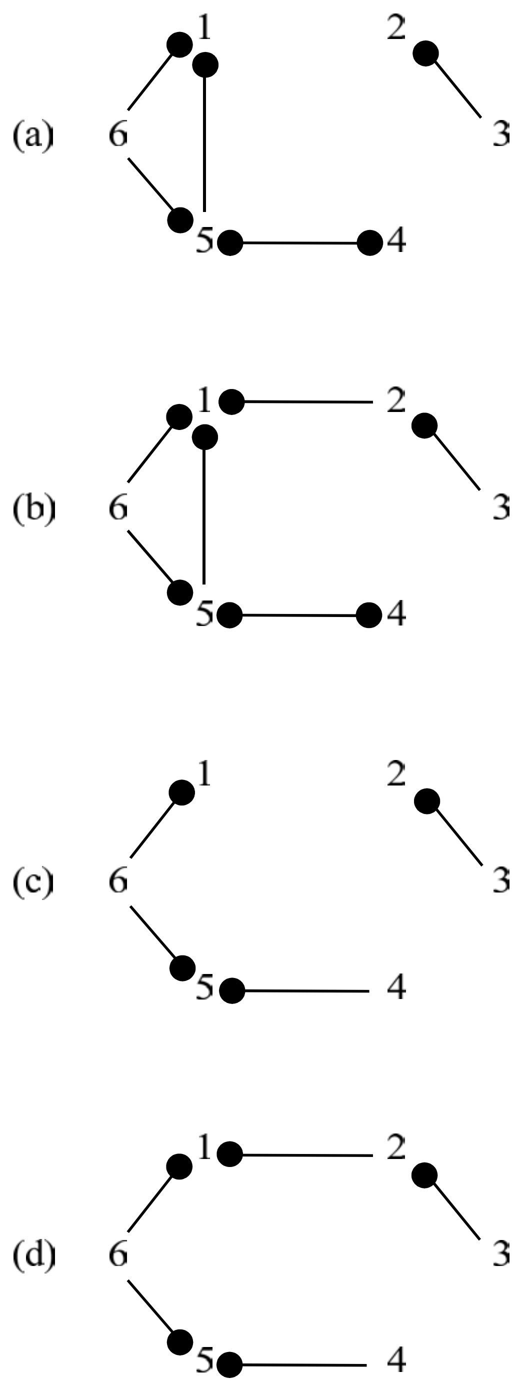


Figure 2

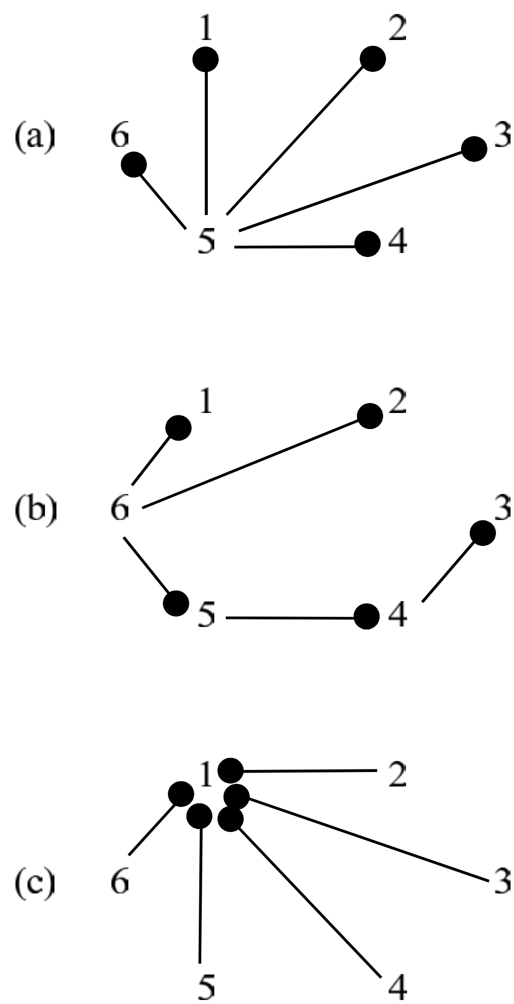


Figure 3

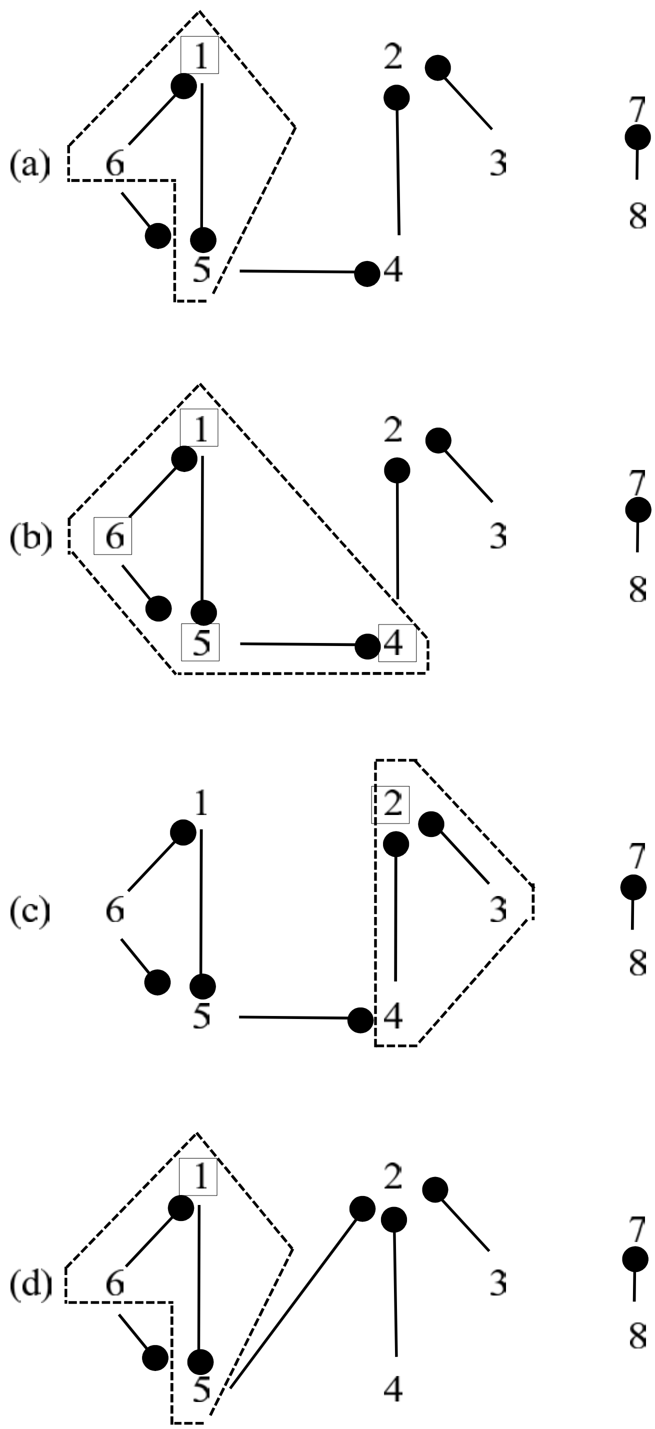
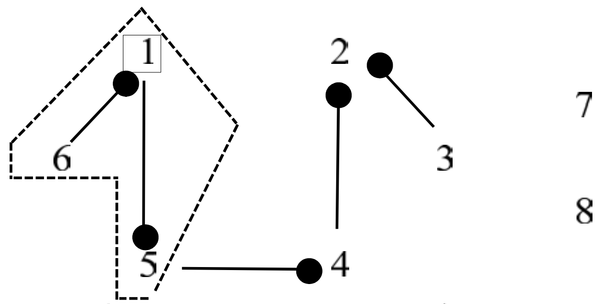


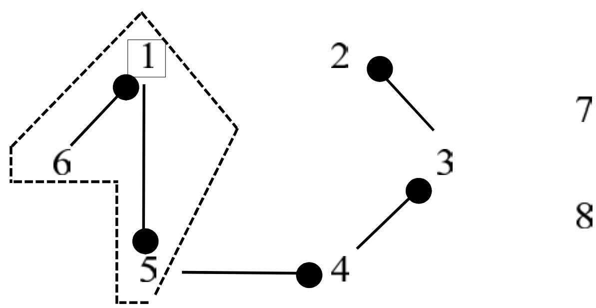
Figure 3 (cont.)

(e)



v' = (v_1, v_2, v_3, v_4, v_5, v_6, e, e)

(f)



v' = (v_1, v_2, v_3, v_4, v_5, v_6, e, e)

Figure 4

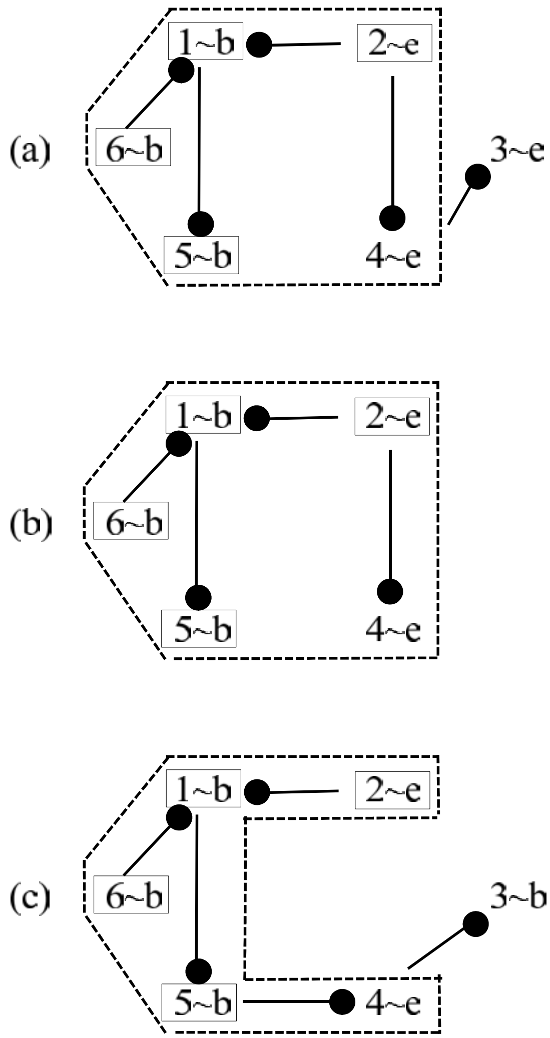
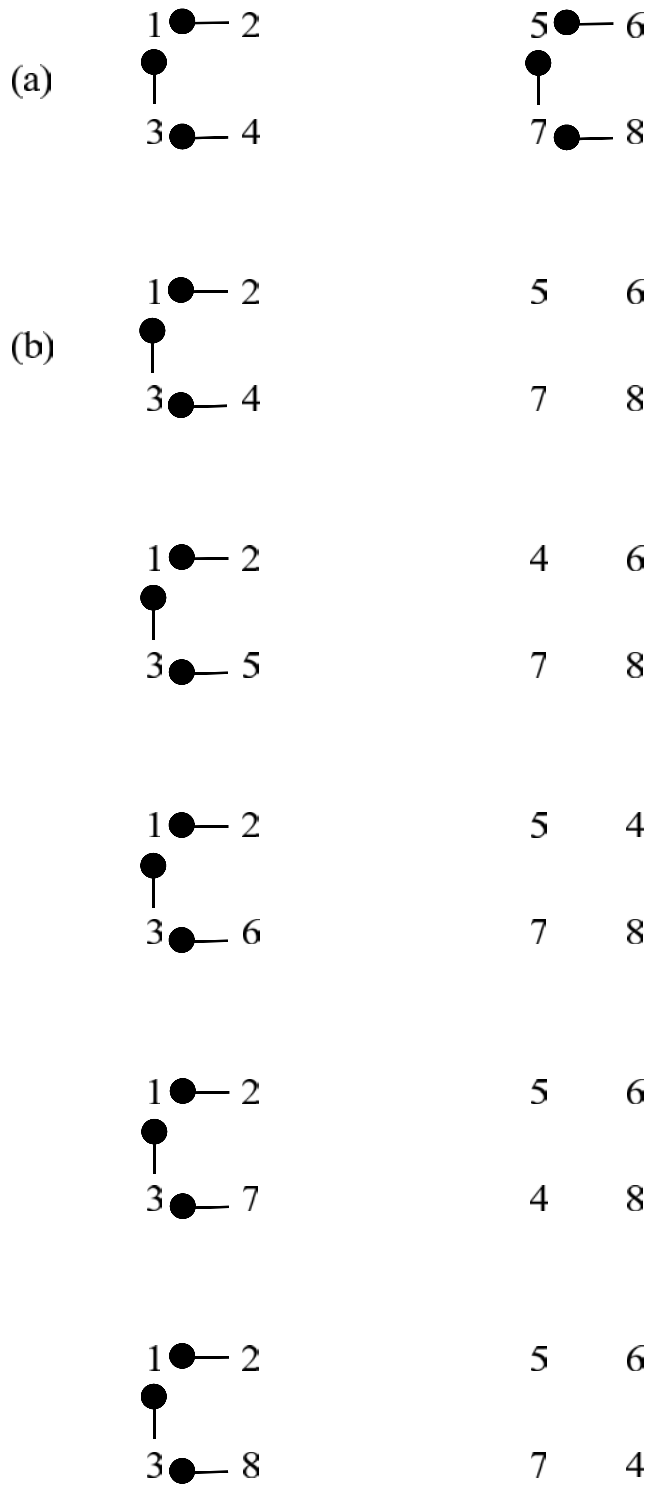


Figure 5


(a)
1 2
3 4
5 6
7 8
(b)
1 2
3 4
5 6
7 8
1 2
3 5
4 6
7 8
1 2
3 6
5 4
7 8
1 2
3 7
5 6
4 8
1 2
3 8
5 6
7 4