ISSN 1936-5098

CAE Working Paper #09-07

What Goes Around Comes Around: A Theory of Indirect Reciprocity in Networks

by

Maximilian Mihm
Russell Toth
and
Corey Lang

May 2009

What Goes Around Comes Around:
A theory of strategic indirect reciprocity in networks (extended version)*

Maximilian Mihm,[†] Russell Toth, and Corey Lang

Cornell University

August 15, 2009

Abstract

We consider strategic interaction on a network of heterogeneous long-term relationships. The bilateral relationships are independent of each other in terms of actions and realized payoffs, and we assume that information regarding outcomes is private to the two parties involved. In spite of this, the network can induce strategic interdependencies between relationships, which facilitate efficient outcomes. We derive necessary and sufficient conditions that characterize efficient equilibria of the network game in terms of the architecture of the underlying network, and interpret these structural conditions in light of empirical regularities observed in many social and economic networks.

JEL classification: D85, C73, D82

Key words: network enforcement, private monitoring, small-worlds, triadic closure

*We thank Chris Barrett, Kaushik Basu, Dan Benjamin, Larry Blume, David Easley, Ani Guerdjikova and Willemien Kets. We also benefited from discussions at an early stage with Adam Brandenburger and Fernando Vega-Redondo, and from comments of participants at the 2009 Econometric Society Summer Meetings in Boston, the 2009 SAET Conference in Ishcia, Italy, and the 20th International Game Theory Conference at Stony Brook. Toth gratefully acknowledges support from the NSF Expeditions in Computing grant on Computational Sustainability, # 0832782. Any errors are the responsibility of the authors.

†Corresponding author. Department of Economics, Cornell University, Ithaca, NY 14853, USA. Email addresses: mam293@cornell.edu (Mihm), rdt28@cornell.edu (Toth), cl395@cornell.edu (Lang).

1 Introduction

Relationships in which individuals exchange favors, share information or trade risks are ubiquitous in social and economic life. These relationships rarely occur in isolation and networks provide a useful way to represent a system of such relations. But do networks also play a substantive role in determining how the relationships function? The interactions between individuals in a network often seem to rely on the principle of indirect reciprocity – the idea that *"You scratch my back and I'll scratch someone else's" or "I scratch your back and someone else will scratch mine"* (Nowak and Sigmund, 2005) – and this suggests that networks may play an important institutional role in coordinating behavior and pooling resources across relationships. But how does indirect reciprocity amongst self-interested economic actors work? How do we make sense of this kind of behavior in a strategic network environment?

In this paper, we show that a network of long-term relationships can facilitate strategic indirect reciprocity when there is heterogeneity in the net benefits from bilateral relations. We derive endogenous strategic connections between relationships and show that a network may sustain efficient outcomes even when each relationship in isolation could not. The relevant constraints are determined by the architecture of the underlying network and the monitoring capabilities across relationships. Our focus is on the case of private information, where individuals only observe the outcomes in their own relationships. The resulting network game can therefore be viewed as a dynamic game of perfect private monitoring (see Kandori, 2002). Under this information structure, we characterize equilibria of a network game in terms of structural properties of the network and provide an intuitive rationale for the small-worlds network structure observed in many real-world networks.[1] Essentially, we show that strategic indirect reciprocity via a network can help to sustain efficient outcomes in a system of asymmetric relations, and that small-worlds are particularly conducive to such network effects because close connectivity enables robust enforcement with low demands on individual monitoring.

An implicit intuition from much of the existing literature on network games is that networks matter if and only if individuals are required to take the same action in all bilateral connections (see, e.g., Galeotti et al., forthcoming; Ballester et al., 2006; Bramoulle and Kranton, forthcoming; Goyal and Moraga-Gonzalez, 2001; Jackson, 2008, chap. 9).[2] As

[1] The key features of small-worlds networks include significant clustering of nodes, small average distances between nodes, and degree distributions with small mean but high variance. Such network structures are the most pervasive regularity observed in empirical network studies (see, e.g., Watts and Strogatz, 1998).

[2] Two important exceptions are Goyal et al. (2008) and Goyal (2005), in which the players can choose

an exogenous constraint on how individuals act, this seems inappropriate in many of the network settings we are interested in. We impose no *a priori* constraints on how individuals behave in different relationships. Instead, the critical feature of the network in our model is that it connects heterogeneous components. All constraints on how individuals act are derived endogenously from the strategic network interaction. As in other work on network games, our view of the underlying network follows the structural view in sociology, where networks are regarded as a primitive of the social or economic environment. Our interest is in the institutions (or norms) that can evolve on such a network, and how these change the incentives regarding behavior in individual relationships. There is an additional growing literature in economics that studies network formation games, modeling the creation of network ties as an organic process (see, for example, Jackson 2005, 2008; Bala and Goyal, 2000). In many settings, insights from network games and network formation games are complementary. For example, the underlying network in our model could be viewed as the outcome of a network formation process that precedes the analysis in this paper. How the resulting network influences strategic interactions is then important for understanding incentive constraints in the formation process. However, modeling an explicit formation process is a distinct exercise which we hope to inform but do not directly address.[3]

Our approach leads to a new perspective on the role of networks in supporting cooperative outcomes, which is applicable in a number of relevant settings. Consider, for example, a network of bilateral relationships which serve as a medium for the exchange of "goods" that cannot be traded in markets, either because the nature of the good makes it arduous to write and enforce formal contracts (such as for information, goods of uncertain quality, or political favors) or because market institutions are underdeveloped (such as the insurance markets in developing countries). In the absence of a central market mechanism, exchange in isolated relationships would be restricted to satisfy a double coincidence of wants because outcomes must be incentive compatible given direct strategic reciprocity between individuals. However, a network provides opportunities for strategic indirect reciprocity across relations, allowing *quid pro quo* to be established across a number of relationships together rather than one relationship at a time. Whenever bilateral exchange does not exhaust the trading opportunities presented by collective exchange, the network can therefore play an important institutional role. The network can pool asymmetries across relationships while still allowing verification of informal agreements to occur within the bilateral relations.[4]

link-specific actions.

[3]Vega-Redondo et al. (2005) and Vega-Redondo (2006) study models which go in this direction. Both study network formation processes driven by the idea that the resulting network plays an institutional role in enforcement mechanisms.

[4]Kranton (1996) analyzes the conflict between direct reciprocity (through repeated bilateral interaction)

As a specific example of the type of network environment we have in mind, consider the networks of information-sharing relationships that exist between firms in many industries and regions, particularly high-innovation industries. Generally, the purpose of these business relationships is to share information acquired in the course of daily business operations, including knowledge of consumer tastes, experience in hiring workers, managerial and technological know-how, as well as information about cultural, political and legal norms (Jarillo, 1988; Bloch, 2008; Mowery et al., 1996).[5] It is clear that the information shared in these relationships is valuable, but sharing the information also involves substantial costs, including direct costs of acquisition, storage and transmission, time spent identifying valuable information with one's partners, as well as the considerable indirect cost of giving up competitive advantage (Easterby-Smith et al., 2008). The fact that information is shared with remarkable ease and frequency between firms in many industries is somewhat of a paradox because the contractual enforcement of agreements is usually impossible in this context (Hagedoorn, 2002). While the persistent exchange of valuable information above and beyond a bilateral *quid pro quo* could be attributed to a number of factors, including altruism or bad management practice, the bilateral perspective taken in much of the management science literature does not take account of the fact that such relationships are often embedded in dense clusters of other, similar relations.[6]

A network perspective allows us to rationalize cooperative behavior, like information-sharing between firms, informal insurance or the exchange of favors, even when there is no apparent bilateral *quid pro quo*. The basic intuition can be explained informally by looking at the networks in Figure 1. Figure 1(a) depicts an isolated bilateral relationship. The nodes (i and j) are players. The arc ij represents a long-term relationship that is asymmetric in

and decentralized monetary trade as two institutions of exchange in an economy. We emphasize that reciprocal exchange may be necessary for goods that are not easily traded in decentralized markets (such as information or favors), or when such markets are underdeveloped, and formalize *indirect* reciprocity in a network of individuals as a means of realizing more efficient exchange by trading off asymmetries in wants; in fact, Kranton suggests this intuition in her conclusion (pg. 846).

[5]Formal and informal information-sharing relationships of this type have become so commonplace that historical scholars now refer to systems of such relations as a primary form of industrial organization, on par with the hierarchical structure and the competitive market (Thorelli, 1996; Lamoreaux et al., 2003).

[6]For a network-based view of strategic business alliances built on dyadic exchanges, see Gulati, 1998. Other examples of networks in which "exchange" is (at least in part) organized through repeated pair-wise interactions include informal social insurance networks in developing countries (Bloch et al, 2008; Bramoulle and Kranton, 2007; Fafchamps and Gubert, 2007), cartels and other collusion networks (Bernheim and Whinston, 1990), political alliances (Knoke, 1990), systems of international relations and trade agreements (Rauch, 1999), and the interconnections between financial institutions (Economides, 1993; Mayer, 2008; Allen and Babus, 2008).

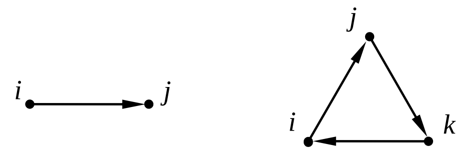


Figure 1: (a) Bilateral relationship (left), (b) Strong triangle network (right).

the sense that it is more valuable to player j than to player i. To fix ideas we can think of ij as an information-sharing relationship between firms, where in each period $t = 0, 1, ..., \infty$ firms can decide whether to share information with their partner, benefits are derived from the information obtained from one's partner, and the transmission of information in each period is costly. The relationship is asymmetric if the net benefits of information sharing do not coincide for the two partners. If all agreements must be enforced through direct reciprocity, the incentive constraints on information-sharing are determined by the payoffs of the player who benefits less from information-sharing (here player i). However, consider what can happen if relationship ij is embedded in a network like the "strong triangle" of Figure 1(b). If all firms in this network share information with both partners, each firm receives low value from one relationship and high value from another. Hence each firm is potentially better off under information sharing than the firm i in the isolated bilateral relationship ij. Also, a firm that reneges on an informal agreement can now be punished by direct reciprocity (j punishes i for i's deviation on ij) *and* indirect reciprocity (j reneges on k after i's deviation on ij, and k then reneges on i following j's deviation on jk). Strategic indirect reciprocity implies that firm i gets punished on the relationship ki, which is of high value to firm i. Even if *a priori* firm i can choose whether to share information with firms j and k independently, the network mechanism imposes an endogenous strategic constraint on these choices that can change the incentives in each relationship.

We formulate a stylized network model that allows us to precisely identify relevant constraints on such *network enforcement mechanisms*. In our model, relationships are long-term and players are self-interested and forward looking. To focus on the network interaction we keep bilateral relationships simple. The only decision facing each partner of a specific bilateral interaction is whether to *maintain* the relationship, or *sever* it. As a result, at time t, player i faces a decision with respect to all relationships in his neighborhood that were maintained by both partners up to $t-1$. On all such relationships, player i can decide to continue the relationship to $t+1$ or sever it permanently. In some sense, the assumption that severance is permanent can be viewed as fixing a particular direct reciprocity mechanism (how players respond directly to a deviation by their partner), allowing us to focus on the role of strategic indirect reciprocity (how players respond in *other* relationships following deviation from a partner).[7] The assumption is not needed to illustrate the possibility of equilibria with network enforcement, but it allows us to draw sharper inferences about the topology of networks on which such mechanism are particularly viable.

In addition to allowing players to choose different actions in each of their bilateral interactions, we assume in our main analysis that play in each bilateral relationship is *private information* to the two parties involved.[8] This assumption captures a constraint that is important in the settings we have in mind, where the nature of bilateral exchange makes the monitoring of relationships difficult for a third party. In the strong triangle network example from Figure 1(b) private monitoring means that player i observes the status of arcs ij and ki, but does not observe the status of arc jk. The status of the unobserved arc does not affect player i's realized payoffs, but when there are strategic network externalities it can be crucial to his expected payoffs from a given strategy. When a player maintains low value relationships because of the possibility of indirect punishment on high value relationships, private information implies that strategies are *conditional* on the unobserved status of other relations. However, players can infer this status from observations on the relationships in their own neighborhood. Interdependencies between even distant relationships in the network are thereby established by strategic connections in the overlapping neighborhoods of individuals.

An equilibrium of the network game requires that individual incentives in local neighborhoods be consistent with the global interdependencies of a network enforcement mechanism, and vice versa. The connection between local and global interactions implies that the network plays a dual role in our analysis. On the one hand, networks encode institutions that support indirect reciprocity via network enforcement mechanisms, and thereby *foster* cooperation in bilateral relationships. On the other hand, the network is an underlying structure

[7]Our model could be translated into an multi-player, infinitely repeated game with private monitoring, in which bilateral relationships are infinitely repeated Prisoner's Dilemmas (IRPDs) if we allowed for severance to be reversible. This framework would increase the complexity of the analysis without yielding significant additional insights. There are papers in the literature on dynamic games with private monitoring where similar irreversibility assumptions are crucial to demonstrate existence of equilibrium (see, e.g., Compte 2002), but it is easily verified that this is not the case in the network game at hand. In particular, the insights from all examples considered in this paper are easily replicated when bilateral relationships are modeled as IRPDs. One can also replicate all the existence results given in the paper to show that infinite cooperation is possible in a network even when it is not incentive compatible in a bilateral, asymmetric IRPD. The relevant sufficient conditions easily follow from the results given in this paper.

[8]Private information in this paper pertains to knowledge about the status of other relationships in a known network. This is in contrast to a few recent papers that consider incomplete knowledge about the network itself, including Galeotti et al., forthcoming, Galeotti (2006), and McBride (2006).

that *constrains* the local and global interaction opportunities of individuals. Equilibrium conditions in the network game clearly identify this duality and relate incentives in localized interactions to the architecture of the global network. In small networks, the effect of private monitoring constraints is limited, but as networks become large (in terms of the distance between nodes) private information makes network enforcement increasingly difficult. The idea that the small-worlds structure of real networks has an institutional justification is found, informally, in a large literature in sociology, management science and elsewhere. But, to our knowledge, our paper is the first to establish formally how network interactions can support the enforcement of indirect reciprocity when there are no exogenous restrictions on choices across relationships, and to relate the small-worlds structure to the fact that closely connected networks allow this to be achieved with a more robust form of decentralized monitoring.[9]

The remainder of the paper is organized as follows. Section 2 presents the model; we begin with the formal model of a bilateral relationship and then embed bilateral relationships in a network. To account for the dynamic structure and informational asymmetries, we analyze the game using perfect Bayesian equilibrium and a belief-free refinement. It is instructive to also consider the game under *public information* as a benchmark, where perfect Bayesian equilibrium is equivalent to the requirement of subgame perfection. Section 3 gives an extended example to highlight some key features of network enforcement, and to contrast the implications of public and private network enforcement. Section 4 gives the general analysis of the network game under public information, and Section 5 presents the full analysis of network enforcement under private information. Section 6 considers the belief-free refinement. Section 7 concludes. Proofs and a number of additional results are given in an Appendix.

[9]Raub and Weesie (1990) is a seminal reference from the sociology literture on networks, which has a similar motivation to our work. The network enforcement mechanisms we study are also related to the community enforcement mechanisms studied in Kandori (1992) in the sense that both approaches provide strategic foundations for indirect reciprocity (see also Ghosh and Ray, 1996, and Deb, 2008, which relax the strong monitoring assumptions in Kandori). However, the environment in our paper is quite different from the random matching environment studied in this literature. We study strategic interactions in a fixed network of long-term bilateral relationships. Indirect reciprocity is important in this type of environment when relationships are asymmetric, such that cooperation in individual relationships is not incentive compatible because the long term benefits are not high enough to induce cooperation from one of the partners.

2 Model

2.1 Bilateral relationships

We first give the formal definition of a bilateral interaction. In a bilateral relationship, two individuals (the partners) can exchange favors over time, $t = 0, ..., \infty$. Individuals maintain relationships to receive a benefits, which are normalized to 1 (without loss of generality) and bestowed on both partners in every period the relationship is maintained. An isolated bilateral relationship is depicted graphically by an *arc* connecting two *nodes* (the partners) as in Figure 1(a). The *direction* in the graphical representation reflects an asymmetry: One partner, j, pays a low cost $\underline{c} \in (0, 1)$ to maintain the relationship in every period, while the other partner, i, pays a higher cost $\bar{c} \in (\underline{c}, 1)$. This implies that the net benefit of a maintained relationship differs across partners. In Figure 1(a), the net benefit of relationship ij is greater for j than for i.[10] Partners choose whether to maintain or sever a relationship simultaneously in each period. A player who unilaterally severs in some period $\tau \geq 0$ obtains benefit without cost in τ, while their partner incurs cost without benefit. If both sever, both incur no cost and receive no benefit. In either case, a severing by either party eliminates all possibility of interaction between i and j for $t > \tau$. This reflects the idea that if an informal agreement is broken, the bilateral relationship as it currently stands ends. Finally, we assume that partners are impatient but not myopic, and discount exponentially with common discount factor $\delta \in (0, 1)$.

An isolated bilateral relationship can be viewed as a simple two-player game. Both partners, i and j, need to choose some time period $\tau \in \{0, 1, ..., \infty\}$ in which to sever the relationship ij. The payoffs if, for example, player i severs in τ and j severs in $\tau' > \tau$ are illustrated in Table 1.

time period	payoff to player i	payoff to player j
$t < \tau$	$1 - \bar{c}$	$1 - \underline{c}$
$t = \tau$	1	$-\underline{c}$
$t > \tau$	0	0

Table 1: payoff table for ij (example)

Note that when $\delta \geq \bar{c}$ both players are willing to maintain the relationship *ad infinitum*

[10] Asymmetry is at the heart of our analysis, but the specific form chosen here is not essential. We could allow for asymmetries in benefits without any change in the results, because only net benefits matter for the analysis. With some additional effort it is also possible to allow for the case $\bar{c} > b$.

as long as they believe their partner will also do so. If $\delta < \underline{c}$ both players will sever the relationship in any case, even though the symmetric outcome in which both do so is strictly Pareto dominated. Finally, there exists a robust range of discount factors, $\Delta := (\underline{c}, \bar{c})$, in which – absent strategic considerations – j would be willing to maintain the relationship with i, but severance is a strictly dominant strategy for i. In this last case, maintenance of the relationship is not an equilibrium outcome precisely because of the asymmetry in the relative value of the relationship.

In this model of a bilateral relationship we have reduced the strategic problem to one of maintaining the relationship (e.g., upholding an information-sharing agreement) or severing (e.g., reneging on the agreement). This model of bilateral relationships is simple enough to make the network analysis tractable and to capture two stylized features of economic and social relationships in many network environments: (1) a trade-off between individual incentives and cooperative outcomes, and (2) an asymmetry in the net benefits from the relationship. In information-sharing relationships, for example, there is often an inherent tension between the short-term competitive advantage that can be obtained by withholding information from a partner firm, and the benefits of cooperative information sharing over the long-run. Moreover, asymmetries arise because firms often specialize in very different types of information. While two firms could adjust the information they share to try and establish a *quid pro quo*, empirical observations suggest that firms in real information sharing networks often do not (Vicente et al., 2008; Sammarra and Biggiero, 2008). To make sense of information sharing under such circumstances, we recognize that bilateral relationships are not isolated but are embedded in networks, and argue that *quid pro quo* can be established by pooling asymmetries across relationships.

2.2 Network enforcement

In our model of a network game, bilateral relationships are embedded in a network of other, similar relations. A network can be represented graphically by a directed graph (or *digraph*) as in the strong triangle network of Figure 1(b). Nodes in the digraph represent players (here players i, j and k), and arcs represent bilateral relationships (here relationships ij, jk and ki). The direction of an arc indicates the asymmetry in each relationship. For example, in the strong triangle network, i incurs a higher cost than j to maintain ij but a lower cost than k for the maintenance of relationship ki. The objective of our analysis is to identify the range of discount factors for which any subset of relations in a given underlying network can be *fully maintained*, i.e. maintained for all $t = 0, ..., \infty$. Note that for $\delta < \bar{c}$ full maintenance of any relationship requires network enforcement because full maintenance of a strategically isolated bilateral relationship is not incentive compatible. We characterize the critical discount factor that determines whether a subset of relations can be fully maintained in an equilibrium in terms of the parameters $\underline{c}$ and $\bar{c}$, and the structural properties of the network architecture. We also determine when the critical discount factor is in Δ, where network enforcement is essential. The analysis is self-contained but requires some notation. We therefore precede a formal definition of the network game with a short example that illustrate the nature of the network analysis.

To illustrate how embedding relationships in a network can effect incentives, we return to the example of a strong triangle network from Figure 1(b). Let $\bar{c} = 0.9$ and $\underline{c} = 0.5$, so that an isolated bilateral relationship (such as ij in Figure 1(a)) can be maintained if and only if $\delta \geq \bar{c} = 0.9$. Now consider the strong triangle network, replicated in Figure 2(a), and suppose for now that actions on all arcs are publicly observed. The strategy profile where all players "maintain all arcs as long as no severance has been observed anywhere in the network, and sever all arcs immediately otherwise" is feasible give this information structure. It is also easily verified that this strategy profile is a subgame perfect equilibrium of the network game (defined formally in Section 2.4) on the strong triangle network as long as $\delta \geq (\bar{c} + \underline{c})/2 = 0.7$. Hence, for $\delta \in [0.7, 0.9)$, any relationship in the strong triangle network can be fully maintained if and only if the whole network is fully maintained.

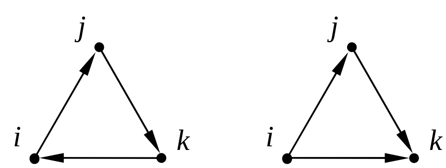


Figure 2: (a) Strong triangle network (left), (b) Weak triangle network (right).

To illustrate why network structure might impose constraints on network enforcement, consider now the weak triangle network in Figure 2(b). This network differs from the strong triangle by the direction of the asymmetry in the relationship between i and k. We say that a node is a *source* with respect to a digraph if it has at least one out-arc but no in-arcs in the network, and that a node is a *sink* when it has at least one in-arc but no out-arcs in the network. Hence, in Figure 2, the strong triangle has no sources or sinks, while node i is a source with respect to the weak triangle and node k is sink. Note that for the source in the weak triangle, node i, maintenance of either relationship is incentive compatible only if $\delta \geq 0.9$, regardless of the strategy followed by other players. This is exactly the incentive condition for maintenance of an isolated bilateral relationship and, under the assumption that players have a common discount factor, network enforcement therefore does extend the opportunity for maintenance in the weak triangle.

Of course, the insight from the weak triangle network generalizes and an absence of sources is therefore a necessary condition for equilibria featuring network enforcement. While this condition looks like a local network property (i.e., one that can be verified by looking only at the neighborhoods of individual nodes), it does have global implications. We say that a consecutive sequence of nodes and arcs is a *cycle* if it begins and ends with the same node, and we can travel from one node to another following the direction of the arcs in the cycle (a formal definition is given in Section 2.3). We get a global implication from the absence of sources condition because all (non-trivial) digraphs that have no sources must have a cycle subdigraph. To see why, suppose that D has no sources and start with an arbitrary node, j_1, that has at least one connection to another node in D (the existence of such a node is meant by non-trivial). Node j_1 has an in-arc, j_2j_1 because j_1 is not a source. Since j_2 is not a source, there must also exist j_3j_2 in D. Likewise there must exist j_4j_3, and so on. However, since the digraph D is finite, eventually we must find an arc $j_{n+1}j_n$ where $j_{n+1} \in \{j_1, ..., j_{n-1}\}$. Hence, we have found at least one cycle in D, and the existence of a cycle is therefore a minimal condition on global network architecture derived by looking only at the incentives of individuals in their own network neighborhoods. The objective of the sequel is to determine exactly what additional local and global restrictions apply to network enforcement, especially when we introduce restrictions on information. These constraints are crucial to a good understanding of network enforcement because, while public information may be a reasonable approximation in small networks such as the strong triangle, extrapolation based on this assumption to arbitrarily large networks places untenable monitoring requirements on individuals.

2.3 Network notation

This section introduces some basic network notation that will be used in later sections. A network is represented formally by a digraph, D, which consists of two finite, non-empty sets: A *set of nodes* $N(D)$ representing players, and a *set of arcs* $A(D)$ representing bilateral relationships.[11] Arcs inherently represent an asymmetric relationship between two nodes, and are graphically represented by an arrow. $ij \in A(D)$ is a generic arc, where $i, j \in N(D)$, and where the order ij indicates that i is the head of the arc and j is the tail (see the digraph in Figure 1(b)). $ij \in A$ is an *out-arc* for i and an *in-arc* for j, and we say that i is *adjacent* to j, and j is adjacent from i (or, simply, that i and j are adjacent in D). The *node*

[11] Where no confusion will result, we write $D = (N(D), A(D)) = (N, A)$. Likewise, for subdigraph D' we use the notation $A' = A(D')$ and $N' = N(D')$.

assignment function $\iota : 2^{A(D)} \to 2^{N(D)}$ assigns to each subset of arcs of digraph D the set of nodes that are adjacent to or from an arc in that set. We denote the set of all possible arc sets connecting nodes $N(D)$ by $\mathcal{A}(D)$. $D' \subset D$ is called a *subdigraph* if $N(D') \subset N(D)$, $A(D') \subset A(D)$ and $\iota(A(D')) \subset N(D')$.

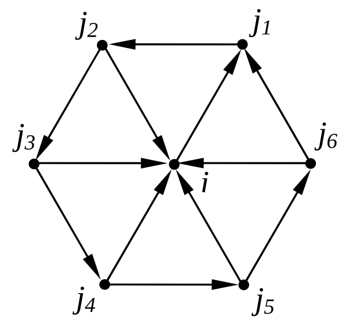


Figure 3: Wheel ($D^W_{6,1}$).

As an example of a network consider the digraph $D^W_{6,1}$ depicted in Figure 3. This digraph is an example of a wheel network. Wheel networks can be viewed as stylized models of many information sharing networks which often have a centralized structure with one node at the center (sometimes called a *sponsor*) connected to many nodes on the outside which are less densely connected to each other (called *peripheral nodes*). We will use wheel networks to illustrate key insights of strategic interactions on networks in later sections, but for now we use the particular network $D^W_{6,1}$ to introduce some important digraph notation. For example, we see that $D^W_{6,1}$ has an interesting subdigraph, $D^{W\prime}_{6,1} = \left(N\left(D^{W\prime}_{6,1}\right), A\left(D^{W\prime}_{6,1}\right)\right) = (\{j_1, ..., j_6\}, \{j_1j_2, ..., j_5j_6, j_6j_1\})$, which is the cycle network connecting nodes on the periphery. Both of these networks satisfy the only *ex-ante* restriction we impose on network structure, given by the following assumption.

Assumption 1 *In a network game on D, (1) $A(D)$ has no self-loops, (2) $A(D)$ has no parallel arcs, and (3) $A(D)$ has no anti-parallel arcs.*

(1) rules out arcs of the form ii, which would be nonsensical in our context. (2) and (3) imply that the bilateral relationship between any nodes i and j is always unique. If nodes represent firms and arcs are information-sharing relationships, (2) and (3) express the idea that all interactions between two firms are summarized in one relationship, and that all relationships in the network exhibit the same asymmetries. This saves us from distinguishing between cases in stating results, which would be cumbersome without leading

to any additional insights. Given this assumption, it is also unambiguous to denote by $\underline{ij}$ either ij or ji, depending on which of these is in the network.

The following notation is used to describe means of traveling through a digraph. Let i and j be (not necessarily distinct) nodes of a digraph D. A finite, alternating and directed sequence

$$i = i_0, i_0j_1, j_1, \ldots, j_{k-1}, j_{k-1}j_k, j_k = j$$

of nodes and arcs, is called a *ij-path* (of length k) if no node is repeated. It is called a *cycle* (of length k) if $i = j$ but no other node is repeated. In digraph $D^W_{6,1}$, the sequence $j_6, j_6j_1, j_1, j_1j_2, j_2$ is a j_6j_2-path (of length 2), while the sequence

$$j_6, j_6j_1, j_1, j_1j_2, j_2, \ldots, j_5, j_5j_6, j_6$$

is a cycle (of length 6).

The *underlying graph* $G(D)$ of digraph D is obtained by replacing the set of arcs of D with a set of undirected edges. Most of the terms defined for digraphs have a natural counterpart for undirected graphs. A node i is said to be connected to a node j in a graph G if there exists an ij-path in G.[12] A graph G is connected if every two of its nodes are connected. A digraph D is *connected* (or weakly connected), if $G(D)$ is connected. The relation 'is connected to' is an equivalence relation on the set of nodes of a graph G, and every subgraph induced by the nodes in a resulting equivalence class is called a connected component of G, or simply a component.

For a connected digraph D, the *distance* $d_D(i,j)$ between two nodes i and j is the minimum of the lengths of the ij-paths of D. So, for instance, in $D^W_{6,1}$, $d_{D^W_{6,1}}(j_1, j_4) = 3$. If there is no ij-path in D, we define $d_D(i,j) = \infty$.[13] $d_D(i, j|\, A')$ is the distance between nodes i and j when travel is restricted to paths that use every arc in A' exactly once (set $d_D(i, j|\, A') = \infty$ when this is impossible). $d_D(i, j|\, \neg A')$ is the distance between nodes i and j when travel is restricted to paths that do not use any arc in A' (set $d_D(i, j|\, \neg A') = \infty$ when this is impossible). In digraph $D^W_{6,1}$, While $d_{D^W_{6,1}}(j_4, j_1) = 2$, $d_D(j_4, j_1|\, \{j_4j_5\}) = d_D(j_4, j_1|\, \neg\{ij_1\}) = 3$. Additionally, note that $d_D(j_4, j_1|\, \neg\{ij_1, j_6j_1\}) = \infty$. The undirected distance functions $d_{G(D)}(i,j)$, $d_{G(D)}(i, j|\, A')$, $d_{G(D)}(i, j|\, \neg A')$ on a (undirected) graph are given in the analogous way by following the shortest undirected distance.

The following notation is used to describe the neighborhood of a node i in a digraph. The *outdegree* $od_D(i)$ is the number of nodes that are adjacent from i in subdigraph D, i.e.

[12] An ij-path in a graph G is defined as for a digraph but without the restriction on direction.

[13] For the statement of results, it is convenient for us to work in the extended integers $\mathbf{N} \cup \{\infty\}$. We follow the convention that $f(\infty)$ denotes $\lim_{x\to\infty} f(x)$ and use this notation only when the limit is well-defined.

$od_D(i) = |\{j \in N(D) - i : ij \in A(D)\}|$. The *indegree* $id_D(i)$ is the number of nodes adjacent to i, i.e., $id_D(i) = |\{j \in N(D) - i : ji \in A(D)\}|$. The *ratio* of outdegree to indegree is denoted $r_D(i) := od_D(i)/id_D(i)$. To ensure that the ratio is well defined, let $r_D(i) = \infty$ if $od_D(i) > id_D(i) = 0$ and $r_D(i) = 0$ if $od_D(i) = id_D(i) = 0$. The η-*neighborhood* of node i in D is $NE_D^\eta(i) = \{lk \in A| \min\{d_{G(D)}(i,k|\{lk\}), d_{G(D)}(i,l|\{lk\})\} \leq \eta\}$. For example, in $D_{6,1}^W$, $od_{D_{6,1}^W}(j_1) = 1$, $od_{D_{6,1}^W}(j_1) = 2$, hence $r_{D_{6,1}^W}(j_1) = 1/2$, while $NE^1_{D_{6,1}^W}(j_1) = \{j_1j_2, ij_1, j_6j_1\}$.

2.4 Network game

We are primarily interested in the network game under private information, but in order encompass public information as a benchmark we introduce the idea of a *radius of information*. Note first that a complete history of the game up to some period t will state, for every arc of the underlying network, whether the corresponding relationship is currently maintained, or when and by whom it was severed. The radius of information, denoted $\rho \in \{1, 2, ...\}$, is the length of paths (in the network) on which a player observes the history of play. For example, when $\rho = 1$ players only know the history of play on adjacent arcs (private information), while $\rho = \infty$ implies public information (where the entire history of play is common knowledge).

The relevant history of play on an arc ij at time t is summarized by $h_{ij}^t = (\mathfrak{s}, \mathfrak{t}, \mathfrak{p})_{ij}^t \in \{\mathcal{M}, \mathcal{S}\} \times \{0, 1, \ldots, t-1\} \times \{\{i\}, \{j\}, \{i,j\}\}$, where $\mathfrak{s} \in \{\mathcal{M}, \mathcal{S}\}$ denotes whether an arc is "maintained" or "severed" at the start of time t, $\mathfrak{t} \in \{0, 1, \ldots, t-1\}$ denotes when the arc was severed (if a severance has occurred; otherwise denote $\mathfrak{t} = \varnothing$), and $\mathfrak{p} \in \{\{i\}, \{j\}, \{i,j\}\}$ denotes which player severed the arc (if a severance has occurred, otherwise denote $\mathfrak{p} = \varnothing$). $h_i^{t,\rho} = \{h_{lk}^t | lk \in NE_D^\rho(i)\}$ is the *i-observable t-history*, i.e., the t-history of all arcs observed by player i. $h^t = \{h_{ij}^t\}_{ij \in A}$ is a complete *history* of the game at time t. Note that the time zero history is a collection of 3-tuples $(\mathcal{M}, \varnothing, \varnothing)_{ij}^0$, which summarizes the underlying network. Capital H's denote the corresponding collections (e.g., $H^t = \left\{\{h_{ij}^t\}_{ij \in A}\right\}$, the set of all possible t-histories on digraph D).

If ρ is small, players cannot distinguish between all histories. To formalize this, let $h^t \sim_{i,\rho} \hat{h}^t \iff h_i^{t,\rho} = \hat{h}_i^{t,\rho}$, where $\sim_{i,\rho}$ is an equivalence relation denoting the observational equivalence of two histories for player i. $\langle h^t \rangle_{i,\rho}$ denotes an equivalence class of t-histories from player i's perspective, i.e., $\hat{h}^t \in \langle h^t \rangle_{i,\rho} \iff \hat{h}^t \sim_{i,\rho} h^t$. $[H^t]_{i,\rho} = \left\{\langle h^t \rangle_{i,\rho}\right\}$ is the collection of all observational equivalence classes of t-histories from player i's perspective. Denote by $\langle h_i^{t,\rho} \rangle$ the set of complete t-histories that are observationally consistent with $h_i^{t,\rho}$, i.e., $\langle h_i^{t,\rho} \rangle = \{\hat{h}^t | \hat{h}_i^{t,\rho} = h_i^{t,\rho}\}$. We omit ρ when it is clear from the context.

We concentrate on equilibria in pure strategies. To define strategies formally, suppose ρ is given. Let $m_i^{t,\rho} : H_i^{t,\rho} \to \mathcal{A}(D)$ give, for any i-observable t-history, the set of arcs that are maintained within i's radius of information at time t; i.e., $m_i^{t,\rho}\left(h_i^{t,\rho}\right) = \left\{ij \in NE_D^{\rho}(i)\,|h_{ij}^t = (M,\varnothing,\varnothing)_{ij}^t\right\}$. Define an action function $s_{i,j}^t : H_i^{t,\rho} \to \{M, S\}$. A strategy of player i should state, for each i-observable t-history, which of any maintained arcs i intends to maintain into the next period. For this, let $s_i^t\left(h_i^{t,\rho}\right) = \left(s_{i,j}^t\left(h_i^{t,\rho}\right)\right)_{\underline{ij} \in m_i^{t,\rho}\left(h_i^{t,\rho}\right)}$. Then a *strategy of player* i is an infinite tuple, $s_i = (s_i^t)_{t \geq 0}$, and a *strategy profile* is a strategy for every player, $s = (s_i)_{i \in N}$. Denote the continuation of player i's strategy s_i after history $h_i^{t,\rho}$ by $s_i|_{h_i^{t,\rho}}$. Capital S's denote the corresponding collections (e.g., S_i is the set of all player i strategies). $\sigma_{ij}^t : S \to H_{ij}^{t+1}$ denotes the status of ij at time $t+1$ given strategy profile s. Analogously, $\sigma^t(s)$ is the status of the underlying network, and $\sigma(s)$ is the *path of play* of the infinite game under strategy s.

The payoffs of a player, i, in a given period, t, is the sum of payoffs across each of i's relationships realized in that period. On in-arcs this is 0 in relationships that were severed by either player prior to t, $(1-\underline{c})$ in relationships maintained in t by both players, $-\underline{c}$ in relationships unilaterally severed by the partner, and 1 on relationships unilaterally severed by i. Payoffs on out-arcs are determined likewise but with cost $\bar{c}$ for maintenance. The total period t payoff is the sum of payoffs from each bilateral relationship in t. Hence, the payoff realized by player i at time t under strategy profile s is:

$$\begin{aligned} \pi_i^t(s) \;=\;& (1-\bar{c})\left(\sum_{ij \in A} I_{(M,\varnothing,\varnothing)}\left(\sigma_{ij}^t(s)\right)\right) + (1-\underline{c})\left(\sum_{ji \in A} I_{(M,\varnothing,\varnothing)}\left(\sigma_{ji}^t(s)\right)\right) \\ &+ \left(\sum_{ij \in A} I_{(S,t,i)}\left(\sigma_{ij}^t(s)\right) + \sum_{ji \in A} I_{(S,t,i)}\left(\sigma_{ji}^t(s)\right)\right) \\ &-\bar{c}\left(\sum_{ij \in A} I_{(S,t,j)}\left(\sigma_{ij}^t(s)\right)\right) - \underline{c}\left(\sum_{ji \in A} I_{(S,t,j)}\left(\sigma_{ji}^t(s)\right)\right), \end{aligned} \tag{1}$$

where I is an indicator function that takes a value one according to which of the states $(M,\varnothing,\varnothing)$, (S,t,i), (S,t,j), $(S,t,\{i,j\})$ is chosen by the partners on a given active arc in period t, under s. The average continuation payoff to player i under strategy profile s at time t is $\pi_i^t(\delta, s) = (1-\delta)\sum_{\tau=t}^{\infty} \delta^{\tau-t}\pi_i^{\tau}(s)$, which exists because $\pi_i^{\tau}(s)$ is bounded and $\delta < 1$.[14]

[14]Note that the realized payoff of player i only depends on the status of arcs in his own network neighborhood. Private information therefore implies that players are able to condition play in the network game on exactly the relationships which are directly payoff relevant.

2.5 Equilibrium

As usual in dynamic games, Nash equilibria of the network game on a digraph D can be supported by non-credible threats off the path of play. To rule out such play we consider a refinement of Nash equilibrium that accounts for the dynamic structure of the game. Our focus is on network games under private information, so that t-histories are only partially observed by any one player. In equilibrium, players form beliefs about what happens in parts of the network they do not observe, and we impose Bayesian consistency on these beliefs.

To define beliefs when t-histories are only partially observed, denote by $\mu_i^t\left(\left\langle h_i^{t,\rho}\right\rangle\right)$ the probability distribution on $\left\langle h_i^{t,\rho}\right\rangle$, which represents player i's beliefs given the observation of t-history $h_i^{t,\rho}$. One component of this is denoted $\mu_i^t\left(h^t|\, h_i^{t,\rho}\right)$ (note that $\mu_i^t\left(h^t|\, h_i^{t,\rho}\right)=0$ if $h_i^{t,\rho}\notin \langle h^{t,\rho}\rangle_{i,\rho}$). $\mu_i^t=\left(\mu_i^t\left(\left\langle h_i^{t,\rho}\right\rangle\right)\right)_{\left\langle h_i^{t,\rho}\right\rangle\in[H^t]_{i,\rho}}$ is a system of period t beliefs of player i (one set of beliefs for every history that player i could have observed at time t); $\mu^t=\left(\mu_i^t\right)_{i\in N}$ is a system of beliefs for each player at time t of the game; and $\mu_i=\left(\mu_i^t\right)_{t\geq 0}$ is a system of beliefs for player i. We call $\mu=\left(\mu_i\right)_{i\in N}=\left(\mu^t\right)_{t\geq 0}$ a *system of beliefs*. Finally, at time t the set of histories is finite and so the expected continuation payoff of player i under strategy s given i-observable t-history $h_i^{t,\rho}$ has been observed is given by an inner product,

$$R_i^t\left(s,\mu_i^t|h_i^{t,\rho}\right)=\mu_i^t\left(\left\langle h_i^{t,\rho}\right\rangle\right)\cdot\pi_i^t\left(\delta,\left(s_j|_{h_j^{t,\rho}}\right)_{j\in N(D)}\right)\ . \tag{2}$$

We say that a system of beliefs μ is *consistent* with a given strategy profile s if it is in the following set

$$\Psi(s)=\left\{\mu|\mu_i^t\left(\sigma^t(\hat{s}_i,s_{-i})|h_i^t\right)=1 \text{ if } \left\langle\sigma^t(\hat{s}_i,s_{-i})\right\rangle=h_i^t,\,\forall t\geq 0,\,\forall\hat{s}_i\in S_i,\,\forall i\in N\right\}\ . \tag{3}$$

This consistency condition is equivalent to the idea that players Bayesian update beliefs regarding t-histories from their beliefs about the strategy profile s at every information set reached with strictly positive probability under s. Given that s is a pure strategy profile, the condition $\mu\in\Psi(s)$ implies (1) if i has observed no deviations from the path of play defined by s, then i should put point mass on the path of play, and (2) if i deviates from s but observes no other deviations from the continuation path prescribed by s, then i should put point mass on the continuation path implied by s, given i's deviation from s. However, if i observes deviations from some player in $N-\{i\}$, then $\mu\in\Psi(s)$ imposes no restrictions on subsequent beliefs because these are all associated with probability zero events under s. We use this notion of consistency to define an appropriate equilibrium concept for the game.

Definition 1 (Perfect Bayesian Equilibrium) *A strategy profile s and system of beliefs μ are a perfect Bayesian equilibrium (PBE) if $\mu\in\Psi(s)$ and*

$$R_i^t\left(s,\mu_i^t|h_i^{t,\rho}\right)\geq R_i^t\left(\hat{s}_i,s_{-i},\mu_i^t|h_i^{t,\rho}\right), \tag{4}$$

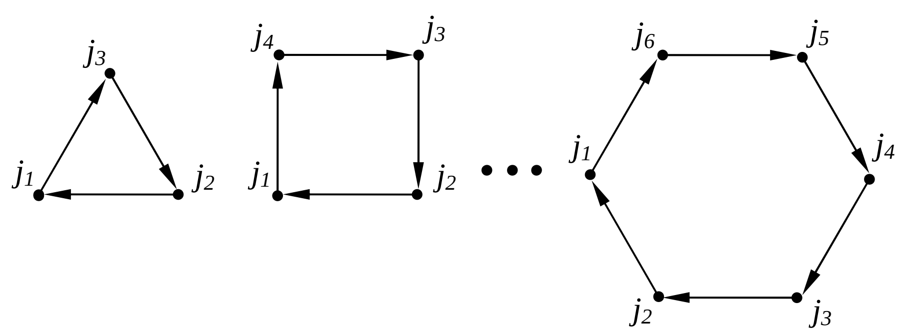
for all $h_i^{t,\rho} \in H_i^{t,\rho}$, $\hat{s}_i \in S_i$, *and* $i \in N(D)$.

Hence, a strategy profile s can be a PBE if and only if *there exists* a system of beliefs consistent with s, such that all players are maximizing expected continuation payoffs at every information set. Given that many of these information sets will be reached with zero probability under s, PBE is the weakest refinement that is consistent with Bayesian rationality and rules out non-credible threats in equilibrium. In the game at hand, PBE is a refinement of subgame perfect equilibrium and the two solution concepts coincide exactly when the network game has public information. We also consider a strong belief-free refinement of PBE in Section 6, to highlight network structures on which network enforcement is particularly robust, but we defer the definition of a belief-free equilibrium to that section.[15]

3 An Example

We first look at network enforcement on a particular subset of digraphs satisfying Assumption 1, namely the class of cycle networks of length n. We denote a typical member of this class $D_n^C = (\{j_1, ..., j_n\}, \{j_1j_2, ..., j_nj_1\})$, and the initial members, D_3^C though D_6^C, are depicted in Figure 4. Note that D_3^C is the strong triangle network from Figures 1(b) and 2(a).

Figure 4: Cycle of length 3 (D_3^C) - Cycle of length 6 (D_6^C).



[15] We also consider sequential equilibrium in Appendix C. Sequential equilibrium refines on PBE by imposing more structure on beliefs, and is refined by belief-free equilibrium. The qualitative results on global network structure under PBE and belief-free equilibrium already go in the same direction, and so sequential equilibrium does not highlight many new insights here. Moreover, while we give some intuitive existence results for network enforcement in a sequential equilibrium, we also demonstrate by example a number of unappealing implications of sequential equilibrium that suggest why this is not a suitable solution concept for the game at hand.

Cycle networks are characterized by two properties: (1) there is only one component, and (2) each node has exactly one in- and one out-arc, i.e., each player in the network has one relationship where they incur cost $\underline{c}$ for maintenance, and one relationship where they incur cost $\bar{c}$. The observation that for $\delta < \bar{c}$ fully maintained digraphs must contain a cycle (see Section 2.2) has an immediate implication for cycle networks: No relationship in a cycle network, D_n^C, can be maintained in any period unless the whole network is fully maintained.[16] For a network game on a cycle network it therefore only remains to show when the whole network can be fully maintained in an equilibrium. This depends on the information structure of the game and we consider public and private information in turn.

3.1 Public network enforcement on cycle networks

To characterize the critical discount factor as of which the whole cycle network D_n^C can be fully maintained in an PBE under public information, we define the worst-punishment strategy of the network game. Let s^{pub} be the strategy profile where all players "maintain all arcs as long as no severance has been observed anywhere in the network, and sever all arcs immediately otherwise". Given this strategy profile, the optimal deviation for any one player is to sever both of their arcs in the network immediately. Hence, s^{pub} is a subgame perfect equilibrium in which a whole cycle network is fully maintained if and only if $\delta \geq (\bar{c}+\underline{c})/2 =: \delta^c_{pub}$. Since s^{pub} specifies the worst punishment, no other strategy profile could achieve full maintenance for $\delta < \delta^c_{pub}$. The critical discount factor as of which full maintenance can be achieved on network D_n^C under public information is simply the average of $\bar{c}$ and $\underline{c}$. If, for example, $\bar{c} = 0.9$ and $\underline{c} = 0.5$, any one relationship in D_n^C can be maintained if and only if $\delta \geq 0.9$, while the whole network can be maintained if and only if $\delta \geq 0.7$. Hence, full maintenance of D_n^C involves network enforcement when $\delta \in [0.5, 0.9)$. This is precisely the range identified in our discussion of the strong triangle network in Section 2.

Note that the critical discount factor δ^c_{pub} does not depend on n, the length of the cycle. However, while the assumption of public information becomes less plausible as the cycle size increases, it also becomes increasingly important for the enforcement mechanisms that utilize

[16]This follows from a simple backward induction argument: Suppose that some relationships in D_n^C are fully maintained, but the whole network is not. Then there exists some time period τ in which the maintained digraph does not contain a cycle, i.e., it has a source. A source will always sever their remaining relationship, but then the source's partner would become a source. Hence, in an equilibrium, the partner should anticipate severance and sever all remaining relationships in $\tau - 1$, and so on. As a result, unless the whole network is fully maintained, all relationships must be severed in period 0. This condition is unique to the class of cycle networks. In any other network, if the whole network is fully maintained it must contain some cycles, and at least one of those could also be maintained without maintenance of any other relationship in the network.

the network. In fact, while we do not impose that players use all the information regarding history of play, in every equilibrium in which a cycle network is fully maintained all players *must* be using all of this information. This again follows from a simple backward induction argument. For $\delta < \bar{c}$ maintenance of any relationship $j_k j_{k+1} \in D_n^C$ must be conditional on maintenance of $j_{k-1}j_k$, otherwise j_k has no incentive to maintain. By the same argument, maintenance of $j_{k-1}j_k$ must be conditional on maintenance of $j_{k-2}j_{k-1}$, and so on. But now suppose that any arc in the network is severed in some period τ, then the partner to that relationship must sever in $\tau + 1$, her partner should anticipate this and do likewise, and so on. It follows that if any relationship is severed in any period, all remaining relationships must be severed immediately in the following period. As a result, s^{pub} is, in fact, the unique equilibrium that can achieve full maintenance when $\delta < \bar{c}$ and information is public. But the assumption that players are be able to monitor the status of every relationship in a network becomes increasingly implausible as networks become large. For densely connected networks, such as the strong triangle, public information may be a reasonable approximation, but as cycles get larger constraints on monitoring limit the possibility of the kind of network enforcement mechanism that s^{pub} represents.

3.2 Private network enforcement on cycle networks

Again, we start by defining the worst-punishment strategy for the network game on a cycle under private information. Let s^{priv} be the strategy profile where all players "maintain all arcs as long as no severance has been observed *in their own network neighborhood*, and sever all arcs immediately otherwise". This strategy profile is similar to s^{pub} but explicitly takes account of the private monitoring constraint. The optimal deviation under private information is, however, quite different from the one under public information. For $\delta \in \Delta$, each player would like to sever out-arcs as soon as possible and maintain in-arcs as long as possible. Since severance by j_k of $j_k j_{k+1}$ is observed only by j_{k+1}, this implies that the optimal deviation from s^{priv} for j_k is to sever $j_k j_{k+1}$, but maintain $j_{k-1}j_k$ for the time being. According to s^{priv}, after $j_k j_{k+1}$ has been severed, severance will spread through the network as j_{k+1}, and then j_{k+2}, and so on, sever their out-arcs. Hence, the optimal deviation for j_k is to sever $j_k j_{k+1}$ in period τ and then sever $j_{k-1}j_k$ in $\tau + (n-2)$, the same period in which j_{k-2} will sever her relationship to j_{k-1}, and therefore the last period before j_{k-1} would anyway sever $j_{k-1}j_k$.[17]

[17]Note that this is not a one-shot deviation. The optimal one-shot deviation would be to sever both arcs immediately, but this is sub-optimal under private information, where the optimal deviation continues for $n-2$ periods. We therefore observe that the counterpart to the one-shot deviation principle does not apply

Given the optimal deviation, we find that the critical discount factor under which full maintenance is incentive compatible for s^{priv} solves $(\bar{c} - \delta^c_{priv}) + (\delta^c_{priv})^{n-2}(\underline{c} - \delta^c_{priv}) = 0$. Clearly, this condition has a solution in Δ, i.e., $\delta^c_{priv} < \bar{c}$. But unlike under public information, δ^c_{priv} is strictly increasing in n, and converges to $\bar{c}$ as n goes to infinity. To give an idea about numerical significance, suppose that $\bar{c} = 0.9$ and $\underline{c} = 0.5$. Then Table 2 gives δ^c_{priv} for various values of n. We see from the example that, while for small n the difference between public and private is not large, as n becomes large the range of discount factors for which network enforcement via s^{priv} is significant diminishes quickly.

n	3	4	5	6	7	8	12	24
$\delta^c_{priv}(n)$	0.73	0.75	0.77	0.79	0.80	0.81	0.84	0.89

Table 2: Critical discount factor as a function of cycle length

Since s^{priv} is the worst punishment strategy available under private information, no other strategy profile could achieve full maintenance for $\delta < \delta^c_{priv}$. To get a characterization of the critical discount factor it therefore remains to show that there exists a system of beliefs under which s^{priv} is a PBE strategy profile. Let μ^{priv} be the belief system in which every player believes that the whole network is fully maintained if they have observed nothing to suggest the contrary, and that every relationship outside of their own neighborhood was severed in the last period otherwise. It is easily verified that $\mu^{priv} \in \Psi(s^{priv})$ and, for $\delta \geq \delta^c_{priv}$, the strategy profile s^{priv} is a mutual best response given μ^{priv} at every information set. Hence, (s^{priv}, μ^{priv}) is a PBE in which the whole cycle network D^C_n is fully maintained if and only if $\delta \geq \delta^c_{priv}(n)$. We therefore have necessary and sufficient conditions for the existence of a PBE with network enforcement, given in terms of the parameters of a bilateral relationship, $\bar{c}$ and $\underline{c}$, and the structural parameter of the network, n. Moreover, while we find that network enforcement does not require public information, we observe that as the cycle length increases network enforcement is relevant on an ever smaller range of discount factors.[18]

in this game. That is not due to the dynamic structure because it is easy to verify that the same observation would hold if every relationship was an infinitely repeated Prisoner's Dilemma. Rather, the usual recursive structure breaks down because of the asymmetric information. Players can exploit private information when deviating from the equilibrium path, and the conditions that are sufficient to rule out one-shot deviations are therefore not sufficient for equilibrium.

[18]It is also worth noting how we have used the common knowledge assumption regarding the network in this argument. We need to assume that players have "knowledge" of the network structure and strategy profile in order to determine the incentive compatibility constraints of individuals in an equilibrium, just as common knowledge is required to determine incentive constraints in a standard two-player game. However, the strategy of any individual player makes no reference to any relationship outside of his neighborhood, and

In the sequel we turn to network enforcement in the general class of networks satisfying Assumption 1. Results are stated generally, but even when it is not mentioned explicitly, it should be understood that $\delta \in \Delta$ in the discussion, i.e., we focus on the range of discount factors where network enforcement is both essential and easily identified.

4 Public Network Enforcement

We start with the case of public information. Under public information the network game is a dynamic game of perfect information. Players are able to condition maintenance of any bilateral relationship in their neighborhood on the status of any other bilateral relationship in the entire network.

Assumption 2 *The radius of information is* $\rho = \infty$.

4.1 Balancing roles as a measure of contribution

Recall that in the class of cycle networks public network enforcement did not depend on the specific structural properties of a network (i.e., the length of a cycle). This is no longer true when we leave the class of cycle networks, because nodes in a digraph D can now have different ratios (of out-arcs to in-arcs), and these are important when considering incentive compatibility constraints. For $\delta \in \Delta$ full maintenance of a given subdigraph $D' \subset D$ requires that players balance the incentive to sever less valuable relationships in D' against their desire to maintain more valuable relationships. If individuals are willing to balance relationships in their neighborhood, the network can act as an institution that pools asymmetries and relaxes overall incentive constraints. The *balancing role* occupied by player i is measured by i's ratio in D'. This ratio can be interpreted as i's contribution to full maintenance of D'. Players with a lower ratio benefit more from maintenance of the network, while players with higher ratios would benefit more from a deviation from maintenance. Hence, the opportunity cost of full maintenance is higher for players who are required to play a greater balancing role to support an equilibrium in which D' is fully maintained.

In Section 3 we found that for the class of cycle networks, full maintenance is only possible when every player makes the same contribution. In general, however, nodes can

the belief system of an individual involves no specific beliefs about what happens on specific relationships outside of his neighborhood. Hence, we do not really require that players know the network in an objective sense, only that they act *as if* they know the network when we determine incentive constraints.

play quite different balancing roles in an equilibrium with network enforcement, and it is the maximal contribution made by any player that determines the incentive compatibility of full maintenance. Theorem 1 establishes this formally by relating equilibria in which a subdigraph D' is fully maintained for $\delta \in \Delta$, to a sufficient statistic of the digraph D', namely the maximum ratio of any node in N'. To state the theorem, we denote by Φ_{Pub} the function

$$\Phi_{Pub}(D) = \frac{r(D)\,\bar{c} + \underline{c}}{[r(D) + 1]}, \tag{5}$$

where $r(D) = \max_{i \in N} r_D(i)$ is the maximum ratio of any node in digraph D.[19] Φ_{Pub} is a weighted average of $\bar{c}$ and $\underline{c}$, and takes values in $[\underline{c}, \bar{c}]$ according to the maximum ratio of digraph D. In the special case of a cycle network, $r(D) = 1$ and we get the incentive constraint from Section 3.1. More generally, Φ_{Pub} determines whether the strategic indirect reciprocity that is required when $\delta \in \Delta$ is incentive compatible for all players in the network. To understand the theorem, it is important to note that Φ_{Pub} is not "monotone": If $D'' \subset D'$ are subdigraphs of D, $\Phi_{Pub}(D'')$ can be greater, less than or equal to $\Phi_{Pub}(D')$. The following theorem gives necessary and sufficient conditions for full maintenance of any subdigraph D' of a network D in terms of the relation between discount factor δ and $\Phi_{Pub}(D')$.

Theorem 1 *Suppose that the network game on a digraph D satisfies Assumptions* 1 *and* 2.

1. *If $\delta' \geq \Phi_{Pub}(D')$, then there exists a PBE, (s, μ), of the network game on D under which $D' \subset D$ is fully maintained for all $\delta \in [\delta', 1)$.*
2. *If subdigraph $D' \subset D$ is fully maintained in a PBE, (s, μ), of the network game on D, then there exists $D'' \supset D'$ such that D'' is fully maintained under s and $\delta \geq \Phi_{Pub}(D'')$.*

Proof. The proof is given in the Appendix. ■

To interpret the content of the theorem, note that parameters $\bar{c}$ and $\underline{c}$ are defined purely in terms of the strategic interaction in an isolated bilateral relationship. Theorem 1 takes these as given and is concerned with two problems of the strategic interaction at the network level. *Problem 1* is to minimize δ such that a given subdigraph D' can be fully maintained in a PBE of the network game on $D \supset D'$. *Problem 2* is to maximize the subdigraph D' that can be fully maintained in a PBE of the network game on $D \supset D'$ for a given δ. Theorem 1 emphasizes the duality between *Problem 1* and *Problem 2* and, by exploiting this duality, identifies a simple network statistic, $r(.)$, that is necessary and sufficient to address

[19]Note that $r_D(.)$ is a function on nodes for given digraph D (see Section 2.3), while $r(.)$ is an operator on the set of all possible digraphs.

both problems. The second condition of the theorem generalizes the insight from the cycle examples (Section 3.1): For a given δ, a subdigraph D' may only be fully maintained in equilibrium if it is contained in a larger subdigraph $D'' \supset D$ that is itself fully maintained in equilibrium. That is the sense in which network interaction can be crucial. The theorem illustrates that it is the contribution that each player would be required to make towards full maintenance of a given subdigraph that alone determines the incentive constraints on public network enforcement. We return to the example of the wheel digraph to illustrate this point.

4.2 Public network enforcement on wheel networks

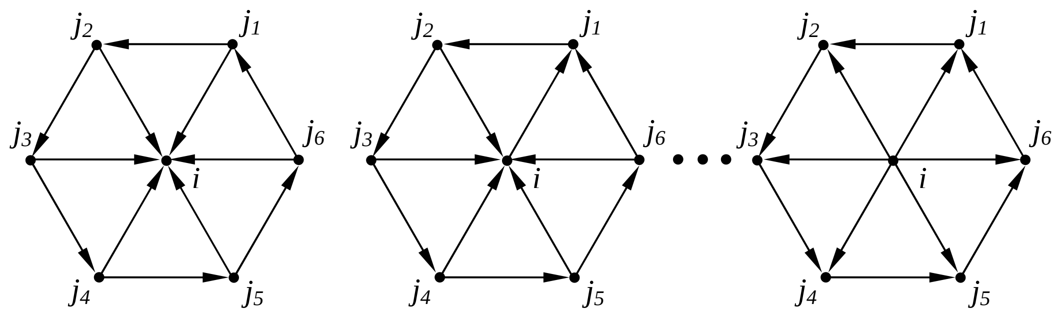


Figure 5: Wheel $D^W_{6,0}$ (far left) – Wheel $D^W_{6,6}$ (far right).

Figure 5 illustrates the first, second and last member of the class of wheel digraphs with periphery cycle of length 6. Consider digraph $D^W_{6,0}$ first. The whole digraph can be fully maintained in a SPE for $\delta \geq (2\bar{c} + \underline{c})/3$ ($= \Phi_{Pub}(D^W_{6,0})$). In an equilibrium with full maintenance of this network, player i makes no contribution. All other nodes in the network are required to play a balancing role. While every player in $N - \{i\}$ has the incentive to sever their relationship to i, they are not able to do so because deviations are publicly observed and, under an appropriate generalization of s^{pub} from Section 3.1, every deviation acts as a public coordination device. However, if full maintenance of $D^W_{6,0}$ is possible in an equilibrium, it follows immediately from Theorem 1 that the subdigraph in which only the cycle connecting the peripheral nodes is fully maintained must also be an equilibrium. In some sense, node i therefore appears to be a free-rider. The payoff of maintaining relationships to i does not compensate players on the periphery for foregoing the one-off severance payoff, and they are willing to do so only because they receive compensation from other players in the network (on each of their respective in-arcs).

Network $D^W_{6,1}$ is obtained from $D^W_{6,0}$ by replacing $j_1 i$ with $i j_1$. This increases the ratio of player i but does not change the maximum ratio in the digraph. Hence, the same range of δ's allows for full maintenance of $D^W_{6,1}$. Likewise $r(D^W_{6,2}) = r(D^W_{6,3}) = r(D^W_{6,4}) = r(D^W_{6,0})$, so that the critical discount factor is not altered. However, going from $D^W_{6,4}$ to $D^W_{6,5}$, player i's ratio increases above $r(D^W_{6,0})$ and the critical discount factor as of which $D^W_{6,5}$ can be fully maintained is $\delta^c = (5\bar{c} + \underline{c})/6 > \Phi_{Pub}(D^W_{6,0})$. Here, it is now the sponsor of the wheel who makes the greatest contribution to full maintenance, but i may still be willing to forgo deviation because of her one remaining in-arc. However, in $D^W_{6,6}$ player i is a source, and so there does not exist any $\delta \in \Delta$ for which $D^W_{6,6}$ can be fully maintained. The subdigraph $D^{W\prime}_{6,6} = D^C_6 \cup \{i\}$ can be fully maintained for $\delta > (\bar{c} + \underline{c})/2$ but full maintenance of the whole network $D^W_{6,6}$ is no longer an equilibrium outcome. While even player i would strictly prefer the outcome where the whole network is maintained to any outcome that is realized in equilibrium, full maintenance is not possible because other players in the network are aware that full maintenance is not incentive compatible for the sponsor.

4.3 Limits on network enforcement

The apparent discontinuity in going from $D^W_{6,5}$, where the whole digraph can be fully maintained for a range of δ's in Δ, to $D^W_{6,6}$, where half of the relationships in the network can not be maintained for any $\delta \in \Delta$, is somewhat misleading. In fact, if we consider the class of all wheel digraphs of the form $D^W_{n,m} = (\{i\} \cup \{j_k\}, \{ij_k | k \leq m\} \cup \{j_k i | m < k \leq n\} \cup \{j_k j_{k \oplus (n) 1} | k \leq n\})$ (where $n > 2$ and $m \leq n$) we have a convergence condition in the following sense: For any $\delta \in [(2\bar{c} + \underline{c})/3, \bar{c})$ there exist an n and an $\bar{m}$ such that for $m < \bar{m}$ the whole digraph can be fully maintained in a network game on $D^W_{n,m}$, and for $m > \bar{m}$ there does not exist a PBE in which the whole digraph is fully maintained in the network game on D^n_m. Corollary 1 establishes this result formally by highlighting that an absence of sources – already identified as a necessary condition for network enforcement in Section 2.2 – is, in fact, the only structural property that is required to get network enforcement on a digraph for some subset of discount factors in Δ.

Corollary 1 *Suppose that the network game on a digraph D satisfies Assumptions* 1 *and* 2*, and that $\delta \in \Delta$.*

1. *If $D' \subset D$ has no sources, then D' can be fully maintained in a SPE, s, of the network game on D for δ sufficiently large.*
2. *If subdigraph $D' \subset D$ is fully maintained in a SPE s of the network game on D, then there exists $D'' \supset D'$ such that D'' is fully maintained under s and has no sources.*

Proof. The proof is given in the Appendix. ■

Theorem 1 and Corollary 1 highlight that balancing alone determines the incentive constraints on public network enforcement. This is a local incentive constraint and while it implies some constraints on global network structure (such as the existence of a cycle in the network), these come from the coordination on public signals rather than from strategic reasoning about the architecture of the network. Empirically and intuitively, the idea that network enforcement should be able to support full maintenace on arbitrarily large networks does not seem plausible. Most networks observed in empirical network studies are found to exhibit a small-worlds structure, with short distances between nodes. On such networks, public information may, in fact, be a reasonable approximation, but Theorem 1 and Corollary 1 suggest that it should also be possible to achieve the same kind of strategic indirect reciprocity observed in small-worlds, on arbitrarily large networks. As the examples of Section 3 illustrate, this often implies that players *must* condition on all information available. But in real networks such information is usually decidedly difficult to come by. We argue, therefore, that the extrapolation to large networks suggested by Theorem 1 and Corollary 1 is flawed because it does not account for monitoring constraints which seem, intuitively, to become increasingly important as the distance between nodes increases. The next section formalizes this intuition.

5 Private Information

We now give the analysis of network enforcement under private information. Under private information the network game is a dynamic game of imperfect information. Players are able to condition maintenance of any bilateral relationship in their neighborhood only on the status of other bilateral relationship in their network neighborhood.

Assumption 3 *The radius of information is* $\rho = 1$.

By restricting the information players can condition on, private information decreases the opportunities for network enforcement. However, decreasing the radius of information, ρ, is *not* an equilibrium refinement. We have demonstrated this already in the example of cycle networks in Section 3, where we observed that the equilibrium strategy profile s^{priv} – which can achieve full maintenance on a cycle network for some $\delta \in \Delta$ under private information – is not an equilibrium of the game under public information. We begin the analysis of private network enforcement with an example to highlight some of the constraints that arise when we require fully decentralized monitoring.

5.1 Gatekeeping, structural holes and transmission

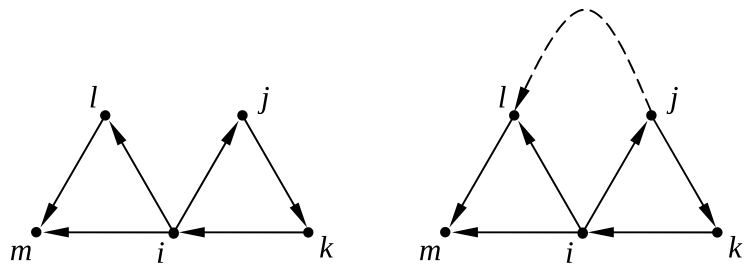


Figure 6: (a) Twin peak network D^{TP} (left), (b) Appended twin peak network D_n^{aTP} (right).

Figure 6(a) illustrates a "twin peak" network, D^{TP}. This network can be fully maintained under public information if and only if $\delta \geq (3\bar{c} + \underline{c})/4$, i.e., for a possibly large subset of discount factors in Δ. However, the twin peak network cannot be fully maintained under private information for $\delta \in \Delta$. Under private information, players j and k find out about the status of arcs in $\{il, im, lm\}$ only if player i reveals this through actions chosen on $\{ij, ik\}$. Hence, conditioning play on $\{ij, ki\}$ on $\{il, im\}$ (and vice versa) can never be part of an equilibrium because optimal deviations by player i will always treat these subdigraphs as separate. Player i therefore partions D^{TP} into two strategically indepedent subdigraphs: $(\{i, j, k\}, \{ij, jk, ki\})$ and $(\{i, l, m\}, \{il, im, lm\})$. The first of these is strong triangle from Sections 2.2 and 3.1, which can be fully maintained in a PBE on a range of discount factors in Δ. The second is D^{wT} from Section 2.2, which we know cannot be maintained for any $\delta < \bar{c}$. Hence, D^{TP} cannot be fully maintained by private network enforcement.

In D^{TP}, player i is said to be a *global gatekeeper*. Global gatekeepers lie on a unique path between two or more parts of a network and, under private information, control the flow of information between these parts of the network. A set of arcs is said to be an *arc-cut set* of a digraph if any path from a node not adjacent to an arc in the set, to a node adjacent to an arc in the set, must pass through a global gatekeeper. Hence, in the twin-peaks network, i is the only global gatekeeper, and $\{il, lm, im\}$ and $\{ij, jk, ki\}$ are the only non-trivial arc-cut sets. As the example indicates, a global gatekeeper will partition a network into its arc-cut sets and these will therefore always be part of strategically independent subdigraphs. A generalization of the notion of global gatekeeping can be used to interpret the additional incentive constraints that arise when we require private network enforcement.

Definition 2 (Local Gatekeeper) [20] *For digraph $D = (N, A)$, suppose that $\{\underline{ij}, \underline{ik}\} \subset N$.*

[20]Local gatekeepers are closely related to the concept of local bridges from the social network literature

Node i is a local gatekeeper with respect to nodes j and k if i is on the (strictly) shortest, undirected path between j and k, i.e., if $d_{G(D)}(j,k|\neg\{\underline{ij},\underline{ik}\}) =: n^i_D(j,k) > 2$. In that case, $n^i_D(j,k)$ is called the order of i as a local gatekeeper between j and k.[21]

To illustrate the relation between local and global gatekeepers, consider the appended twin-peaks network of Figure 6(b), D^{aTP}_n, obtained from D^{TP} by appending a path of length n connecting nodes j and l. Then i is a local gatekeeper of order n between j and l, and becomes a global gatekeeper as n goes to infinity. Our main result in this section demonstrates that it is not only player i's ratio that determines the incentive compatibility of private network enforcement, but also player i's position as a local gatekeeper between in- and out-arcs in i's neighborhood. This is because, for $\delta \in \Delta$, every maintained relationship in network D must be conditioned on the maintenance of other relationships in D. Suppose now that i severs arc $ij \in A(D)$ and that this represents a deviation from the path of play. By Assumption 1, player j is not in a position to punish i for the deviation directly. Moreover, under private information, i's deviation is not observed by any player other than j. Enforcing maintenance of ij therefore requires that j be willing and able to sever other relationships to communicate i's deviation to nodes that are in a position to punish i. We call this *transmission*, and how long transmission takes is vital for network enforcement under private information. The importance of gatekeeping has no counterpart when deviations from equilibrium are publicly observed precisely because transmission is not relevant. But when information is private, players can exploit the fact that the spread of severance through the network is constrained by the distance of paths along which severance must travel, and this means that the distance between nodes becomes important for private network enforcement.

As a result, private information has important implications for the role that nodes occupy in the network. While under public information a player's contribution to full maintenance of a network can be measured by their balancing role alone, under private information nodes can also perform a vital information transmission role to counteract the incentives other players have to exploit local gatekeeping positions. This is particularly relevant when the network has (or, rather, would otherwise have) "structural holes", i.e., when there are parts of the network which are sparsely connected or connected only by very long paths. Even nodes which make little contribution to network enforcement in terms of their balancing role can still be crucial for network enforcement if they occupy such structural holes, and transmit information between otherwise disparate parts of the network. Hence, what appear to be free-riders in the network under public information, can actually play a vital role in

(see, e.g., Granovetter 1973).

[21] A local gatekeeper i is a global gatekeeper if there exists a partition of $NE_D(i) = NE^1 \cup NE^2$ such that $j \in NE^1$ and $k \in NE^2$ implies $n^i_D(j,k) = \infty$.

the decentralized monitoring that is essential for private network enforcement. We next illustrate this point with an example.

5.2 Private network enforcement on wheel networks

The transmission role can be highlighted by comparing the cycle networks from Section 3 to the wheel networks analyzed under public information in Section 4. Note that each node in a cycle network of length n, D_n^C, is a local gatekeeper of order $n_D^{j_k}(j_{k-1}, j_{k+1}) = n-1$, and recall that under private information the critical discount factor as of which D_n^C can be fully maintained in a PBE is a strictly increasing function of n. The critical δ as of which a typical wheel network in which the sponsor has ratio 0, $D_{n,0}^W$, can be fully maintained solves $(\bar{c}-\delta)+\delta(\bar{c}-\delta)+\delta^2(\underline{c}-\delta)=0$. While the cycle subdigraph $D_n^C \subset D_{n,0}^W$ must be maintained in order for $D_{n,0}^W$ to be fully maintained, we observe that the critical discount factor for full maintenance of the wheel is independent of the structural parameter n (the length of the periphery cycle). This is because each node on the periphery of $D_{n,0}^W$ is a local gatekeeper of order 2, and these orders do not depend on n. Hence, both the ratio and gatekeeping order are constant in the class $D_{n,0}^W$ as we vary n. For $\bar{c}=0.9$ and $\underline{c}=0.5$, Figure 7 compares how the critical discount factor depends on n for cycle networks, D_n^C and wheel networks, $D_{n,0}^W$.

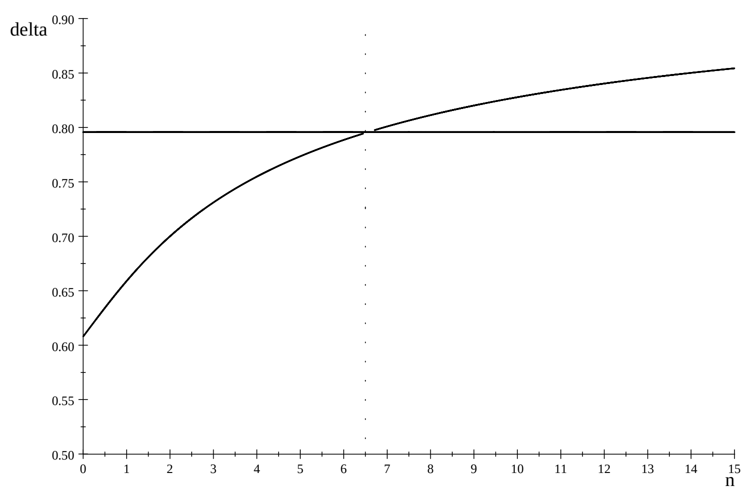


Figure 7: Critical δ as a function of n.

For $n \leq 5$ we observe that there is an open subset of Δ for which $D_{n,0}^W$ cannot be fully maintained, but D_n^C (which is a subdigraph of $D_{n,0}^W$) can be fully maintained in a PBE. But for $n \geq 6$ there is an open subset of Δ for which $D_{n,0}^W$ can be fully maintained in a PBE, but the subdigraph D_n^C cannot. The latter has no counterpart when $\rho = \infty$. But, under private information, despite the fact that node i makes no contribution to full maintenance in terms of a balancing role in $D_{6,0}^W$, node i does play a critical role in network enforcement. By reducing the local gatekeeping order of nodes on the periphery, node i allows information and punishment to spread more quickly through the wheel network. Such networks can be viewed as a stylized model of many information-sharing networks, which often appear to have this centralized structure. Moreover, it is common to find that the sponsor in such networks benefits disproportionately from the network interactions (as suggested here by the direction of the asymmetries in $D_{n,0}^W$). Under public information we are led to believe that the sponsor is free-riding because peripheral nodes can not coordinate severance. However, the analysis of private information highlights that the sponsor may play a vital role in private network enforcement.

5.3 Equilibrium analysis of private network enforcement

We now generalize the insights from the twin peak, cycle and wheel examples. Recall from Section 3 that under private information optimal deviations from a worst-punishment strategy such as s^{priv} do not generally involve a one-shot deviation. As a result, it is generally not possible to solve for a critical discount factor explicitly, but we can characterize the critical discount factor as of which any subset of relations in a digraph D can be fully maintained in terms of an implicit condition that highlights both the balancing and transmission roles observed in the wheel network example. To this end, let $N^+(D)$ denote the set of nodes in a digraph D which have at least one out-arc in D (i.e., $N^+(D) = \{i \in N(D) | od_D(i) > 0\}$), and for each $i \in N^+(D)$ let

$$\Upsilon_D^i \left(\alpha^i, \delta\right) = (\bar{c} - \delta) \sum_{ij \in NE_D(i)} \delta^{\alpha_{ij}} + (\underline{c} - \delta) \sum_{ji \in NE_D(i)} \delta^{\beta_{ji}(\alpha^i)}, \tag{6}$$

where $\alpha^i = (\alpha_{ij})_{ij \in NE_D(i)}$, $\alpha_{ij} \in \mathbf{N} \cup \{\infty\}$ for all $ij \in NE_D(i)$, and

$$\beta_{ji}(\alpha^i) = \min_{ik \in NE_D(i)} \left[\alpha_{ik} + d_D^{ik}\left(i, j | \neg\{ik\}\right) - 1\right] . \tag{7}$$

As a function of δ, $\Upsilon_D^i(\alpha^i, \delta)$ in (6) is defined in terms of parameters from the strategic interaction in a bilateral relationship, i.e., $\bar{c}$ and $\underline{c}$, properties of the network D, and an auxiliary variable α. Given δ, $(\bar{c} - \delta)$ is the net payoff to player i of severing an out-arc when his partner would otherwise have maintained that relationship. $(\underline{c} - \delta)$ is the net payoff to player i of severing an in-arc when his partner would otherwise have maintained that relationship. $\Upsilon_D^i(\alpha^i, \delta)$ can therefore be interpreted as a weighted average of the net benefit of maintaining in- and out arcs, with the number of terms determined by $od_D(i)$ and $id_D(i)$, and the weight on each term determined by α^i and (β_{ij}). The importance of transmission and local gatekeeping is identified by observing that (β_{ij}) in (7) is a function of the distance between out-arcs and in-arcs in player i's neighborhood.

The incentive constraints on full maintenance in a PBE are given by

$$\Phi_{PBE}(i, D, \delta) = \max_{\alpha^i} \Upsilon^i_D \left(\alpha^i, \delta\right) \,, \tag{8}$$

which can be interpreted as a weighted average of net benefits from severance (on in- and out-arcs) along an optimal severance program, given that all nodes in the network follow a specific worst punishment strategy that is very closely related to s^{priv} (see Appendix A for details).[22]

Theorem 2 *Suppose that the network game on a digraph D satisfies Assumptions* 1 *and* 3.

1. *If $\Phi_{PBE}(D', i, \delta') = 0$ for each $i \in N^+(D')$, then there exists a PBE (s, μ) of the network game on D under which D' is fully maintained for all $\delta \in [\delta', 1]$.*
2. *If subdigraph $D' \subset D$ is fully maintained in a PBE (s, μ) of the network game on D, then there exists $D'' \supset D'$ such that D'' is fully maintained under (s, μ) and $\Phi(D'', i, \delta) = 0$ for each $i \in N^+(D'')$.*

Proof. The proof is given in the Appendix. ■

Theorem 2 re-emphasizes the duality between *Problem 1* and *Problem 2* observed in Section 4. Here both problems are solved by looking at a summary statistic, Φ_{PBE}, that is defined only in terms of the parameters $\bar{c}$ and $\underline{c}$, and structural properties of the network architecture. In particular, Φ_{PBE} highlights that the number of in- and out-arcs, as well as the distance between these, determine incentive compatibility constraints on private network enforcement. Heuristically, as the ratio of a player increases, the range of discount factors under which full maintenance is possible decreases. This is related to our findings under public information. The new aspect here is that the opportunity for network enforcement also diminishes as the distances between in- and out-arcs increase. As a result, we find that structural holes – places in the network where connections are not dense – can be a severe impediment to network enforcement. Hence, the importance of the transmission role for players that bridge such structural holes.

[22]Note that the auxiliary variable α is no longer a part of Φ_{PBE}, which can be viewed as a constrained maximization problem where the relevant constraints have already been included via β's dependence on α.

There is a large literature on the implications of "structural holes" in social and economic networks (see, e.g. Burt, 1992 and 2000). In the study of business organizations, for example, structural holes are often viewed as places in a network where the flow of information is restricted by a lack of connectivity. This is hypothesized to reduce efficiency, suggesting that organizations should take efforts to close structural holes. However, it is also often assumed that local gatekeepers acquire benefits from their position bridging structural holes, and therefore resist organizational restructuring. In the context of the network game on a digraph D, Theorem 2 highlights that structural holes can indeed reduce efficiency because they reduce the opportunity for network enforcement mechanisms to counteract double coincidence of wants constraints. However, the equilibrium analysis also illustrates that this does not necessarily imply benefits for local gatekeepers in the network. In an equilibrium, partners of a local gatekeeper fully anticipate the gatekeeper's informational advantage and adjust their behavior accordingly. Hence, the local gatekeeper is never able to exploit the informational advantage of his position in the network.

Finally, it remains to show when $\Phi_{PBE}(D', i, \delta') = 0$ has a solution in Δ, where network enforcement is essential. In a manner analogous to Section 4, this is given by a simple limiting result if we interpret the relevant space correctly. First, as under public information, we need to realize that a source is the limit of a player with increasing ratio. In addition, under private information, we have to perform a similar exercise with respect to path lengths. In particular, we need to realize that global gatekeepers are the appropriately defined limit of local gatekeepers (as illustrated in the twin-peaks network). With these interpretations, the following corollary to Theorem 2 is immediate and gives a possibility result for private network enforcement.

Corollary 2 *Suppose that the network game on a digraph D satisfies Assumptions* 1 *and* 3*, and that $\delta \in \Delta$.*

1. *If the digraph $\left(\iota\left(\hat{A}\right), \hat{A}\right)$ has no sources for every arc-cut set $\hat{A} \subset A'$ of subdigraph $D' \subset D$, then D' can be fully maintained in a PBE of the network game on D for δ sufficiently large.*
2. *If subdigraph $D' \subset D$ is fully maintained in a PBE (s, μ) of the network game on D, then there exists $D'' \supset D'$ such that D'' is fully maintained under (s, μ), and $\left(\iota\left(\hat{A}\right), \hat{A}\right)$ has no sources for every arc-cut set $\hat{A} \subset A''$.*

Proof. The proof is given in the Appendix. ■

5.4 The value of information and public monitoring institutions

The analysis of public and private network enforcement allows us to highlight the value of public monitoring institutions in a network environment. Assuming for the moment that cooperative equilibria can be selected, suppose that players in a network had the opportunity to pay some amount C to fund a public monitoring institution. How much would they be willing to pay? For any discount factor $\delta \in \Delta$ and network D, Theorems 1 and 2 give us the maximal subdigraph of D that can be maintained if, respectively, players do fund such an institution, compared to the maximal fully maintained subdigraph if they do not. Hence, for given values of $\underline{c}$ and $\bar{c}$, their willinginess to pay, C, will be a function of the network D and the discount factor δ. We know already that public information increases the opportunity for network enforcement and hence $C(D, \delta)$ is non-negative. Moreover, Theorems 1 and 2 allow us to reduce the problem of calculating $C(D, \delta)$ to a straightforward computational exercise. For example, if $\bar{c} = 0.9$ and $\underline{c} = 0.5$, then for a cycle network of length $n = 7$ we get

$$C\left(D_7^C, \delta\right) = \begin{cases} 0 & \text{if } \delta \in (0, 0.7] \cup [0.8, 1) \\ 0.6 & \text{if } \delta \in (0.7, 0.8) \end{cases}, \tag{9}$$

because for $\delta \leq 0.7$ neither public nor private network enforcement can fully maintain the network, for $\delta \geq 0.8$ private network enforcement fully maintains the cycle, but for $\delta \in (0.7, 0.8)$ public information would allow players to maintain the whole network while private network enforcement fails. Hence, players may be willing to pay up to 60% of the benefit from a maintained relationship for a public monitoring institution that supports network enforcement in this network.

Institutions that appear to play a public monitoring role are observed in a number of network environments. The system of international trade relations is an example of a network environment in which trade is organized, essentially, via bilateral relationships, and in which in many dimensions the conduct of each partner is difficult for outsiders to the trade relationship to monitor. It is also difficult to demonstrate that each bilateral trade agreement really is mutually beneficial for the agents involved. In fact, one explicit justification given for multilateral agreements is that there is a sense that with some indirect reciprocity, free-trade is beneficial for all even if individual trade agreements are not necessarily beneficial for both parties. In this context, at least one function of the WTO is to monitor and report publicly on trade behavior (WTO, 2009). Of course, the WTO also plays other roles, e.g., by enforcing a multilateral agreement it helps to establish the right norm (free-trade) from the number that are available. But if we view international trade as a network of bilateral agreements, it is clear that the WTO also has an essential role to play in monitoring and coordinating responses to a deviation. Public monitoring institutions are also observed in collusion networks. An example is provided by the collusion case brought against the US airlines industry in the 1980's. Airlines colluded on prices by funding a public price board and then positing and implicitly trading forward prices on this board. One difficulty plaintiffs encountered in bringing a case against the airlines industry was that, in the absence of a theoretical framework in which to formulate and analyze indirect reciprocity, they had to establish the mutual benefit of transactions on a bilateral level. Our model gives a framework for thinking about what indirect reciprocity would mean in such a context, and enables us to relate indirect reciprocity to incentives of individuals, (potentially observable) network structure and the monitoring institutions available.

To summarize the results of this section, we find that the analysis of PBE of the network game under private information is useful to demonstrate and delimit the possibilities of indirect reciprocity on a given network. In particular, we find that the distance between nodes in the network is important for incentive compatibility of network enforcement. As a result, we are able to make sense of the role certain individuals play by virtue of a central network position, and to assess the value of public monitoring institutions in network environments.

6 Belief-free Network Enforcement

In this section we want to consider a refinement of PBE that is useful for an alternative interpretation of the model, by which networks are formed partly under the anticipation of strategic indirect reciprocity. This alternative interpretation is closer to some of the existing literature in economics, which is concerned with incentives on network formation, and is also suggested by a number of applied literatures. For example, in the strategic management literature it is now in vogue to view network formation as a guided process, i.e., to view networks themselves, not just the firms that constitute a network, as the outcome of entrepreneurial design. This perspective is partly motivated by the striking regularities observed in network data, such as the small-worlds property, and by the observation of network institutions such as indirect reciprocity. In this context, a closely related paper is Haag and Lagunoff (2006), who focus on a planner's problem in the design of real-world neighborhoods when there are strategic network interactions between inhabitants. Their modeling framework is close to ours and the tractability restrictions they impose in their equilibrium concept lead to conclusions that have the same flavor as the results derived in this section.

The preceding analysis already suggests a relation between strategic indirect reciprocity and the small-worlds structures, but the dependence of perfect Bayesian equilibrium on particular belief systems is problematic if we want to view the network game as the second-stage in a network formation process. Under this interpretation of the model, the network is formed in some explicit network formation process between players in stage 0, with the understanding that the network game outlined in Section 2 will then be played on the resulting digraph. Anticipation of indirect reciprocity can influence the incentives in the formation process, but it does not seem reasonable that these institutions should then depend on particular belief systems. We introduce the notion of a belief-free equilibrium to address these concerns and identify network structures that allow for a more robust form of private network enforcement. Our main finding is a very stark small-worlds prediction.

6.1 Belief-free equilibrium

Recall that a strategy profile s is a PBE if there exists a system of beliefs $\mu \in \Psi(s)$ such that s is a mutual best response at every information set given the beliefs μ (see Section 2.5). In a *belief-free equilibrium* (BFE) we require that s be a mutual best response at every information set *for all* beliefs consistent with s (i.e., all beliefs μ in $\Psi(s)$).[23] This is a strong refinement that has the satisfying property that the equilibria we construct no longer depend on any particular system of beliefs. As a result, the equilibrium notion is better suited to an interpretation of the model in which the network is first formed under the anticipation of indirect reciprocity opportunities. In fact, the belief-free equilibrium concept is motivated, in part, by the literature on mechanism design (see, e.g., Bergemann and Morris 2007). [24]

To illustrate the concept of a BFE, consider the strong triangle network from Figure 4 in Section 3 again, and suppose that players follow the strategy profile s^{priv} from that section. It is easily verified that the continuation of s^{priv} after observing a severance is a best response regardless of which beliefs player any player in D_3^C entertains in $\Psi(s^{priv})$. Hence, whenever s^{priv} is a PBE it is, in fact, a belief-free equilibrium (BFE). However, consider going from D_3^C to D_4^C. The strategy profile s^{priv} cannot be a BFE of a cycle network of length 4 because we can identify two different belief systems consistent with s^{priv} in which the best response by

[23]Formally, a strategy s is a belief-free equilibrium (BFE) of a network game on D with radius of information ρ if for every $\mu \in \Psi(s)$

$$R_i^t\left(s, \mu_i^t | h_i^{t,\rho}\right) \geq R_i^t\left(\hat{s}_i, s_{-i}, \mu_i^t | h_i^{t,\rho}\right), \tag{10}$$

for all $h_i^{t,\rho} \in H_i^{t,\rho}$, $\hat{s}_i \in S_i$, $i \in N(D)$.

[24]The notion of belief-free equilibrium is also used in the literature on repeated games with imperfect private monitoring (see, e.g., Horner and Olszewski, 2006, and Ely et al. 2005). For a survey of the use of belief-free equilibrium concepts in game theory see Olszewski (2007).

a player in D_4^C following a deviation by one of their partners differ.[25] In fact, for $\delta < \bar{c}$ there does not exist a BFE on any cycle network except the strong triangle. For $\bar{c} = 0.9$, $\underline{c} = 0.5$ the whole network D_3^C can be fully maintained in a BFE if and only if $\delta \geq 0.73$, but for $n > 3$ any relationship in D_n^C can be maintained if and only if $\delta \geq 0.9$. This is in contrast to our findings under PBE where the range of discount factors for which network enforcement is incentive compatible decreases quickly as n increases, but this range vanishes only as n goes to infinity. The reason there are no BFE in large cycles is an absence of "connectivity": In contrast to D_3^C there are players in D_n^C who are not connected to each other by a bilateral relationship when $n > 3$. This suggests that belief-free equilibria require that networks be tightly connected. The appropriate formalization of this idea is given by a notion of *triadic closure*.

Definition 3 (Triadic closure) *A digraph $D = (N, A)$ satisfies triadic closure if $\{lk, \underline{kj}\} \subset A$ implies $\underline{lj} \in A$.*

The terminology of triadic closure is borrowed from social network theory (see, e.g., Grannovetter, 1973). In sociology it expresses the idea that if two people share a common acquaintance it is likely that they will also know each other.[26] We show in Theorem 3 (below) that network enforcement is only possible in a belief-free equilibrium if two players with a common partner also have a bilateral relationship. In some sense, this provides a justification for triadic closure that – to our knowledge – has not yet been established formally. When individuals share acquaintances, they have access to very robust enforcement mechanisms that overcome private monitoring constraints because deviations in any one relationship can quickly be communicated to the whole community.

[25]Suppose, for example, that j_2 severs the relationship $j_1 j_2$. Player j_1 could believe that relationships $j_2 j_3$ and $j_3 j_4$ are severed or maintained. If both arcs are currently maintained, according to the continuation strategy $j_2 j_3$ will be severed next period and $j_3 j_4$ in the period after that. The arc $j_4 j_1$ would therefore be severed three periods into the future and, if $\delta \in \Delta$, j_1's best response given these beliefs is to sever $j_4 j_1$ in two periods. If, however, $j_2 j_3$ and $j_3 j_4$ have already been severed, player j_1 should sever the relationship to j_4 immediately. There are at least two systems of beliefs that are consistent with s^{priv}. The latter can support s^{priv} as a PBE for $\delta \in \Delta$. But s^{priv} cannot be a BFE since there exist j_1-observable histories under which player j_1's best responses are not the same under two distinct belief systems in $\Psi(s)$.

[26]Kossinets and Watts (2006) provide empirical evidence in support of triadic closure using e-mail data.

6.2 Belief-free network enforcement and triadic closure

To characterize belief-free network enforcement, we again define an appropriate weighting function, $\Phi_{BFE}(D)$, as follows,

$$\Phi_{BFE}(D) = \frac{-(r(D) - \underline{c}) + \sqrt{(r(D) - \underline{c})^2 + 4r(D)\bar{c}}}{2}. \tag{11}$$

Φ_{BFE} is a weighted average of $\bar{c}$ and $\underline{c}$ taking values in $[\underline{c}, \bar{c}]$ according to $r(D)$. While Φ_{Pub} was derived by considering the payoff from severing a set of arcs A that would otherwise be maintained in one period, Φ_{BFE} is derived by considering the payoff from severing all out-arcs in A in one period, and then severing all in-arcs in A one period thereafter.

Theorem 3 *Suppose that the network game on a digraph D satisfies Assumptions* 1 *and* 3.

1. *If subdigraph $D' \subset D$ satisfies triadic closure and $\delta' \geq \Phi_{BFE}(D')$, then there exists a belief-free equilibrium s of the network game on D in which D' is fully maintained for all $\delta \in [\delta', 1)$.*
2. *If $D' \subset D$ is fully maintained in a belief-free equilibrium s of the network game on D, then there exists $D'' \supset D'$ such that D'' is fully maintained under s, D'' satisfies triadic closure and $\delta' \geq \Phi_{BFE}(D'')$.*

Proof. The proof is given in the Appendix. ■

The duality between *Problem 1* and *Problem 2* also underlies Theorem 3. As under public information, the maximal ratio, $r(.)$, is a sufficient statistic. If D' satisfies triadic closure, coordination with respect to severance is enough to induce maintenance without the need to form beliefs about where deviations originated. On the other hand, if a network does not satisfy triadic closure, players who observe a deviation must form beliefs about who else in the network has already observed a deviation, and best responses depend on these beliefs. As in private and public network enforcement, belief-free network enforcement requires that players in the fully maintained network play a balancing role, and the incentive compatibility constraint given by Φ_{BFE} in Theorem 3 indicate that the balancing role may again be quite different across nodes. The information transmission role does not appear in Φ_{BFE}, but is in the restriction to networks satisfying triadic closure. In a BFE all players must be in a position to directly transmit information regarding a deviation from equilibrium to all players in the network who occupy a balancing role with respect to them, or who play a balancing role with respect to someone who plays a balancing role with respect to them, and so on. Transmission is then similar to a public signal except that there is a one period delay and transmission is via many private signals as opposed to a single public signal. Hence, on very densely connected networks, private network enforcement can operate in a similar way to public network enforcement, but while increasing the distance between nodes has no effect on public network enforcement, it rules out belief-free network enforcement altogether.[27]

As a corollary to Theorem 3, triadic closure and an absence of sources are both necessary and sufficient for full maintenance in a BFE for some $\delta \in \Delta$.

Corollary 3 *Suppose that the network game on a digraph D satisfies Assumptions* 1 *and* 3*, and that $\delta \in \Delta$.*

1. *If $D' \subset D$ has no sources and satisfies triadic closure, then there exists a belief-free equilibrium s of the network game on D in which D' is fully maintained for δ sufficiently large.*
2. *If subdigraph $D' \subset D$ is fully maintained in a belief-free equilibrium s of the network game on D, then there exists $D'' \supset D'$ such that D'' is fully maintained under s, has no sources and satisfies triadic closure.*

Proof. The proof is given in the Appendix. ■

An important class of networks that satisfy triadic closure is the class of tournaments. A digraph D satisfying Assumption 1 is called a *tournament* if there is an arc connecting any two nodes in the network (i.e., $i, j \in N(D)$ implies $\underline{ij} \in A(D)$). While triadic closure is a local network property, in Appendix B we give a complete global characterization of networks satisfying triadic closure that is closely related to tournaments, and which we therefore call a quasi-tournament network structure. Basically, a digraph is a quasi-tournaments if it can be partitioned into strongly-connected tournaments, and these parts satisfy a global version of triadic closure (see Appendix B for details).

Tournaments (and quasi-tournaments) are common in economic environments. One example is the network of interconnections between large financial trading institutions, transacting in financial products such as commercial paper and credit default swaps.[28] The structure of interactions in this network is dense, with bilateral flows of transactions between

[27]In the comparison between wheel and cycle digraphs in Section 5, we observed that sinks perform an important task: The sink in a typical wheel network $D^W_{n,0}$ bridges the "structural hole" that otherwise exists between nodes that are far apart in the network. However, sinks in a digraph satisfying triadic closure do not play the same role. In fact, the subdigraph in which all arcs connected to a sink are severed still satisfies triadic closure. Since the ratio of each node in the remaining subdigraph must be lower, existence of BFE in which the whole subdigraph is fully maintained implies existence of a BFE when relationships to any sink are severed.

[28]See Economides (1993) for a simple description of financial networks.

firms that are only partially observed by outsiders, and which often imply highly asymmetric net obligations. Although the transactions are organized via formal markets, it has long been recognized that trust and informal enforcement between trading partners are crucial to the functioning of these markets (see, e.g. Mayer, 2008; or Allen and Babus, 2008, who also argue that tight networks may be important for monitoring in financial networks).[29] Theorem 3 demonstrates that tournament networks can serve as an institution to assist the market in pooling risk and asymmetries in a particularly robust sense. However, Theorem 3 also illustrates a sense in which belief-free network enforcement is fragile. The only strategy that constitutes a BFE on a tournament is "maintain as long as all observed arcs are maintained, and sever all remaining arcs immediately otherwise." While this provides a powerful inducement for others to maintain, it also highlights the potential for swift and crippling severance of cooperative ties if the prescribed path of play is ever left. Tournament networks support enforcement mechanisms that pool asymmetries even with very limited information, but do so by requiring large-scale coordination in punishments.

7 Conclusion

This paper analyzes enforcement mechanisms which foster cooperative behavior between individuals in a network. We show that strategic network externalities can be crucial to achieve efficient outcomes in a network of *a priori* independent relationships, and study constraints on such network effects coming from the structure of the underlying network and the monitoring capabilities across relationships. Intuitively, network effects arise when individuals understand that the way they behave in their own relationships influences the way their partners behave towards others in the network. If individuals believe that what goes around *can* come around, concerns about contagion become a powerful inducement to cooperate. Moreover, if individuals cooperate in some relationships only because of contagion concerns, they in turn become vehicles for further contagion if the incentive to cooperate is ever removed. As a result, social institutions can arise on the network and establish correlations between *a priori* independent bilateral relationships. Our model allows for a sharp characterization of such correlations in terms of equilibrium conditions of the dynamic network game. Specifically, we find that the network occupies a dual role. On the one hand, network institutions that utilize strategic interdependencies between network relations foster cooperation in long-term relationships. Hence, behavior observed in individual relationships

[29]Baker et al. (2004) also observe that in many R&D information-sharing networks the central core of companies that share the most information tends to be very tight, like a quasi-tournament, though they do not use that terminology.

that is difficult to rationalize when these are viewed as isolated entities makes sense when we account for the embeddedness in a network. On the other hand, network structure also imposes constraints on network enforcement mechanisms, and this leads to a clear connection between the incentives of individuals and the global network architecture.

Restrictions on monitoring are crucial to a good understanding of the interconnection between network structure and individual incentives. On networks exhibiting a small-worlds structure we find that network enforcement can be effective, even when monitoring is fully decentralized. However, under private information strategic indirect reciprocity is constrained by the inability of individuals to monitor that informal agreements in distant relationships are being upheld. As a result, we find that network enforcement is only effective on networks exhibiting a small-worlds structure. The conclusion that strategic indirect reciprocity requires small-worlds is further underscored by looking at belief-free network enforcement. Here we find that networks must exhibit a form of triadic closure, which implies a very stark small-worlds prediction. While the perfect Bayesian and belief-free analysis are motivated in part by different interpretations of the model, they lead to qualitatively similar conclusions on global network structure. In particular, the conclusions we reach for private and belief-free network enforcement are much closer than our findings under public information. In all three cases we observe the incentive constraints that arise due to a balancing role that individuals must play in network enforcement. However, unlike under public information, the analysis of private information also highlights the information transmission role in network enforcement, and this leads to the small-worlds structure. As a result, our model suggests an institutional rationale for this pervasive structure of social and economic interactions.

A Proofs

A.1 Proofs for public information

Proof of Theorem 1. 1. We first construct a strategy profile and beliefs $(\hat{s}, \hat{\mu})$, and show that these are a PBE if the conditions given in Theorem 1 are satisfied. Let every player $i \in N$ play strategy $\hat{s}_i(D')$ with respect to D: sever every arc $\underline{ik} \in A - A'$ in period 0, and play the following strategy with respect to the remaining arcs in D': "maintain every adjacent arc in D' until a severing is observed anywhere in D', and then sever on each active adjacent arc in D'"; formally, for each $i \in N$, and arc $\underline{ij}$ in player i's set of active adjacent arcs in D', denoted $m_i^{t,\infty}\left(h_i^{t,\infty}\right) \cap A'$, in period t,

$$\hat{s}_{i,j}^{t}\left(h_i^{t,\infty}\right)(D') = \begin{cases} M & \text{if } m_i^{t,\infty}\left(h_i^{t,\infty}\right)\cap A' = A' \\ S & \begin{array}{c}\text{if } \exists\ lm \in A', \text{ such that } lm \notin m_i^{t,\infty}\left(h_i^{t,\infty}\right)\cap A' \\ \text{and } \underline{ij} \in m_i^{t,\infty}\left(h_i^{t,\infty}\right)\cap A',\end{array} \end{cases} \tag{12}$$

with $\hat{s}_{i,j}^{0}\left(h_i^{0,\infty}\right)(D') = S$ for all $\underline{ik} \in A - A'$. With $\rho = \infty$ beliefs $\hat{\mu}$ are trivial–$\hat{\mu}\left(h^t\right) = 1$, where h^t is the realized path of play on D.

To show that $(\hat{s}, \hat{\mu})$ is a PBE, we need to show that (1) $\hat{\mu}$ is consistent (i.e., $\hat{\mu} \in \Psi$), (2) the severance response specified by $\hat{s}$ is optimal after every history, and (3) players are willing to maintain all of their arcs in D' if the incentive compatibility condition, $\delta \geq \Phi_{Pub}(D')$ for all $i \in N'$, is satisfied, and all opponents maintain all arcs in D'. First, the consistency of $\hat{\mu}$ is trivial. Secondly, the severance response is optimal after every history. Severance of an arc in D' serves as a public coordination device, which makes simultaneous severance of all remaining arcs in D' a mutual best response for all $i \in N'$. This is true after every history. Further, it is obvious that the period-0 severances are mutual best responses for all players.

Thirdly, we look at the incentive compatibility condition, $\delta \geq \Phi_{Pub}(D')$ for all $i \in N$. Notice that $\Phi_{Pub}(D')$ emerges from the optimal deviation from a full maintenance equilibrium that i could follow, given that his opponents play $\hat{s}_{-i}$. It is clear that the optimal deviation involves complete and simultaneous severance of all active arcs, since severance of any arc triggers severance of all remaining arcs in the subsequent period, under $\hat{s}$. Hence $\delta \geq \Phi_{Pub}(D')$ is derived from the following equation,

$$od_{D'}(i)\frac{1-\bar{c}}{1-\delta} + id_{D'}(i)\frac{1-\underline{c}}{1-\delta} \geq od_{D'}(i) + id_{D'}(i), \tag{13}$$

which formalizes the incentive compatibility constraint comparing the discounted payoff from a full maintenance path of play against the payoff from immediate severance of all of i's arcs in D'. Hence if $\delta \geq \Phi_{Pub}(D')$ then each $i \in N$ is willing to maintain all arcs in D' in a full maintenance equilibrium, since even the optimal deviation brings a net loss in payoff. This concludes the proof that the stated conditions are sufficient for the existence of a PBE on a digraph D under which D' is fully maintained.

2. Suppose that (s, μ) is an equilibrium of the network game on D, where we denote the set of fully maintained arcs in D under (s, μ) by

$$\hat{A} = \left\{ij \in A | \sigma_{ij}^{t}(s) = (M, \varnothing, \varnothing)\ \forall\ t \geq 0\right\}. \tag{14}$$

Namely, $\hat{D} = \left(\iota\left(\hat{A}\right), \hat{A}\right)$ is the largest subdigraph that is maintained under (s, μ). Under $\rho = \infty$ beliefs μ are trivial. Clearly, if D' is to be fully maintained under (s, μ) then it must be the case that $D' \subset \hat{D}$, and $\hat{D}$ must satisfy any necessary conditions for the existence of a full maintenance equilibrium. For every $ij \in A - \hat{A}$ let τ_{ij} be the time period in which ij is severed, which is given in $\sigma_{ij}^{\infty}(s)$, and clearly must be finite. Let $\tau = \max_{ij \in A - \hat{A}} \tau_{ij}$, which exists because D is finite.

Now we show that $\delta \geq \Phi_{Pub}(D')$ for all $i \in N'$ is necessary for a full maintenance equilibrium to exist on $\hat{D}$. Let us denote by $s_{\tau}\hat{s}\left(\hat{D}\right)$ the concatenation of strategy s up until period τ with the strategy $\hat{s}$ from part 1. of this proof for all subsequent time periods. Notice that this is a slight abuse of notation; by construction all arcs on $A - \hat{A}$ will have been severed as of period τ, and hence the period-0 severance aspect of $\hat{s}$ is superfluous. Notice that Φ_{Pub} is constructed under the assumption that players play the actions dictated by $\hat{s}\left(\hat{D}\right)$ with respect to the arcs in $\hat{A}$. We now proceed to argue that $\hat{s}$ prescribes exactly the most severe possible deviation response, and hence if $\delta \geq \Phi_{Pub}\left(\hat{D}\right)$ under $s_{\tau}\hat{s}\left(\hat{D}\right)$, then it must be the case that $\delta \geq \Phi_{Pub}\left(\hat{D}\right)$ under $s_{\tau}s = s$.

Notice that any strategy $s_{\tau}s' \neq s_{\tau}\hat{s}$ will (weakly) increase the right-hand-side of the equation

$$od_{D'}(i)\frac{1-\bar{c}}{1-\delta} + id_{D'}(i)\frac{1-\underline{c}}{1-\delta} \geq od_{D'}(i) + id_{D'}(i), \tag{15}$$

because $s_{\tau}\hat{s}$ specifies the fastest response on any given arc, to any deviation by i, and hence any alternative deviation response implies that payoffs are received in some of i's arcs for some additional period(s). This shows that if $\delta \geq \Phi_{Pub}\left(\hat{D}\right)$ under $s_{\tau}\hat{s}\left(\hat{D}\right)$, then it must be the case that $\delta \geq \Phi_{Pub}\left(\hat{D}\right)$ for any alternative strategy profile $s_{\tau}s' \neq s_{\tau}\hat{s}$, and in particular for strategy profile s. ■

Proof of Corollary 1. 1. Suppose there does not exist $i \in N'$, such that i is a source with respect to D'. Then let all players $i \in N'$ play the equilibrium strategy $\hat{s}$ in the proof to Theorem 1, namely in equation 12 above. Then as shown in Theorem 1, the critical discount factor necessary to support full maintenance solves $\delta = \Phi_{Pub}(D')$. Since $r(D')$ is finite, we have the result because $\Delta \neq \emptyset$.

2. If digraph $\hat{D}$ is fully maintained under a PBE, then rearranging the payoff condition shown to be necessary for full maintenance in Theorem 1,

$$\Phi\left(\hat{D}\right)=\delta=\frac{r\left(\hat{D}\right)\bar{c}+\underline{c}}{\left[r\left(\hat{D}\right)+1\right]}, \tag{16}$$

we obtain

$$r\left(\hat{D}\right)\leq\frac{\delta-\underline{c}}{\bar{c}-\delta}, \tag{17}$$

which clearly cannot hold if there is a source in $\hat{D}=D'$ or, if there exists a source in D', in any $\hat{D}=D''\supset D'$, since $r_{\tilde{D}}(i)=\infty$ for a source with respect to a digraph $\tilde{D}$. ■

A.2 Proofs for private information

The following preliminary result proves two properties of Φ_{PBE}, which are useful for the proof of Theorem 2: that for all non-sinks Φ_{PBE} exists (is finite) for all $\delta\in[0,1)$, and there exists a unique δ which separates the range of δ under which $\Phi_{PBE}>0$ and $\Phi_{PBE}=0$.

Lemma 1 *Suppose digraph D satisfies Assumption 1. Then*

(i) Φ_{PBE} *exists for all* $i\in N^{+}(D)$ *and* $\delta\in[0,1)$,

(ii) There exists a unique $\delta^{c}\in[\underline{c},\bar{c}]$ *such that* $\forall\ i\in N^{+}(D)$

$$\delta\in[0,\delta^{c})\implies\Phi_{PBE}(i,D,\delta)>0$$
$$\delta\in[\delta^{c},1)\implies\Phi_{PBE}(i,D,\delta)=0.$$

Proof. (i) We consider three mutually exclusive and collectively exhaustive cases in the interval $[0,1)$. First, for all $\delta\in[\bar{c},1)$, it is clear that each $i\in N^{+}(D)$ optimally sets $\alpha^{i}=(\infty)$, and hence $\Phi_{PBE}=0$. Second, for all $\delta\in[0,\underline{c}]$, it is clear that each $i\in N^{+}(D)$ optimally sets $\alpha^{i}=(0)$, and hence 0 ¡ $\Phi_{PBE}<\infty$. Finally, suppose $\delta\in(\underline{c},\bar{c})$. We show that $\Phi_{PBE}=\max_{\alpha^{i}}\Upsilon_{D}^{i}(\alpha^{i},\delta)$ is bounded. Notice that

$$\begin{aligned}\left|\max_{\alpha^{i}}\Upsilon_{D}^{i}\left(\alpha^{i},\delta\right)\right| &= \left|\max_{\alpha^{i}}\left\{(\bar{c}-\delta)\sum_{ij\in NE_{D}(i)}\delta^{\alpha_{ij}}+(\underline{c}-\delta)\sum_{ji\in NE_{D}(i)}\delta^{\beta_{ji}\left(\alpha^{i}\right)}\right\}\right| \quad (18)\\ &\leq \max_{\alpha^{i}}|\bar{c}-\delta|\sum_{ij\in NE_{D}(i)}\delta^{\alpha_{ij}}+\max_{\alpha^{i}}|\underline{c}-\delta|\sum_{ji\in NE_{D}(i)}\delta^{\beta_{ji}\left(\alpha^{i}\right)} \quad (19)\\ &\leq od_{D}(i)\,|\bar{c}-\delta|+id_{D}(i)\,|\underline{c}-\delta|<\infty. \quad (20)\end{aligned}$$

Since by convention $\delta^{\infty} = 0$, it follows that $\max_{\alpha^i} \Upsilon^i_D\left(\alpha^i, \delta\right)$ is well-defined, and hence Φ_{PBE} exists for all $i \in N^+(D)$ and $\delta \in [0, 1)$.

(ii) First, we show that for all $i \in N^+(D)$, equilibrium strategies involve conditioning play in any arc-cut set $\hat{A} \subset D$ only on arcs in $\hat{A}$. If i's arcs are all in one arc-cut set, then this claim is trivial. If player i has arcs in any two arc-cut sets $\hat{A}_1, \hat{A}_2$, then it must be the case that i is a global gatekeeper, and hence the only paths in D between nodes in $\hat{A}_1$ and nodes in $\hat{A}_2$ must travel through i. With $\rho = \infty$, i can never improve its payoff by playing strategies that condition play in one arc-cut set on the path of play in another arc-cut set. Any enforcement actions that i carries out in an arc-cut set must be carried out on i's arc within that arc-cut set alone, because severing arcs in any other arc-cut set cannot reach the arc-cut set in question, due exactly to the structure of the digraph. Hence it follows that the result must be shown arc-cut set by arc-cut set for each $i \in N^+(D)$. Notice that for it to be the case that $\Phi_{PBE}(i, D, \delta) > 0$, it is only necessary that $\Phi_{PBE}(i, \left(\iota\left(\hat{A}\right), \hat{A}\right), \delta) > 0$ for some arc-cut set $\hat{A} \subset A$, and hence we focus on Φ_{PBE} on an arc-cut set by arc-cut set basis.

We consider two mutually exlusive and collectively exhaustive cases, to prove the existence of a unique $\delta^c \in [\underline{c}, \bar{c}]$, such that the stated conditions hold for all $i \in N^+(D)$. First, suppose that i is a source with respect to D. Then $\delta^c = \bar{c}$, since

$$\Upsilon^i_D\left(\alpha^i, \delta\right) = \max_{\alpha^i} (\bar{c} - \delta) \sum_{ij \in NE_D(i)} \delta^{\alpha_{ij}} > 0 \tag{21}$$

for all $\delta < \bar{c}$, since $\alpha^i < \infty$ will clearly be chosen at a maximum. Hence $\Phi_{PBE}(i, D, \delta) > 0$ for all $\delta \in [0, \bar{c})$ and $\Phi_{PBE}(i, D, \delta) = 0$ for all $\delta \in [\bar{c}, 1)$, if $i \in N^+(D)$ is a source with respect to $\left(\iota\left(\hat{A}\right), \hat{A}\right)$.

Second, suppose that $i \in N^+(D)$ is not a source with respect to any arc-cut set $\hat{A} \subset A$, on subdigraph $\hat{D} = \left(\iota\left(\hat{A}\right), \hat{A}\right)$. We show that there exists a unique $\delta^c \in [\underline{c}, \bar{c}]$ such that $\delta \in [0, \delta^c) \implies \Phi_{PBE}(i, \hat{D}, \delta) > 0$ and $\delta \in [\delta^c, 1) \implies \Phi_{PBE}(i, \hat{D}, \delta) = 0$. First note that $\Phi_{PBE}(i, \hat{D}, \delta)$ is a maximum of polynomials and is therefore continuous. For $\delta \geq \bar{c}$, $\Phi_{PBE} = 0$ because every term in $\Upsilon^i_{\hat{D}}$ is non-positive and hence at the maximum $\alpha^i = (\infty)$. For $\delta < \underline{c}$, $\Phi_{PBE} > 0$ because every term in $\Upsilon^i_{\hat{D}}$ is non-negative and at least one term is strictly positive. Hence by the Intermediate Value Theorem $\exists\ \bar{c} \in \Delta$, such that $\Phi_{PBE} > 0$ for all $\delta < \bar{c}$. We need to show that if $\Phi_{PBE}(i, \hat{D}, \delta') > 0$ there does not exist $\delta'' \in (\underline{\delta}, \delta')$ such that $\Phi_{PBE}(i, \hat{D}, \delta'') = 0$. To show this, we show that Φ_{PBE} is non-decreasing in δ.

Now consider any $\delta \in \Delta$. Let $\hat{\alpha}^i = \arg\max_{\alpha^i} \Upsilon^i_{\hat{D}}\left(\alpha^i, \delta\right)$. Rewrite $\Upsilon^i_{\hat{D}}$ as follows:

$$\Upsilon^i_{\hat{D}}\left(\alpha^i,\delta\right) = (\bar{c}-\delta)\sum_{ij\in NE_{\hat{D}}(i)}\delta^{\alpha_{ij}} + (\underline{c}-\delta)\sum_{ji\in NE_{\hat{D}}(i)}\delta^{\beta_{ji}\left(\alpha^i\right)} \tag{22}$$

$$= l(\bar{c}-\delta) + \sum_{k\geq 1}\left[(\bar{c}-\delta)\sum_{i_k=1}^{n_k}\delta^{\alpha_{i_k}} + (\underline{c}-\delta)\sum_{j_k=1}^{m_k}\delta^{\beta_{j_k}}\right] \tag{23}$$

$$= l(\bar{c}-\delta) + \sum_{k\geq 1}\delta^{\alpha^k}\left[(\bar{c}-\delta)\sum_{i_k=1}^{n_k}\delta^{\alpha_{i_k}-\alpha^k} + (\underline{c}-\delta)\sum_{j_k=1}^{m_k}\delta^{\beta_{j_k}-\alpha^k}\right], \tag{24}$$

where $l>0$, $\alpha^k := \max_{\alpha_{i_k}} \leq \min \beta_{j_k}$ for all $k \geq 1$, and $\alpha_{1_{k+1}} > \beta_{m_k}$. A re-arrangement of terms and change of variables like this is always possible. ij terms in $\Upsilon^i_{\hat{D}}$ are strictly positive for $\delta \in \Delta$ while ji terms are strictly negative. Hence, for every α exponent, there must exist a strictly greater β exponent or $\hat{\alpha}^i$ does not maximize $\Upsilon^i_{\hat{D}}$ because the α exponent can be reduced without constraint (and therefore should be reduced). Note also that $\sum_{k\geq 1}\left[(\bar{c}-\delta)\sum_{i_k=1}^{n_k}\delta^{\alpha_{i_k}} + (\underline{c}-\delta)\sum_{j_k=1}^{m_k}\delta^{\beta_{j_k}}\right] < 0$ because it too can be moved one period earlier without constraint (and therefore should be moved one period earlier unless it is negative). Then, differentiating $\Upsilon^i_{\hat{D}}$ with respect to δ,

$$\begin{aligned}\frac{\partial \Upsilon^i_{\hat{D}}}{\partial\delta}\left(\hat{\alpha}^i,\delta\right) \quad = \quad & -l + \sum_{k\geq 1}\left(\alpha^k\delta^{\alpha^k-1}\left[(\bar{c}-\delta)\sum_{i_k=1}^{n_k}\delta^{\alpha_{i_k}-\alpha^k} + (\underline{c}-\delta)\sum_{j_k=1}^{m_k}\delta^{\beta_{j_k}-\alpha^k}\right]\right. \\ & +\delta^{\alpha^k}\left[-\sum_{i_k=1}^{n_k}\delta^{\alpha_{i_k}-\alpha^k} + (\bar{c}-\delta)\sum_{i_k=1}^{n_k}\left(\alpha_{i_k}-\alpha^k\right)\delta^{\alpha_{i_k}-\alpha^k-1}\right. \\ & \left.\left.-\sum_{j_k=1}^{m_k}\delta^{\beta_{j_k}-\alpha^k} + (\underline{c}-\delta)\sum_{j_k=1}^{m_k}\left(\beta_{j_k}-\alpha^k\right)\delta^{\beta_{j_k}-\alpha^k-1}\right]\right). \end{aligned}\tag{25}$$

Since $\alpha_{i_k} < \alpha^k < \beta_{j_k}$ and $\underline{c} < \delta < \bar{c}$, every term in the derivative is negative. Hence, for fixed $\hat{\alpha}^i$, $\Upsilon^i_{\hat{D}}$ is decreasing in δ. Hence $\Phi_{PBE}(i,\hat{D},\delta') > 0 \Longrightarrow \Phi_{PBE}(i,\hat{D},\delta) > 0\ \forall\ \delta < \delta'$. Hence, since there exists a unique $\delta^c \in [\underline{c},\bar{c}]$ such that $\delta \in [0,\delta^c) \implies \Phi_{PBE}(i,\hat{D},\delta) > 0$ and $\delta \in [\delta^c, 1) \implies \Phi_{PBE}(i,\hat{D},\delta) = 0$. ■

Proof of Theorem 2. 1. We first construct a strategy profile and beliefs $(\hat{s},\hat{\mu})$, and show that these are a PBE if the stated conditions are satisfied for all $i \in N$. Let every player $i \in N$ play strategy $\hat{s}_i(D')$ with respect to D: sever every arc $\underline{ik} \in A - A'$ in period 0, and play the following strategy with respect to the remaining arcs in D': "maintain every adjacent arc in D' until a severing is observed anywhere in D', and then sever on each active adjacent arc in D'"; formally, for each arc $\underline{ik}$ in player i's set of active adjacent arcs in D', denoted $m_i^{t,1}\left(h_i^t\right)$, in period t,

$$\hat{s}_{i,k}^{t}\left(h_i^{t,1}\right)(D') = \begin{cases} M & \text{if } m_i^{t,1}\left(h_i^{t,1}\right) = NE_{D'}^{1}(i) \\ \\ S & \text{if } \exists\ lm \in NE_{D'}^{1}(i)\text{, s.t. } lm \notin m_i^{t,1}\left(h_i^{t,1}\right)\text{, and } \underline{ik} \in m_i^{t,1}\left(h_i^{t,1}\right), \end{cases} \tag{26}$$

with $\hat{s}_{i,k}^{0}\left(h_i^{0,1}\right)(D') = S$ for all $\underline{ik} \in A - A'$. Then we specify beliefs $\hat{\mu}$ as follows: if a player $i \in \iota(D')$ observes a severing on any of his adjacent arcs in D', $\hat{\mu}\left(\hat{h}^t\right) = 1$, where $\hat{h}^t$ is the period-t history under which the i-observable history occured, along with every other non-i-observable arc having been severed, which was not severed previously.

To show that $(\hat{s}, \hat{\mu})$ is a PBE, we need to show that (1) $\hat{\mu}$ is consistent (i.e., $\hat{\mu} \in \Psi$), (2) the severance response specified by $\hat{s}$ is optimal after every observable t-history, and (3) players are willing to maintain all of their arcs in D' if the incentive compatibility condition, $\Phi_{PBE}(i, D', \delta) = 0$ for all $i \in N^+(D')$, is satisfied, and all opponents maintain all arcs in D'. First, $\hat{\mu}$ are clearly consistent, as beliefs exactly follow the conditions in Ψ along the equilibrium path of play, while after a severance of some arc in D' beliefs about unobserved arcs can be arbitrarily specified, since beliefs are unconstrained off the path of play under PBE. Secondly, the severance response is optimal after every observable t-history. Notice that Ψ dictates that players believe that opponents will play in accordance with the continuation strategy dictated by $\hat{s}$ off the equilibrium path of play, in response to the believed structure of the maintained digraph. Hence if $\hat{\mu}_i^t$ describes the beliefs of player i after observation of a deviation in time period t, that all unobserved arcs have been severed, then continuation play under $\hat{s}$ dictates that all of i's remaining arcs will be severed in period $t+1$, and hence it is optimal for i to sever all remaining arcs in period $t+1$. This holds after any deviation, and therefore after any observable t-history following a deviation. It is obvious that the period-0 severances are mutual best responses for all players.

Thirdly, we look at the incentive compatibility condition, $\Phi_{PBE}(i, D', \delta) = 0$ for all $i \in N^+(D')$. First, notice that $\Phi_{PBE}(i, D', \delta)$ emerges from the optimal deviation from a full maintenance equilibrium that i could follow, given that his opponents play $\hat{s}_{-i}$. Against this strategy and for $\delta \in \Delta$ player i wants to sever out-arcs in D' as soon as possible and defer severance on in-arcs in D' as long as possible, subject to the constraint that $\underline{ik}$ must always be severed by i before it is severed by k; these conditions emerge directly from the incentives with respect to each arc when $\delta \in \Delta$. For any $i \in N^+(D')$, $\max_{\alpha^i} \Upsilon_{D'}^{i}(\alpha^i, \delta)$ describes the net gain in expected discounted payoff to player i from following an optimal deviation program, over the payoff $od_{\hat{A}}(j)(1-\bar{c}) + id_{\hat{A}}(j)(1-\underline{c})$ derived from full maintenance, which is implicitly defined by the max operator over all possible deviation programs. If i is a sink with respect to D', then full maintenance is always an optimal response to full maintenance by one's opponents for $\delta \in \Delta$, by incentive compatibility. Given that opponents play $\hat{s}_{-i}$, in-arcs are severed one period prior to being reached as severance is transmitted along pathways emanating from the out-arcs. Recall that we showed in Lemma 1 that Φ_{PBE} has a solution for all $\delta \in [0,1)$ and there exists a unique δ^c as of which $\Phi_{PBE}(i, D', \delta) > 0$ for all $\delta < \delta^c$ for any player $i \in N^+(D')$. Hence if $\Phi_{PBE} = 0$ player $i \in N^+(D')$ is willing to maintain all arcs D', since even the optimal deviation program brings a net gain of zero. This concludes the proof that the stated conditions are sufficient for the existence of a PBE on a digraph D under which D' is fully maintained.

2. Suppose that (s, μ) is an equilibrium of the network game on D, where we denote the set of fully maintained arcs in D under (s, μ) by

$$\hat{A} = \left\{ ij \in A | \sigma_{ij}^t(s) = (M, \varnothing, \varnothing) \;\; \forall \, t \geq 0 \right\}. \tag{27}$$

Namely, $\hat{D} = \left(\iota\left(\hat{A}\right), \hat{A}\right)$ is the largest subdigraph that is maintained under (s, μ). We fix consistent beliefs μ, the only possible beliefs in an equilibrium, under which each player i believes that the path of play dictated by s is followed with probability 1. Clearly, if D' is to be fully maintained under (s, μ) then it must be the case that $D' \subset \hat{D}$, and $\hat{D}$ must satisfy any necessary conditions for the existence of a full maintenance equilibrium. For every $ij \in A - \hat{A}$ let τ_{ij} be the time period in which ij is severed, which is given in $\sigma_{ij}^{\infty}(s)$, and clearly must be finite. Let $\tau = \max_{ij \in A - \hat{A}} \tau_{ij}$, which exists because D is finite.

Now let us show that $\Phi_{PBE}(i, \hat{D}, \delta') = 0$ for all $i \in N^+\left(\hat{D}\right)$ is necessary for a full maintenance equilibrium to exist on $\hat{D}$. Let us denote by $s_\tau \hat{s}\left(\hat{D}\right)$ the concatenation of strategy s up until period τ with the strategy $\hat{s}$ from part 1. of this proof for all subsequent time periods. Notice that this is a slight abuse of notation; by construction all arcs on $A - \hat{A}$ will have been severed as of period τ, and hence the period-0 severance aspect of $\hat{s}$ is superfluous. Notice that Φ_{PBE} is constructed under the assumption that players play the actions dictated by $\hat{s}\left(\hat{D}\right)$ with respect to the arcs in $\hat{A}$. We now proceed to argue that $\hat{s}$ prescribes exactly the most severe possible deviation response, and hence if $\Phi_{PBE}\left(i, \hat{D}, \delta'\right) > 0$ for some $i \in N^+\left(\hat{D}\right)$ under $s_\tau \hat{s}\left(\hat{D}\right)$, then it must be the case that $\Phi_{PBE}\left(i, \hat{D}, \delta'\right) > 0$ for i under $s_\tau s = s$. Finally, notice that we do not need to consider nodes that are sinks with respect to $\hat{D}$, since for $\delta \in \Delta$ incentive compatibility always dictates that full maintenance is always a best response to full maintenance by one's opponents.

Notice that any strategy $s_\tau s' \neq s_\tau \hat{s}$ will (weakly) increase the value of $\Upsilon^i_{\hat{D}}(\alpha^i, \delta)$ for all $i \in N^+\left(\hat{D}\right)$. Notice that $s_\tau \hat{s}$ specifies the fastest response on any given arc, to any deviation by i. In other words, any alterative strategy profile $s_\tau s'$ implies a deviation response that (weakly) increases the value of $\beta_{ij}(\alpha^i)$ for some $ij \in NE_{\hat{D}}(i)$, and hence (weakly) increases the value of $\Upsilon^i_{\hat{D}}(\alpha^i, \delta)$, which (weakly) increases the value of $\Phi_{PBE}\left(i, \hat{D}, \delta'\right)$. This shows that if $\Phi_{PBE}\left(i, \hat{D}, \delta'\right) > 0$ under $s_\tau \hat{s}\left(\hat{D}\right)$, then it must be the case that $\Phi_{PBE}\left(i, \hat{D}, \delta'\right) > 0$ for any alternative strategy profile $s_\tau s' \neq s_\tau \hat{s}$, and in particular for strategy profile s. ∎

Proof of Corollary 2. 1. Suppose that for every arc-cut set $\hat{A} \subset A'$, there does not exist $i \in \iota\left(\hat{A}\right)$ such that i is a source with respect to $\hat{D} = \left(\iota\left(\hat{A}\right), \hat{A}\right)$. Hence both summation terms in $\Upsilon_{\hat{A}}\left(i, \hat{D}, \delta\right)$ are non-trivial or i is a sink with respect to $\hat{D}$, in which case full maintenance is a best response to full maintenance by one's opponents. Let all players play the equilibrium strategy $\hat{s}(D')$ in the proof to Theorem 2, namely in Equation 26 above. We have shown in Lemma 1 above that there exists a unique $\delta^c \in \Delta$ such that $\Phi_{PBE}(i, \hat{A}, \delta^c) = 0$ and $\Phi_{PBE}(i, \hat{A}, \delta) > 0$ for all $\delta < \delta^c$ when i is not a source or sink with respect to $\hat{A}$. If i is a sink with respect to $\hat{A}$, then $\Phi_{PBE}(i, \hat{A}, \delta) = 0$ for all $\delta \in \Delta$. Hence there exists $\delta' \in \Delta$ such that $\Phi_{PBE}(i, \hat{A}, \delta') = 0$ and for all $\delta \in [\delta', 1)$, $\hat{s}(D')$ is a PBE of the network game on D under which $\hat{A}$ is fully maintained, for all $\hat{A} \subset A'$. Hence D' can be fully maintained in a PBE of the network game on D for all $\delta \in [\delta', 1)$.

2. By contrapositive, suppose that there does not exist $D'' \supset D'$ such that D'' is fully maintained under (s, μ), or for all $D'' \supset D'$ there is some arc-cut set $\hat{A} \subset A''$ which has a source with respect to $\hat{A}$. If there does not exist $D'' \supset D'$ such that D'' is fully maintained under (s, μ), then the fact that D' is not maintained is trivial, since $D'' \supset D'$ includes the case $D'' = D'$. Alternatively, suppose that for all $D'' \supset D'$ there is some arc-cut set $\hat{A} \subset A''$ which has a source with respect to $\hat{A}$. In this case,

$$\Upsilon^i_{\hat{A}}\left(\alpha^i, \delta\right) = (\bar{c} - \delta) \sum_{ij \in NE_D(i)} \delta^{\alpha_{ij}} > 0 \tag{28}$$

for all $\delta \in \Delta$. Clearly, this implies that i is never willing to play a full maintenance equilibrium on $\hat{A}$, and hence there does not exist $D'' \supset D'$ which is fully maintained in a PBE of the network game on D. Hence we have shown the necessary conditions. ∎

A.3 Proofs for belief-free network enforcement

Proof of Theorem 3. 1. Let us construct a strategy profile $\hat{s}$ and show that this is a BFE with respect to all $\mu \in \Psi(\hat{s})$ if the stated conditions are satisfied. Let every player $i \in N$

play strategy $\hat{s}_i(D')$ with respect to D, where strategy $\hat{s}$ is given in the proof to Theorem 2. Note that we focus on the case $|N'| \geq 3$ for any component of D, since if $|N'| = 2$ then $\Phi_{BFE}(D') = \bar{c}$ and hence full maintenance is only possible in absence of network enforcement (the case $|N'| = 1$ is ruled out by the assumption that the set of arcs is nonempty). Hence triadic closure will apply to any component.

First, notice that $\Psi(\hat{s})$ is non-empty because $\hat{\mu}$ from the proof to Theorem 2 is always in Ψ. Then, to show that $\hat{s}$ is a BFE with respect to all $\mu \in \Psi(\hat{s})$, we need to show that (1) the severance response specified by $\hat{s}$ is optimal for all $\mu \in \Psi(\hat{s})$, and (2) players are willing to maintain all of their arcs in D' if the incentive compatibility condition, $\delta \geq \Phi_{BFE}(D')$ for all $i \in N'$, is satisfied, and all opponents maintain all arcs in D', for all $\mu \in \Psi(\hat{s})$. First, let us show that the severance responses are optimal, given triadic closure. Consider any two nodes $i, j \in N'$, with $\underline{ij} \in A'$, and consider any severing initiated by j on $\underline{ij}$. Let us consider the two possible cases. First, suppose $ij \in A'$. By triadic closure, for any $jk \in A'$, it must be the case that $ik \in A'$. Hence clearly $\hat{s}$ specifies a mutual best-response with respect to nodes k such that $jk \in A'$. Notice also that for any $l \in A'$, such that $\underline{jl} \notin A'$ but $\underline{il} \in A'$, it must be the case that $il \in A'$, since if it were the case that $li \in A'$, then triadic closure would imply $\underline{jl} \in A'$, which has already been assumed not to hold. But if under $\hat{s}$ it is optimal for i to sever immediately on all arcs connected to nodes that j is also connected to, then we have shown that i will become a source with respect to all remaining arcs, and hence should sever all remaining (out-)arcs immediately (it has been shown previously, in Corollary 2 that if a node becomes a source with respect to all remaining arcs and $\delta < \bar{c}$, then immediate severance is optimal). Second, suppose $ji \in A'$, and j severs arc ji. By triadic closure, for every $\underline{ik} \in A'$, there exists $\underline{jk} \in A'$, and hence immediate severance on all arcs $\underline{ik}$,$\underline{jk}$ is a mutual best response for i and j. Hence we have shown that the deviation-response conditions are optimal; notice that they were shown to be optimal without regard to μ.

Second, let us verify the incentive compatibility condition. Notice that under $\hat{s}$, the optimal deviation program involves severance of all out-arcs simultaneously, and severance of all in-arcs in the subsequent period. This is because, by triadic closure, on all out-arcs one is always connected to all of one's opponents, hence the above deviation specification is optimal, since one can wait one period on all in-arcs and no longer in order to still obtain payoff 1 on each severed in-arc, while severing out-arcs immediately is of course optimal. On all in-arcs, one's opponent is connected to all of one's remaining opponents, and hence since one has at least one out-arc that one severs immediately, the remaining opponents will hence observe severance in the following period, so it is optimal to sever in-arcs in the following period. The only remaining case is if one is a sink with respect to D', but in that case one will never want to sever; given that one is a sink there must be other players in the graph who are not sinks with respect to D', hence sinks will never determine the critical value of δ necessary for full maintenance on D' to be incentive compatible. Hence $\Phi_{BFE}(D')$ is derived from the following equation, based on the optimal deviation program described,

$$\delta^2 + \delta\left(r\left(D'\right) - \underline{c}\right) - r\left(D'\right)\bar{c} \geq 0. \tag{29}$$

Hence if $\delta \geq \Phi_{BFE}(D')$ then each $i \in N$ is willing to maintain all arcs in D' in a full maintenance equilibrium, since even the optimal deviation brings a net loss in payoff. This concludes the proof that the stated conditions are sufficient for the existence of a BFE on a digraph D under which D' is fully maintained.

2. Suppose that s is an equilibrium of the network game on D with respect to all $\mu \in \Psi(s)$, where we denote the set of fully maintained arcs in D under s by

$$\hat{A} = \left\{ij \in A | \sigma_{ij}^{t}(s) = (M, \varnothing, \varnothing)\ \forall\ t \geq 0\right\}. \tag{30}$$

Namely, $\hat{D} = \left(\iota\left(\hat{A}\right), \hat{A}\right)$ is the largest subdigraph that is maintained under s. Clearly, if D' is to be fully maintained under (s, μ) then it must be the case that $D' \subset \hat{D}$, and $\hat{D}$ must satisfy any necessary conditions for the existence of a full maintenance equilibrium. For every $ij \in A - \hat{A}$ let τ_{ij} be the time period in which ij is severed, which is given in $\sigma_{ij}^{\infty}(s)$, and clearly must be finite. Let $\tau = \max_{ij \in A - \hat{A}} \tau_{ij}$, which exists because D is finite.

In order to show the necessity of the incentive compatibility and graph structure conditions for the maintenance of $\hat{D}$, we proceed in two steps: (1) Argue that any BFE on D under which $\hat{D}$ is fully maintained must involve all players following strategy $\hat{s}\left(\hat{D}\right)$ from the proof to Theorem 2 as of period τ, (2) show that under $\hat{s}$, the necessary conditions must hold. First, we argue that any BFE on D under which $\hat{D}$ is fully maintained must involve all players following strategy $\hat{s}\left(\hat{D}\right)$ from the proof to Theorem 2 as of period τ. This must hold, because Ψ always includes the particular system of beliefs $\hat{\mu}$ that has each player believe that *all* unobserved arcs have been severed, upon observing any deviation from full maintenance by an opponent. This is because Ψ, characterizing all beliefs consistent with PBE, puts no structure on beliefs off the equilibrium path; under full maintenance equilibria, any observed severing is off the equilibrium path. As has already been argued in the proof to Theorem 2, given that one believes that all unobserved arcs have been severed, it is always optimal to sever all remaining arcs. This is because under PBE one always believes that one's opponents will follow the equilibrium continuation strategy with respect to the remaining structure of the graph, and hence severing all remaining active arcs would be incentive compatible. Hence we have shown that for subdigraph $\hat{D}$ to be fully maintained, players must play strategy $\hat{s}\left(\hat{D}\right)$ as defined in the proof of Theorem 2.

Second, now that we have fixed the equilibrium strategy $\hat{s}$, we show the necessary conditions on graph structure and incentive compatibility for $\hat{s}$ to be optimal for all $\mu \in \Psi$. First, by way of contradiction, suppose that there exists a BFE of the network game on D, under which $\hat{D}$ is fully maintained, but $\hat{D}$ does not satisfy triadic closure. Namely, there exist nodes $j, k, l \in \hat{N}$, such that $\left\{lk, \underline{kj}\right\} \subset \hat{A}$, but $\underline{lj} \notin \hat{A}$. As already noted, it has been shown in Corollary 2 that there can be no sources with respect to the maintained graph in a PBE with $\delta < \bar{c}$; hence node l must have at least one in-arc. Now, suppose that j severs arc $\underline{kj}$. As argued in Step (1), in a belief-free equilibrium k must respond by severing all arcs in the following period; this is optimal under the belief system $\hat{\mu}_k$, under which k puts full probability mass on the history under which all of l's unobserved arcs were severed simultaneously with $\underline{kj}$. But given that arc $\underline{lj}$ does not exist, this response is not necessarily optimal. Namely, k can also entertain belief system $\tilde{\mu}_k$, under which k puts full probability mass on the history under which only arc $\underline{kj}$ was severed, and all remaining arcs in the digraph are maintained. Given that it is optimal for k to maintain arc lk as long as possible, and any further severance from player j would take at least two periods to reach player l under $\tilde{\mu}_k$, k's best response is clearly to maintain arc lk for at least one more period. But two different strategies being optimal under two different consistent beliefs, $\hat{\mu}, \tilde{\mu} \in \Psi$, violates the belief-free condition. Hence $\hat{D}$ must satisfy triadic closure.
Next, we show that $\delta \geq \Phi_{BFE}\left(\hat{D}\right)$ is also necessary. To show this, as usual, notice that $\hat{s}$ is the most severe deviation response, since it implies that one is punished as soon as possible on all possible arcs, for any deviation from full maintenance. Hence $\hat{s}$ implies the lowest δ as of which each player is willing to not deviate, which is formalized in Φ_{BFE}. This shows that if $\delta \geq \Phi_{BFE}\left(\hat{D}\right)$ under $s_\tau \hat{s}\left(\hat{D}\right)$, then it must be the case that $\delta \geq \Phi_{BFE}\left(\hat{D}\right)$ for any alternative strategy profile $s_\tau s' \neq s_\tau \hat{s}$, and in particular for strategy profile s. ■

Proof of Corollary 3. 1. We have already shown in Theorem 3 that if D' satisfies triadic closure then it may be possible to have a full-maintenance BFE of the network game on D under which D' is fully maintained, for sufficiently large $\delta < \bar{c}$, with an additional incentive compatibility constraint. Let us show that the absence of sources combined with triadic closure with respect to D' allows for an interval of $\delta < \bar{c}$ under which full maintenance is possible. Notice that the condition characterizing the δ as of which at least one player wants to deviate under $\hat{s}$ (from the proof to Theorem 2) can be simplified to

$$\delta \frac{(\underline{c} - \delta)}{(\delta - \bar{c})} \geq r\left(D'\right). \tag{31}$$

If there exists a source with respect to D', then $r\left(D'\right) = \infty$, and hence the inequality can only hold at $\bar{c}$, where network enforcement is no longer necessary for full maintenance. However, if there are no sources with respect to D', then $r(D')$ is finite, and hence the above inequality holds for sufficiently large $\delta < \bar{c}$, since $\delta(\underline{c} - \delta) / (\delta - \bar{c})$ is strictly increasing in δ, for $\delta \in \Delta$.

2. By contrapositive, suppose that there does not exist $D'' \supset D'$ such that (i) D'' is fully maintained under s, or (ii) D'' has no sources, or (iii) D'' satisfies triadic closure. That (i) implies that D' is not fully maintained under s is trivial, since $D'' \supset D'$ includes the case $D'' = D'$. That (ii) implies that D' is not fully maintained under s follows from Corollary 2, since BFE is a refinement of PBE. That (iii) implies that D' is not fully maintained under s has already been shown in Theorem 3. Hence we have shown the necessary conditions. ■

B Appendix: Characterization of Triadic Closure

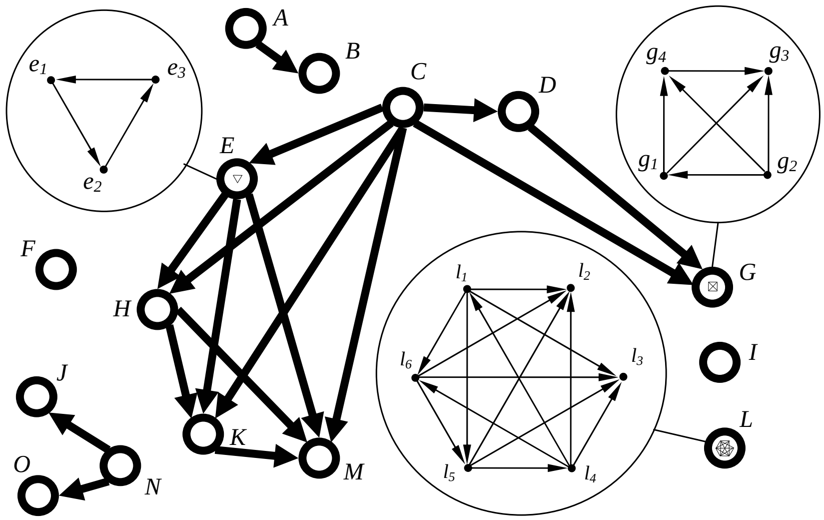
Triadic closure is a local property defined on the neighborhood of individual nodes. We now give a global characterization of networks satisfying triadic closure. To give the characterization, let $\sim$ be a binary relation on $N(D)$ with $i \sim j$ if there exists both ij-path $\subset D$ and ji-path $\subset D$. It is straightforward to verify that $\sim$ is an equivalence relation on the set of nodes of a digraph. Denote by $N_1, \ldots, N_n$ the unique partition of N according to $\sim$, which we call the *strong partition of N*.

Definition 4 (Quasi-tournaments) *Let $N_1, \ldots, N_n$ be the strong partition of $N(D)$. Digraph D is a quasi-tournaments if $i, j \in N_k \Longrightarrow \underline{ij} \in N_k \ \forall \ k$ (that is, each part N_k can form a tournament subdigraph with arcs that exist in $A(D)$), and there exists meta-digraph* $\mathbf{D} = (\mathbf{N}, \mathbf{A}) = (\{N_1, \ldots, N_n\}, \mathbf{A})$, *such that*

1. *Every node of $\mathbf{N}(\mathbf{D})$ represents one N_k ($k \in \{1, \ldots, n\}$) in the strong partition of the nodes of D,*
2. *There exists an arc between two nodes in $\mathbf{D}$, $N_k N_l \in \mathbf{A}$, if and only if $i \in N_k$, $j \in N_l \Longrightarrow ij \in A$, and*
3. *$\mathbf{D}$ satisfies triadic closure.*

This definition can be understood as saying that a digraph is a quasi-tournaments if it can be partitioned into strongly-connected tournaments, and the parts satisfy a global version of triadic closure. The following diagram depicts a meta-digraph representation and is useful for illustrating the properties of a quasi-tournaments network.

Figure 8: Quasi-tournaments.



The dark circles in Figure 8 represent strongly connected tournaments – parts of the digraph in which each node is linked and there is a directed path from any node to any other node. For example, nodes E, G and L are the tournament subdigraphs illustrated. The dark arrows in the meta-digraph represents "flows" of arrows in the digraph, such that if any two nodes of the meta-digraph are connected then there must be an arc from every digraph node in one meta-digraph node, to every digraph node in the other meta-digraph node. This certainly does not preclude any two parts not being connected at all, as with B and C, or being isolated, as with F and I. Finally, we see that quasi-tournaments requires that the meta-digraph satisfy triadic closure. For example, since C is upstream of D, and D is connected to G, C must be connected to G. Also, since D is upstream of G and C is connected to G, D and C must be connected.

Proposition 1 *Suppose that digraph D satisfies assumption 1. Then D satisfies triadic closure if and only if D is a quasi-tournaments.*

Proof. We start with a number of useful definitions. A digraph D has a *hierarchy* if there exists a non-singleton partition $\hat{N}_1, \ldots, \hat{N}_n$ of $N(D)$ such that for every $k, l \in \{1, \ldots, n\}$, $k \neq l$, either (i) $i \in \hat{N}_k$ and $j \in \hat{N}_l \implies \underline{ij} \notin A(D)$ or (ii) $i \in \hat{N}_k, j \in \hat{N}_l$ and $ij \in A(D) \implies j'i' \notin A(D)$ for all $i' \in \hat{N}_k, j' \in \hat{N}_l$, and in addition $i, j \in \hat{N}_k \implies \exists$ ii-cycle $\left(\hat{A}, \hat{N}\right)_{ii}$ such that $j \in \left(\hat{A}, \hat{N}\right)_{ii}$ and $\hat{N} \subset \hat{N}_k$. This says D has a hierarchy if there exists a partition of its nodes such that for any two parts either none of their nodes are connected, or arcs only flow one-way from one part to the other, and within a part all nodes are strongly connected (that is, for any two nodes in a part, the part contains a cycle on which both nodes appear). We will show below that every digraph satisfying assumption 1 can be partitioned into a hierarchy.

A digraph D has a *strong hierarchy* if there exists a non-singleton partition $\hat{N}_1, \ldots, \hat{N}_n$ of $N(D)$ such that the following hold:

(i) $i \in \hat{N}_k, j \in \hat{N}_l$, $k \neq l$, and $ij \in A(D) \implies i'j' \in A(D)$ for all $i' \in \hat{N}_k, j' \in \hat{N}_l$; we denote this as $\hat{N}_k \lhd \hat{N}_l$,

(ii) for any distinct $k, l, m \in \{1, \ldots, n\}$, (1) $\hat{N}_k \lhd \hat{N}_l$ and $\hat{N}_l \lhd \hat{N}_m \implies \hat{N}_k \lhd \hat{N}_m$, and (2) $\hat{N}_k \lhd \hat{N}_l$ and $\hat{N}_m \lhd \hat{N}_l \implies \hat{N}_k \lhd \hat{N}_m$ or $\hat{N}_m \lhd \hat{N}_k$.

Strong hierarchy strengthens hierarchy by requiring the existence of a partition of nodes such that if two parts in a partition are connected, then there is a flow from *all* nodes in one part to *all* nodes in the other part, and that the downstream flow property can be thought of as having a transitivity property and a version of a completeness property. Notice that part (i) of this definition implies that if D has a strong hierarchy then it has a hierarchy. We are now in a position to prove the result.

($\implies$) We proceed in three steps. First, we show that every digraph has a hierarchy partition, then we show by triadic closure that every part of a hierarchy must be a tournament, and finally we show that if every part of a hierarchy is a tournament then by triadic closure the digraph must have a strong hierarchy and hence be a quasi-tournaments.

To show that every digraph D has a hierarchy partition, define the following relation, $R_{N(D)}$, on the nodes of D. For any $i, j \in N(D)$ we say that $iR_{N(D)}j$ if ij-path $(A, N)_{i,j} \subset D \implies ji$-path $(A, N)_{j,i} \subset D$. To show that $R_{N(D)}$ is an equivalence relation, notice that $R_{N(D)}$ is (1) reflexive: if there exists an ii-path then it can be used twice, while if such a path does not exist then the definition is vacuous, (2) symmetric: if existence of an ij-path implies existence of a ji-path, then existence of a ji-path automatically implies existence of an ij-path, (3) if $iR_{N(D)}j$ and $jR_{N(D)}k$, then there exists ik-path $(A, N)_{i,k} \subset D$, which can be constructed by piecing together the ij-, jk-, kj-, and ji-paths together, in that order. Hence $R_{N(D)}$ is an equivalence relation, and so it directly follows that $R_{N(D)}$ implies a partition on $N(D)$. Notice three facts: first, each part of the partition is strongly connected, since there exists a cycle between any two nodes in a part; second, sinks and sources with respect to D form their own, singleton parts; finally, the partition is unique.

Second, we argue by triadic closure that if two nodes i and j are in a part of a hierarchy of a digraph D, then there must be $\underline{ij} \in A(D)$; that is, each part of a hierarchy must be a tournament. Recall from step 1 that we showed that any part of a hierarchy must be strongly connected; namely, there must exist a cycle between any two nodes i and j. We argue that it must be the case that all nodes on any such cycle must share an arc, and hence i and j must also be connected. Without loss of generality, let us label the cycle as follows, taking the cycle that begins and ends at node i:

$$i = j_0, j_0 j_1, \ldots, j_{k-1} j, j = j_k, j j_{k+1}, \ldots, j_{k+l} j_0, j_0 = i. \tag{32}$$

Let us presume that these are the only arcs that exist between the nodes $j_0, j_1, \ldots, j_{k+l}$. If some such arcs do exist, then we can simply ignore their creation in the following argument; this step shows that they *all* must exist. We proceed by finite induction. First, by triadic closure, arcs $j_0 j_2, \ldots, j_{k+l} j_1 \in A(D)$, since $(j_0 j_1, j_1 j_2), \ldots, (j_{k+l} j_0, j_0 j_1) \in A(D)^2$. An an inductive hypothesis, assume that for some $m \leq k+l-2$ the arcs $j_1 j_m, \ldots, j_0 j_{m-1} \in A(D)$.. Then again by triadic closure, it follows that $j_0 j_m, \ldots, j_{(k+l)} j_{m-1} \in A(D)$, since we already had that $j_0 j_1, j_1 j_2, \ldots, j_{k+l} j_0 \in A(D)$ by assumption of looking at a part of a hierarchy partition. Hence any cycle $C \subset D$ in a part of a hierarchy is a tournament, and hence any two nodes i and j in a part must be connected.

Finally, we show that if D has a hierarchy, each part of which is a tournament, then D must have a strong hierarchy. Let us show each part of the definition of strong hierarchy in turn. Recall again that each part of a hierarchy that is not a source or sink with respect to D contains a cycle between any two of its nodes (notice that if a node is a source with respect to D then part (i) is trivial, while if a node is a sink with respect to D then it is the implication of (i) that must hold). In particular, if $\hat{N}_k$ and $\hat{N}_l$ are two distinct parts of a hierarchy partition, with $i \in \hat{N}_k, j \in \hat{N}_l$ and $ij \in A(D)$, then first consider any node $j' \in \hat{N}_l$, $j' \neq j$. Given that $\hat{N}_l$ must be a tournament, by step 2, it follows immediately that $\underline{jj'} \in A(D)$, and by triadic closure that $ij' \in A(D)$ for all $j' \in \hat{N}_l$. Now consider any node $i' \in \hat{N}_k$, $i' \neq i$. By step 1 there must exist a cycle between i' and i, and in particular, the directed path $i' = j_0, j_0 j_1, \ldots, j_{k-1} j_k, j_k = i$. Arguing by triadic closure we can proceed by a similar finite induction argument as in step 2, to show that $i'j \in A(D)$, from which it follows immediately by the argument already given in this step that $i'j' \in A(D)$ for all $i' \in \hat{N}_k, j' \in \hat{N}_l$ (notice that it must always be $i'j'$ and not $j'i'$, else it would have to be that $j' \in \hat{N}_k$ by strong connectivity). Now consider part (ii) of the definition of strong hierarchy. Part (ii)(1) is immediate by triadic closure. Part (ii)(2) again follows by triadic closure, and by the partition of nodes according to strong connectivity.

($\Longleftarrow$) Suppose that D is a quasi-tournaments. Under the quasi-tournaments partition, for any $i, j \in N(D)$, $ij \in A(D)$ only if (1) i and j are in the same part of the partition, $\hat{N}_{ij}$, or (2) i is in one part of the partition, $\hat{N}_i$, and j is in a different part, $\hat{N}_j$. Under

(1) $ij \in A(D)$ and $\underline{jl} \in A(D) \implies \underline{il} \in A(D)$, since it is either the case that $l \in \hat{N}_{ij}$, in which case it is immediate that $\underline{il} \in A(D)$ since $\hat{N}_{ij}$ is a tournament, or else $l \notin \hat{N}_{ij}$, but $\underline{il} \in A(D)$, by part (i) of the definition of strong hierarchy. Under (2), $ij \in A(D)$ and $\underline{jl} \in A(D) \implies \underline{il} \in A(D)$, since it is either the case that $l \in \hat{N}_j$, in which case it is immediate that $\underline{il} \in A(D)$, by part (i) of the definition of strong hierarchy, or l is in a different part from j, in which case it follows that $\underline{il} \in A(D)$, either because $l \in \hat{N}_i$ and $\hat{N}_i$ is a tournament, or $l \notin \hat{N}_i$, which case $\underline{il} \in A(D)$ follows from part (ii) of the definition of strong hierarchy. ■

C Appendix: Sequential Network Enforcement

This appendix considers an alternative solution concept for the network game under private information, to further highlight some aspects of the interplay between balancing and transmission. Under public information balancing is crucial to network enforcement but transmission is not relevant because all deviations from equilibrium are publicly observed. Under private information we observe the same role of balancing in network enforcement, but our analysis in Sections 5 and 6 also highlights the importance of transmission in network enforcement. We have already constructed two classes of equilibria under which transmission is always incentive compatible.

1. In the PBE construction of Theorem 2 transmission is always incentive compatible because, upon observing a deviation, players believe that all others have also observed a deviation from equilibrium. In that sense, they believe that there is nothing for them to transmit, but also that there is no reason for them not to transmit. Hence when considering deviations from full maintenance equilibria, players serving a critical balancing role in the network must do so with the knowledge that the response to any deviation will be swift and complete.
2. In the BFE construction of Theorem 3, players do not need to form beliefs about whether others in the network have (or have not) observed deviations from equilibrium. If player i observes a deviation, the equilibrium strategy prescribes that his opponent will transmit this to all other nodes in the network in the next period. Hence, player i has no reason not to do likewise (which in turn justifies his opponent's response). This type of coordinated transmission is possible in BFE because in a network satisfying triadic closure all players are essentially connected via the network (see Proposition 1). And again, such immediate and coordinated transmission leads critical players in a balancing role to be more willing to play full maintenance equilibria.

Hence, in both the PBE and BFE considered thus far, transmission is incentive compatible by construction. This is not necessarily the case in a *sequential equilibrium* (SE) which we consider in this appendix (Kreps and Wilson, 1982). In a network game under private information, SE is a refinement of PBE, and is itself refined by BFE. The refinement of PBE to SE imposes structural consistency on beliefs, which limits the extent to which a player can believe that extensive severance has occured in unobserved parts of the network, upon observing a deviation. Given such a limitation on beliefs, there are network structures in which transmission is not incentive compatible, and hence some critical players may not be willing to play a balancing role. In this appendix, we first show that with additional structure on beliefs there are network structures in which transmission is not incentive compatible, weakening balancing incentives for some players. We do not give a characterization of sequential equilibria in the style of Theorems 1-3, but we explore its implications in a number of examples. We also give a sufficient condition for the existence of sequential equilibria with network enforcement in terms of an economically meaningful and well-studied network architecture (a generalization of "Eulerian" digraphs).

C.1 Sequential equilibrium

To define a SE we need notation for behavioral strategies. Let $\Delta\left(\{M,S\}^{|m_i^{t,1}(\cdot)|}\right)$ be the simplex (distribution) over action vectors of length $|m_i^{t,1}(\cdot)|$.

$$b_i^t : H_i^t \to \left[\Delta\left(\{M,S\}^{|m_i^{t,1}(\cdot)|}\right)\right], \tag{33}$$

is the behavioral strategy of player i in period t on all of his active arcs. Denote the set of behavioral strategy combinations of player i by B_i, and let $B = \times_{i\in N} B_i$ and denote its interior int$B$. $\sigma^t : B \to \Delta H^{t+1}$, where $\sigma^t(b) = \left(\sigma_{ij}^t(b)\right)_{ij\in D}$ is the distribution over the realized path of play at $t+1$, given behavioral strategy profile $b \in B$. Finally, denote the probability that history h^t is realized under behavioral strategy combination b by $P_b(h^t)$.

We say that a system of beliefs μ is *structurally consistent* with respect to a strategy s of the network game on digraph D if for every $i \in N(D)$ $\exists$ $(b_n^i)_{n\geq 0} \in$ int$B$ such that $b_n \to s$ (where convergence is point-wise), and for all $t \geq 0$

$$\mu_{i,n}^t\left(h^t|h_i^t\right) = \frac{P_{b_n^i}\left(h^t\right)}{\sum_{\hat{h}^t\in\langle h_i^t\rangle} P_{b_n^i}\left(\hat{h}^t\right)} \tag{34}$$

and $\mu_{ni}^t \to \mu_i^{t*}$ $\forall$ t. If, in addition, $b^i = b^j$ for all $i,j \in N$ in the construction of beliefs above, we say that beliefs are *common structurally consistent* (or simply *common*). Denote the set of beliefs that are structurally consistent with respect to a strategy s by $\bar{\Psi}(s)$, and the set of beliefs that are common with respect to s by $\hat{\Psi}(s)$. Note that $\hat{\Psi}(s) \subset \bar{\Psi}(s) \subset \Psi(s)$, and for most $s \in S$ the inclusions are strict.

Definition 5 (Sequential Equilibrium) *A strategy s and a system of beliefs μ are a sequential equilibrium of a network game on D with radius of information ρ if $\mu \in \hat{\Psi}(s)$, and*

$$R_i^t\left(s, \mu_i^t | h_i^{t,\rho}\right) \geq R_i^t\left(\hat{s}_i, s_{-i}, \mu_i^t | h_i^{t,\rho}\right), \tag{35}$$

for all $h_i^{t,\rho} \in H_i^{t,\rho}$, $\hat{s}_i \in S_i$, and $i \in N(D)$.

Consistency with respect to strategy s, which is the constraint on beliefs in a PBE, implies that players should believe that others are following strategy s as long as they have observed nothing to the contrary. Structural consistency implies that if player i observes a deviation from strategy s in period t and initially believes that the history is $h^t \neq \sigma^t(s)$, then in $t' > t$ player i should believe that the history is $\sigma^{t'}(s|_{h^t})$ unless i observes a further deviation to the contrary.[30] In addition, if two players' beliefs do not disagree on the basis of differential observations, then common beliefs requires that their beliefs should not disagree.[31]

To illustrate the relevant additional structure on beliefs implied by SE, consider a PBE on the cycle network D_6^C from Figure 4 and suppose that j_6 severs the arc $j_6 j_1$ in period 0. It is consistent for j_1 to believe that all arcs other than $j_2 j_1$ have been severed, and therefore to sever $j_2 j_1$ in period 1 (in anticipation of j_2's severance). However, this cannot be structurally consistent. Under structural consistency, beliefs are generated by a sequence of behavioral strategies that converges to the pure strategy profile in question. Behavioral strategies can feature strategic interconnections between arcs. However, the randomization in two players' behavioral strategies are independent of each other. Along a sequence of assessments μ_n it is therefore orders of magnitude more likely that $j_3 j_2$ has not been severed yet. As a result, in a SE j_1 cannot sever $j_2 j_1$ in period 1.

C.2 Example of diagonal and weak networks

The networks in Figures 9 and 10 illustrate how structural consistency refines PBE. Here we demonstrate that digraphs D_1^D from Figure 9 and D_1^w from Figure 10 can be fully maintained

[30]Player i's beliefs do not need to put point mass on a single history; we assume this only for the purpose of clearer exposition.

[31]If we consider continuation games after any history as games of incomplete information in which player types encode knowledge about the status of the network, then common structural consistency imposes a common prior condition on beliefs regarding types.

Figure 9: Cycle with diagonal: D_1^D and D_2^D.

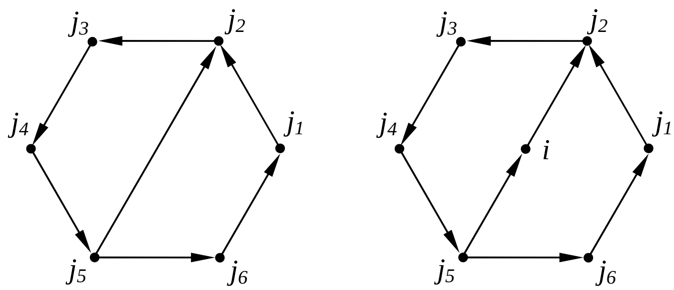
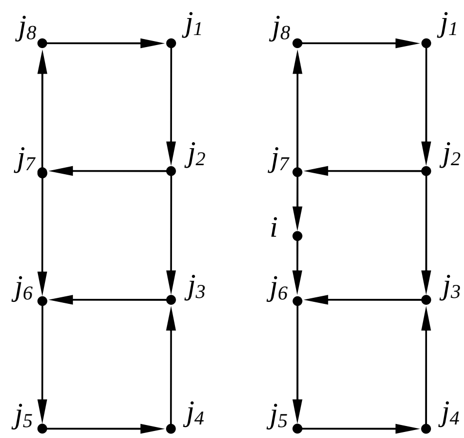


Figure 10: Weak digraphs: D_1^w and D_2^w.



in a SE on an open subset of Δ, while D_2^D and D_2^w cannot be fully maintained in a PBE with structurally consistent beliefs for any $\delta < \bar{c}$.

Consider first the network D_2^D, and suppose that strategy s prescribes full maintenance of the whole network but j_1 severs j_1j_2 in period 0. [32] If beliefs are structurally consistent, j_2 can believe that the fully maintained digraph is now either $(N_2^D, A_2^D - \{j_1j_2\})$ or $(N_2^D, A_2^D - \{j_1j_2, j_6j_1\})$ (or any mixture of these beliefs). However, s cannot be a SE because under any of these beliefs, j_2 will not sever any further relationships. Under private information and structurally consistent beliefs, players j_1, j_2, and j_6 do not believe that anyone else in the network has observed any deviations, and regardless of the exact status of the network, have no incentive to transmit information regarding the severance of j_1j_2. As a result, severing

[32] It is notable that an unappealing feature of SE is that its ability to refine PBE through imposing structural consistency on beliefs strongly depends on deviations occuring early in a network game, and in particular in period 0.

j_1j_2 is optimal for j_1 for any $\delta \in \Delta$ and s cannot be a SE; in other words, due to the restriction on transmission of deviations under SE, j_1 is no longer willing to play a balancing role in D_2^D. On the other hand, D_1^D can be fully maintained in a SE because if j_1 severs j_1j_2, it is structurally consistent for j_2 to believe that arc j_6j_1 has also been severed. In that case, it is possible for j_6 and j_2 to coordinate transmission in the sense that it is optimal for j_6 to transmit when j_2 transmits, and optimal for j_2 to transmit when j_6 does. Hence j_1 is again willing to play a balancing role. A similar argument can be used to demonstrate the difference between networks D_1^w and D_2^w.

C.3 Coordinated transmission and forward induction

The examples of Figures 9 and 10 illustrate the primary channel by which SE refines PBE. If we require that beliefs be structurally consistent, a player i who observes a severance by j on ji in period 0 identifies very precisely where the deviation occurred. Moreover, player i believes that other nodes in the network will continue to maintain relationships unless i, or another node in j's neighborhood who has already observed a deviation, transmits that the fully maintained path of play has been left. Transmission is incentive compatible in two cases:

i. Players *coordinate* on transmission, i.e., it is optimal for player i to transmit if some player $k \in \iota(NE_j)$ transmits, and it is optimal for k to transmit when i does. In this case, i and k essentially play a coordination game. Transmission is a payoff dominated outcome of the coordination between i and k, but full maintenance is incentive compatible because of the threat to coordinate on this low outcome.
ii. Transmission is not incentive constrained *ex-post*, i.e., given that ji has been severed, there is a transmission route now available to i that does not involve a payoff loss given the continuation strategy of nodes in $N - \{j\}$. The simplest case is when i can balance arc ji against exactly one arc ik, in the sense that ik is only maintained because ji is maintained and vice versa. In this case, after ji is severed, there is no reason for i to maintain ik, and we only need to check that there is a global coordination of transmission such that j will in fact be punished for the original deviation. We show below that if the network has an appropriate topology, transmission can be globally coordinated in this way.

We have already identified a class of networks in which coordinated transmission is always possible, namely networks satisfying triadic closure. BFE is a refinement of SE and network enforcement is therefore always possible in a SE on networks satisfying triadic closure. However, SE which require coordinated transmission and are not belief-free often contradict elementary forward induction reasoning.[33] Firstly, if players already believe that a particular node is the unique source of deviation, they may also reason about the nature of that deviation. Consider a strategy s under which D_1^D from Figure 9 is fully maintained by network enforcement. If beliefs are structurally consistent and arc j_1j_2 is severed in period 0, j_2 believes that only j_1 has deviated from s. If j_1j_2 was the only arc that was severed, not transmitting would be the mutual best response. In this case, j_1 is strictly better off than on the path of play under s and j_2 is strictly better off than on the continuation path under s. Hence, strategy s is only a SE if j_2 believes that j_6j_1 has also been severed. These beliefs are structurally consistent, and allow for coordinated transmission between j_2 and j_6, but they imply that j_2 believes that j_1 has acted irrationally, even though under the first scenario j_2's deviation is perfectly rational. It seems to contradict a belief in rationality for j_2 to draw the latter conclusion. Heuristically, this type of forward induction reasoning is closely related to the intuitive criterion applied to signalling games (Cho, 1987). Strong belief in rationality is defined in Battigalli and Sinischalchi (2002), and the coordination subgames they use to illustrate ideas have a similar structure to the two-player continuation games in which transmission is coordinated in the present example.

There is a natural extension to the previous forward induction logic. If players reason about the nature of a deviation, they may also go further and reason about how a node who deviated intends to play the continuation game. Consider again a strategy s that prescribes full maintenance of D_1^D and suppose that beliefs are structurally consistent. Suppose now that j_5 severs j_5j_2 in period 0. j_2 must believe that all players other than j_5 are still following the strategy s. There are two reasons why j_2 would transmit following j_5's deviation: (1) j_2 believes that j_5 has already severed j_5j_6 and that severance is therefore spreading through the network already, or (2) only j_5j_2 has been severed but j_5 intends to transmit this to other players next period, and j_2 should therefore do likewise. Under both scenarios, j_5 is strictly worse off than on the equilibrium path of play. Should j_2 not therefore interpret j_5's deviation as a signal that neither (1) nor (2) apply, i.e., that j_5 has not and does not intend to sever j_5j_6? In this last scenario, j_5 is strictly better off than on the path of play under s, and j_2 is strictly better off than on any of the alternative continuation paths. Hence, full maintenance of D_1^D is being maintained by j_2's punishment threat to coordinate on

[33] BFE and the type of forward induction reasoning we describe here lead to disjoint but intersecting sets of equilibria. As we will see at the end of this section, while cycles with more than 3 nodes cannot be BFE they satisfy our elementary forward induction reasoning; on the other hand, tournaments with a sink can support BFE, while such equilibria can be eliminated by application of elementary forward induction reasoning.

transmission, even though this threat is surely not credible if we argue that j_5 can signal intent with the deviation on $j_5 j_2$ in period 0.

Allowing for forward induction reasoning refines SE in a sensible way. At first sight, there appears to be little difference between networks D_1^D and D_2^D (as well as networks D_1^w and D_2^w) to suggest that they should support fundamentally different strategic outcomes. From a graph-theoretic viewpoint the networks are said to be *topologically equivalent*, and it is surprising that the equilibrium predictions on two topologically equivalent networks should differ in a substantial way (see Chartrand and Lesniak, 2005). However, as the preceding discussion suggests, the SE that involve full maintenance of D_1^D and D_1^w contradict basic forward induction reasoning. SE that accounts for forward induction reasoning better identifies where the relevant incentive constraints on transmission originate. For example, it is the (indirect) relationship between j_5 and j_2 that should be viewed as the essential strategic characteristic of cycle networks with a diagonal, rather than the existence of the additional node i. Likewise, in the weak digraphs of Figure 10, the essential strategic characteristic is the relationship between nodes $\{j_2, j_7\}$ and nodes $\{j_3, j_6\}$.

C.4 Perfectly balanced networks

Using structural consistency and forward induction to allow players to better predict where deviations have occurred and how others will play in the continuation, provides some insight on coordinated transmission and the incentives this places on balancing. In Proposition 3 we show that on an important class of networks, transmission is incentive compatible without any local coordination in punishments. To give and interpret the relevant network topology, recall that network enforcement extends the opportunity for full maintenance because players balance less valuable relationships against more valuable relationships via the strategic interaction at the network level. Proposition 3 shows that if all players occupy the same balancing role in a network, there is a way to organize transmission so that players can respond to severance on an in-arc by severing exactly one out-arc. If this can be done locally, transmission is clearly incentive compatible. With an important result from the graph theoretic literature, we can additionally show that this transmission strategy can achieve the global coordination that is required in an equilibrium of the network game. Moreover, because the SE we construct remove asymmetric balancing roles to ensure that transmission is always incentive compatible, it follows that the equilibria do not contradict the elementary forward induction reasoning we presented in the last Subsection.

Definition 6 (Perfectly Balanced Network) *The digraph D is perfectly balanced if* $r_D(i) =$

1 *for all* $i \in N(D)$.

Perfect balance is defined in terms of a purely local property of the network, namely the ratio of individual nodes. However, the following global characterization of perfectly balanced digraphs is well-known in the graph theory literature (for a proof see Chartrand and Lesniak, 2005), and is used to establish how transmission can be achieved on such networks.

Proposition 2 *Suppose digraph D satisfies Assumption 1. The following are equivalent:*

1. *D is perfectly balanced.*
2. *There exists a partition of $A(D)$, $A_1, ..., A_n$ such that A_i is a cycle for $i = 1, ..., n$. We call such a partition a cycle partition (and denote it generically by $\bar{A}_D$).*

Given this proposition, for a perfectly balanced digraph D let $\bar{\mathcal{A}}(D)$ be the collection of all cycle partitions of D, and define $l(D) = \min_{\bar{A}_D \in \bar{\mathcal{A}}(D)} \max_{A_i \in \bar{A}_D} |A_i|$. Let $\Phi_{SE}(D) = \delta^{l(D)}(\underline{c} - \delta) + (\bar{c} - \delta)$. This is the weighted average of the (normalized) payoff of severing one out- and one in-arc. The weight corresponds to the maximal distance between an out- and in-arc along a cycle in a cycle partition of a perfectly balanced digraph. To maximize the potential for network enforcement we can choose the cycle partition to minimize this cycle length.

Proposition 3 *Suppose the network game on D satisfies Assumptions 1 and 3. If D' is a perfectly balanced subdigraph of D and $\Phi_{SE}(D') \leq 0$, then D' can be fully maintained in a sequential equilibrium of the network game on D.*

Before giving the formal proof of this proposition, we provide the following interpretation. Notice that $\Phi_{SE}(D)$ is decreasing in δ and that, by Corollary 2, $\Phi_{SE}(D) = 0$ has a solution in Δ. Hence, perfectly balanced networks can be fully maintained by network enforcement in a SE. There therefore appears to be an interesting contrast when comparing Proposition 3 to the results on PBE and BFE. For arbitrary D, there is no unique sufficient digraph statistic that determines the critical discount factor, $\delta^c_{PBE}(D)$, for a PBE because of possible interactions between the balancing role and local gatekeeping position of nodes in the network. In a BFE the gatekeeping role is irrelevant because of the opportunity for immediate transmission, and the maximal ratio therefore determines $\delta^c_{BFE}(D)$. In a SE on perfectly balanced networks, the balancing role is constant across all players and the maximal cycle length determines $\delta^c_{SE}(D)$. Recall, however, that the restriction to digraphs that satisfy triadic closure is an implication of BFE, while the restriction to perfectly balanced digraphs in Proposition 3 is an assumption.

Proof. We prove the Proposition by constructing a strategy profile and beliefs $(\hat{s}, \hat{\mu})$, and showing that these are a SE if D' is a perfectly balanced subdigraph of D and $\Phi_{SE}(D') \leq 0$. By Proposition 2, if D' is perfectly balanced then it can be partitioned into cycles. In particular, let us employ the cycle partition $\bar{A}'_D = \arg\min_{\bar{A}_D \in \bar{\mathcal{A}}(D)} \max_{A_i \in \bar{A}_D} |A_i|$. Let every player $i \in N$ play the following strategy with respect to D: sever every arc $\underline{ik} \in A - A'$ in period 0, and play the following strategy with respect to the remaining arcs in D': "maintain every adjacent arc in D' until a severing is observed on an in-arc in D', in which case sever the remaining out-arc on the same cycle immediately; if a severing is observed on an out-arc, sever in a number of periods forward that is three less than the length of the cycle if one's opponent caused the severing, and in a number of periods forward that is two less than the length of the cycle if oneself caused the severing"; formally, for each $i \in N$, and arcs ki,il on each cycle $\hat{A} \subset A'$ such that $i \in \iota\left(\hat{A}\right)$,

$$\left(\hat{s}^t_{ki}\left(h^{t,1}_{il}\right), \hat{s}^t_{il}\left(h^{t,1}_{i}\right)\right)\left(\hat{A}\right) = \left(\begin{cases} S & \text{if } h^t_{il} = \left(S, l, t - \left|\hat{A}\right| - 3\right), \text{ or} \\ & h^t_{il} = \left(S, i, t - \left|\hat{A}\right| - 2\right) \\ M & \text{o.w.} \end{cases}, \begin{cases} S & \text{if } h^t_{ki} = (S, k, t-1), \text{ or} \\ & h^t_{ki} = (S, i, t-1) \\ M & \text{o.w.} \end{cases}\right), \tag{36}$$

with $\hat{s}^0_{i,m}\left(h^{0,1}_i\right)(D') = S$ for all $\underline{im} \in A - A'$. Then we specify beliefs $\hat{\mu}$ as follows, where beliefs about the status of each arc in the graph is only affected by the status of observable arcs on the same cycle: as long as no severing is observed each player i puts full mass on the history under which the entire digraph (i.e., every cycle) is maintained; if a player $i \in \iota\left(\hat{A}\right)$ observes a severing on his in-arc ki on cycle $\hat{A}$ in period t, then he puts full mass on a history under which k severed both his in-arc and his out-arc ki on cycle $\hat{A}$ in t; if a player $i \in \iota\left(\hat{A}\right)$ observes a severing on his out-arc il on cycle $\hat{A}$ in t, then he puts full mass on the history under which l severed both il and lm, l's out-arc on cycle $\hat{A}$, in t; finally notice that if both of a player's arcs with respect to a cycle have been severed then his beliefs with respect to that cycle are irrelevant under the strategy $\hat{s}$ given above. Additionally, the player puts full mass on the history under which the continuation equilibrium strategy profile is followed on unobserved arcs, in response to the observed history.

Now, let us show that the proposed $(\hat{s}, \hat{\mu})$ is an SE. We begin by verifying sequential rationality. First, notice that constructing the strategy by conditioning action choices on a given cycle only on observations on that cycle is optimal, given that each opponent does the same. If all opponents play cycles independently, then one can never improve payoffs by conditioning across cycles, and in some cases could be worse off by doing so. Second, notice that once players condition play on cycles independently, $\delta^{|A_i|}(\underline{c}-\delta)+(\bar{c}-\delta)=0$ characterizes the critical δ as of which any player with out- and in-arc on cycle A_i is willing to maintain both arcs on cycle A_i. Notice that if one is to sever one's arcs on a cycle, given $\hat{s}$, the optimal severance program involves severing one's out-arc immediately, and then severing one's in-arc in $\left|\hat{A}\right|-1$ periods. Hence, for the entire digraph D' to be maintained, it must be that $\Phi_{SE}(D)\leq 0$. Finally, notice that given that playing cycles independently is a best-response in equilibrium, that it is easy to see that $\hat{\mu}$ given above implies that $\hat{s}$ is sequentially rational.

Next, we verify that $\hat{\mu}$ are common structurally consistent; that is, for every $i\in N(D)$ $\exists\left(\hat{b}_n^i\right)_{n\geq 0}\in intB$ such that $\hat{b}_n\overset{p.w.}{\rightarrow}\hat{s}$ and for all $t\geq 0$ $\hat{\mu}_{i,n}^t\rightarrow\hat{\mu}_i^t$. Under the assumption that each player plays each cycle independently, let us define $\hat{b}_n^i$ as follows. For each period t, and each cycle A_j on which i has at least one active arc, the number of which at any period t is given by $\sum_{A_j\in\bar{A}'_D} I\left[m_i^{t,1}\left(h_i^{t,1}\right)\cap A_j\neq\varnothing\right]$, define

$$\hat{b}_n^{i,t}(A_j)=\begin{cases}\left(\frac{\eta_n^{A_j}}{\Gamma\left(\eta_n^{A_j}\right)},\frac{\left(\eta_n^{A_j}\right)^3}{\Gamma\left(\eta_n^{A_j}\right)},\frac{\left(\eta_n^{A_j}\right)^3}{\Gamma\left(\eta_n^{A_j}\right)},\frac{\left(\eta_n^{A_j}\right)^2}{\Gamma\left(\eta_n^{A_j}\right)}\right) & \begin{array}{c}\text{if } h_i^{t,1}\cap h_{A_j}^t=\left[(M,\varnothing,\varnothing)^t\right]^2,\\ \text{where } \Gamma\left(\eta_n^{A_j}\right)=\eta_n^{A_j}+\left(\eta_n^{A_j}\right)^2+2\left(\eta_n^{A_j}\right)^3\end{array}\\ \left(\eta_n^{A_j},1-\eta_n^{A_j}\right) & \begin{array}{c}\text{if } h_i^{t,1}\cap h_{A_j}^t=\left[(S,k,t-1)^t(M,\varnothing,\varnothing)^t\right]\\ \text{or } h_i^{t,1}\cap h_{A_j}^t=\left[(S,i,t-1)^t(M,\varnothing,\varnothing)^t\right]\end{array}\\ \left(1-\eta_n^{A_j},\eta_n^{A_j}\right) & \begin{array}{c}\text{if } h_i^{t,1}\cap h_{A_j}^t=\left[(M,\varnothing,\varnothing)^t,(S,l,\tau)^t\right]\\ \text{or } h_i^{t,1}\cap h_{A_j}^t=\left[(M,\varnothing,\varnothing)^t,(S,i,\tau+1)^t\right],\\ \tau>t-|A_j|-3\end{array}\\ \left(\eta_n^{A_j},1-\eta_n^{A_j}\right) & \begin{array}{c}\text{if } h_i^{t,1}\cap h_{A_j}^t=\left[(M,\varnothing,\varnothing)^t,(S,l,\tau)^t\right]\\ \text{or } h_i^{t,1}\cap h_{A_j}^t=\left[(M,\varnothing,\varnothing)^t,(S,i,\tau+1)^t\right],\\ \tau\leq t-|A_j|-3\end{array}\end{cases}\tag{37}$$

where histories are always given in terms of the status on one's in-arc and then one's out-arc on cycle A_j (where one's in-arc is ki and out-arc is il), and the behavioral strategies are given with the 4-tuple giving the probabilities of not severing either arc, severing the in-arc and not the out-arc and vice versa and severing neither arc, and the 2-tuples indicating first the probability of severing the active arc, and then the probability of not severing. It is always the case that $\eta_n^{A_j}\in(0,1)$, with $\lim_{n\rightarrow\infty}\eta_n^{A_j}=0$. It can easily be seen that the behavioral strategy probabilities on vectors of action profiles sum to one, since by independent play across cycles the behavioral strategies can be factored cycle-by-cycle, and that $\hat{b}_n\overset{p.w.}{\rightarrow}\hat{s}$. Finally, notice that we have that $\hat{\mu}_{i,n}^t\rightarrow\hat{\mu}_i^t$. The above behavioral strategies in $\hat{b}$ clearly put orders of magnitude higher probability mass on actions converging to the strategies in $\hat{s}$, and hence this generates the beliefs given in $\hat{\mu}$:

- Upon observing no severance, players believe the graph is fully maintained, since

$\Gamma\left(\eta_n^{A_j}\right) = \eta_n^{A_j} + O\left(\eta_n^{A_j}\right)$, hence $\lim_{n\to\infty} \eta_n^{A_j}/\Gamma\left(\eta_n^{A_j}\right) = 1$;

- Given observed full maintenance on a cycle, upon observing a severance on an in-arc or out-arc by one's opponent on the cycle players believe the opponent severed both arcs on that cycle, since $\left(\eta_n^{A_j}\right)^3/\Gamma\left(\eta_n^{A_j}\right) \in o\left(\left(\eta_n^{A_j}\right)^2/\Gamma\left(\eta_n^{A_j}\right)\right)$;

- Given that one arc is lost on a cycle, upon observing a severance on the remaining arc by one's opponent on the cycle players believe the opponent severed both arcs on that cycle, again since $\left(\eta_n^{A_j}\right)^3/\Gamma\left(\eta_n^{A_j}\right) \in o\left(\left(\eta_n^{A_j}\right)^2/\Gamma\left(\eta_n^{A_j}\right)\right)$;

- In all other cases players believe that the equilibrium continuation path of play is followed, though these other cases are not relevant for the player's realized choices, since with actions played independently across cycles, once both of one's arcs on a cycle are lost what happens on that cycle is irrelevant, as is what happens on cycles that a player never had arcs on; in any case, such other outcomes are never observed.

■

References

[1] Allen, Franklin and Ana Babus. 2008. "Networks in Finance." Working paper 08-07, Wharton Financial Institutions Center, University of Pennsylvania.

[2] Bala, Venkatesh and Sanjeev Goyal. 2000. "A Non-Cooperative Model of Network Formation." *Econometrica*, 2000, 68(5): 1181-1231.

[3] Ballester, Coralio, Antoni Calvo-Armengol and Yves Zenou. 2006. "Who's Who in Networks. Wanted: The Key Player". *Econometrica*, 74(5): 1403-17.

[4] Battigalli, Pierpaolo and Marciano Siniscalchi. 2002. "Strong Belief and Forward Induction Reasoning." *Journal of Economic Theory*, 106: 356-391.

[5] Bergemann, Dirk and Stephen Morris. 2007. "Belief-free Incomplete Information Games." Cowles Foundation Discussion Paper No. 1629.

[6] Bernheim, Douglas and Michael Whinston. 1990. "Multimarket Contact and Collusive Behavior." *RAND Journal of Economics*, 21(1): 1-26.

[7] Bloch, Francis. 2008. "Business Networks." In *The New Palgrave Dictionary of Economics*, ed. L. Blume and S. Durlauf. Palgrave Macmillan.

[8] Bloch, Francis, Garance Genicot and Debraj Ray. 2008. "Informal Insurance in Social Networks." *Journal of Economic Theory*, 143(1): 36-58.

[9] Bramoulle, Yann, and Rachel Kranton. 2007. "Risk-Sharing Networks." *Journal of Economic Behavior and Organization*, 64 (3-4): 275-94.

[10] Bramoulle, Yann and Rachel Kranton. Forthcoming. "Strategic Experimentation in Networks". *Journal of Economic Theory*.

[11] Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge: Harvard University Press.

[12] Burt, Ronald. 2000. "Structural Holes vs. Network Closure as Social Capital." In *Social Capital: Theory and Research*, ed. Nan Lin, Karen S. Cook, and R.S. Burt. New York: Adline de Gruyter. Pp. 31-56.

[13] Chartrand, Gary and Linda Lesniak. 2005. *Graphs and Digraphs*, 4th ed. Boca Raton, FL: Chapman & Hall/CRC.

[14] Cho, In-Koo. 1987. "A Refinement of Sequential Equilibrium." *Econometrica*, 55(6): 1367-89.

[15] Compte, Olivier. 2002. "On Sustaining Cooperation without Public Observations." *Journal of Economic Theory*, 102: 106-50.

[16] Deb, Joyee. 2008. "Cooperation and Community Responsibility: A Folk Theorem for Repeated Random Matching Games." Working paper.

[17] Easterby-Smith, Mark, Marjorie A. Lyles and Eric W.K. Tsang. 2008. "Inter-Organizational Knowledge Transfer: Current Themes and Future Prospects." *Journal of Management Studies*, 45(4): 677-90.

[18] Economides, Nicholas. 1993. "Network Economics with Applications to Finance." *Financial Markets, Instituitions and Instruments*, 2(5): 89-97.

[19] Ellison, Glenn. 1994. "Cooperation in a Prisoner's Dilemma with Anonymous Random Matching." *Review of Economic Studies*, 61: 567-88.

[20] Ely, Jeffrey, Johannes Hörner and Wojciech Olszewski. 2005. "Belief-free Equilibria in Repeated Games." *Econometrica*, 73: 377-415.

[21] Fafchamps, Marcel and Flore Gubert. 2007. "The Formation of Risk-Sharing Networks." *Journal of Development Economics*, 83(2): 326-50.

[22] Galeotti, Andrea. 2006. "One-Way Flow Networks: The Role of Heterogeneity". *Economic Theory*, 29(1): 163-79.

[23] Galeotti, Andrea, Sanjeev Goyal, Matthew Jackson, Fernando Vega-Redondo, and Leeat Yariv. Forthcoming. "Network games". *Review of Economic Studies*.

[24] Ghosh, Parikshit and Debraj Ray. 1996. "Cooperation in Community Interaction without Information Flows." *Review of Economic Studies*, 63: 491-519.

[25] Goyal, Sanjeev. 2005. "Strong and Weak Links." *Journal of the European Economic Assocation*, 3: 608-616.

[26] Goyal, Sanjeev. 2007. *Connections: An Introduction to the Economics of Networks*. Princeton, NJ: Princeton University Press.

[27] Goyal, Sanjeev, Alexander Konovalov and Jose Luis Moraga-Gonzalez. 2008. "Hybrid R&D." *Journal of the European Economic Assocation*, 6(6): 1309-1338.

[28] Goyal, Sanjeev, and Jose Luis Moraga-Gonzalez. 2001. "R&D Networks." *RAND Journal of Economics*, 32: 686-707.

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

[30] Gulati, Ranjay. 1998. "Alliances and Networks." *Strategic Management Journal*, 19(4): 293-317.

[31] Haag, Matthew and Roger Lagunoff. 2006. "Social Norms, Local Interaction, and Neighborhood Planning." *International Economic Review*, 47(1): 265-96.

[32] Hagedoorn, John. 2002. "Inter-firm R&D Partnerships: An Overview of Major Trends and Patterns Since 1960." *Research Policy*, 31: 477-92.

[33] Hörner, Johannes and Wojciech Olszewski. 2006. "The Folk Theorem for Games with Private Almost-Perfect Monitoring." *Econometrica*, 74(6): 1499-1544.

[34] Jackson, Matthew. 2005. "A Survey of Models of Network Formation: Stability and Efficiency." In *Group Formation in Economics: Networks, Clubs and Coalitions*, ed. by Gabrielle Demange and Myrna Wooders. Cambridge: Cambridge University Press.

[35] Jackson, Matthew. 2008. *Social and Economic Networks*. Princeton University Press.

[36] Jarillo, J. Carlos. 1988. "On Strategic Networks." *Strategic Management Journal*, 9(1): 31-41.

[37] Kandori, Michihiro. 1992. "Social Norms and Community Enforcement." *Review of Economic Studies*, 59: 63-80.

[38] Kandori, Michihiro. 2002. "Introduction to Repeated Games with Private Monitoring." *Journal of Economic Theory*, 102(1): 1-15.

[39] Knoke, David. 1990. *Political Networks: The Structural Perspective*. Cambridge: Cambridge University Press.

[40] Kossinets, Gueorgi and Duncan Watts. 2006. "Empirical Analysis of an Evolving Social Network." *Science*, 311(5757): 88-90.

[41] Kranton, Rachel. 1996. "Reciprocal Exchange: A Self-Sustaining System." *American Economic Review*, 86(4): 830-51.

[42] Kranton, Rachel and Deborah Minehart. 2001. "A Theory of Buyer-Seller Networks." *American Economic Review*, 91(3): 485-508.

[43] Kreps, David and Robert Wilson. 1982. "Sequential Equilibrium." *Econometrica*, 50: 863-94.

[44] Lamoreaux, Naomi R. 2003. "Rethinking the Transition to Capitalism in the Early American Northeast." *The Journal of American History*, 90(2): 437-61.

[45] Mayer, Colin. Sept. 2008. "Trust in Financial Markets." *European Financial Management*, 14(4): 617-32.

[46] McBride, Michael. 2006. "Imperfect Monitoring in Communication Networks." *Journal of Economic Theory*, 126: 97-119.

[47] Mihm, Maximilian A., Russell D. Toth and Corey Lang. 2009. "What Goes Around Comes Around: A theory of indirect reciprocity in networks (extended version)", available at http://www.arts.cornell.edu/econ/cae/09-07.pdf.

[48] Mowery, David C., Joanne E. Oxley and Brian S. Silverman. 1996. "Strategic Alliances and Interfirm Knowledge Transfer." *Strategic Management Journal*, 17: 77-91.

[49] Nowak, Martin A. and Karl Sigmund. 2005. "Evolution of Indirect Reciprocity." *Nature*, 437: 1291-98.

[50] Olszewski, Wojciech. 2007. "A Simple Exposition of Belief-Free Equilibria in Repeated Games." *Economics Bulletin*, 58(3): 1-16.

[51] Raub, Werner and Jeroen Weesie. 1990. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology*, 96(3): 626-54.

[52] Rauch, James. 1999. "Networks Versus Markets in International Trade." *Journal of International Economics*, 48(1):7-35.

[53] Sammarra, Alessia and Lucio Biggiero. 2008. "Heterogeneity and Specificity of Inter-Firm Knowledge Flows in Innovation Networks." *Journal of Management Studies*, 45(4), 800-829.

[54] Thorelli, H.B. 1986. "Networks: Between Markets and Hierarches." *Strategic Management Journal*, 7: 37-51.

[55] United States of America v. Airline Tariff Publishing Company et al., 92-2854 F. Supp. 2d (D.C. Cir. 1992).

[56] Vega-Redondo, Fernando, Matteo Marsili and Frantisek Slanina. 2005. "Clustering, Cooperation and Search in Social Networks." *Journal of the European Economic Association*, 3(2-3): 628-38.

[57] Vega-Redondo, Fernando. 2006. "Building Up Social Capital in a Changing World." *Journal of Economic Dynamics and Control*, 30.

[58] Vicente, Jerome, Pierre-Alexandre Balland and Olivier Brossard. 2008. "Getting into Networks and Clusters: Evidence on the GNSS Composite Knowledge Process in (and from) Midi-Pyrenees." Utrecht University Papers in Evolutionary and Economic Geography.

[59] Watts, Duncan and Steven Strogatz. 1998. "Collective Dynamics of "Small World" Networks." *Nature*, 393: 440-42.

[60] WTO. “The Trade Policy Review Mechanism (TPRM)”. World Trade Organization. 16 Mar. 2009 <http://www.liu.edu/cwis/cwp/library/workshop/citmla.htm>.