Cooperative Networks with Robust Private Monitoring*

Maximilian Mihm† Russell Toth‡

July 31, 2019

Abstract

Social networks support cooperative behavior in a variety of social and economic settings. We study cooperative networks that can be formed when payoff and information asymmetries imply that cooperation relies on unverifiable third-party punishments. Under private monitoring, equilibrium predictions can depend on players' beliefs about unobservable behavior, and our results identify the range of equilibrium network outcomes. In particular, when strategies must be robust to beliefs, we show that network structures satisfy a triadic closure property, providing a strategic rationale for the short paths and high clustering observed in many social and economic networks. We illustrate some of the substantive restrictions identified in our results for risk- and information-sharing networks.

Key words: Belief-free equilibrium; local monitoring; networks; triadic closure.

1 Introduction

Individuals cooperate in a range of social and economic settings even without formal institutions to enforce cooperative behavior. The sharing of information between innovative firms, the exchange of favors among colleagues, or the informal risk-sharing arrangements in developing economies, are examples of cooperative behavior that rely on informal enforcement of cooperative norms. While cooperation can sometimes be sustained bilaterally through infinitely repeated interaction, when there is heterogeneity in preferences, information, or enforcement opportunities, it can be more efficient to pool incentive constraints across groups, communities, or markets

*We thank Chris Barrett, Kaushik Basu, Larry Blume, Adam Brandenburger, David Easley, Ani Guerdjikova, Johannes Hoerner, Willemien Kets, Corey Lang, Debraj Ray and Fernando Vega-Redondo for useful discussions. Toth gratefully acknowledges support from the NSF Expeditions in Computing grant on Computational Sustainability, #0832782.

†New York University Abu Dhabi. Email: max.mihm@nyu.edu.

‡University of Sydney. Email: russell.toth@sydney.edu.au.

(Bernheim and Whinston [1990]). Community enforcement leverages contagion strategies, triggering a spread of non-cooperative behavior following deviations (Kandori [1992], Ellison [1994]). A recent strand of research seeks to formalize the structure of community interactions in terms of networks (Jackson [2009], Goyal [2012]), enhancing the empirical relevance of the theory by generating predictions about observable network structure.

In this paper, we develop a theory of community enforcement that generates predictions about networks that support cooperative behavior under private monitoring. We model network formation and continuation play in a unified framework, with a simple trade-off between investments in the formation of ties and the ongoing benefits of cooperation. In the formation phase, players can make heterogeneous investments in bilateral ties. In the continuation phase, players have the opportunity to cooperate on their bilateral ties, receiving a stream of cooperative payoffs, or permanently ending a bilateral interaction to receive a one-off premium. In a given period, the stage-game on a bilateral tie therefore has the structure of a Prisoner's Dilemma, providing a stylized model of the tension between individual incentives and social efficiency. Investment in a tie during the formation phase has an inverse relationship to the subsequent cost of cooperation, reflected in a "formation technology" that can be different in each bilateral relationship. This generates a trade-off between the effort cost in forming ties, and the effort cost required to cooperate. For a range of parameters, this trade-off can imply that strategically isolated bilateral ties are never incentive compatible, and cooperative behavior can emerge only when ties are embedded in a network that pools incentive constraints.

With perfect monitoring virtually any network structure can emerge under the right configuration of parameters, even when strategically isolated ties are not incentive compatible. In particular, full cooperation can be supported by equilibria that dictate defection on every active relationship for any observed deviation, thereby pooling incentive constraints across the network. However, in a range of network applications, such rich monitoring structures seem implausible. For example, in settings such as rural villages in developing countries, farmers may have a disincentive to broadly disclose their history of wealth shocks. In a network of innovative firms, firms may be reluctant to disclose the content of their relationships in order to protect trade and technological secrets. Our primary analysis therefore focuses on an environment with private monitoring. We introduce monitoring restrictions via a *radius of information* ρ, which specifies the longest network path on which a player can observe actions. When $\rho = 1$, players can only observe actions on their own ties; when $\rho = 2$, players can also observe actions on their neighbor's ties; when $\rho = 3$, players can observe their neighbors' neighbors' ties, and so on.

With private monitoring expected payoffs depend on players' *beliefs* about actions outside their radius of information. As a result, equilibrium predictions can be sensitive to the choice of equilibrium refinement, which imposes structure on beliefs. Our approach to this problem is to provide upper and lower bounds on the set of equilibrium network outcomes by studying arguably the weakest and strongest equilibrium refinements that accommodate monitoring restrictions. To provide an upper bound, we consider Perfect Bayesian Equilibrium, which requires only that, at each information set, every player best-responds given *some* system of beliefs that is consistent with the equilibrium strategy profile. Perfect Bayesian Equilibrium (PBE) is therefore a weak equilibrium refinement that ensures sequential rationality given Bayesian beliefs. To provide a lower bound, we consider Belief-Free Equilibrium (Piccione [2002], Ely and Välimäki [2002], Ely et al. [2005], Hörner and Lovo [2009], Hörner et al. [2011]), which requires that play is sequentially rational for *all* beliefs that are consistent with the equilibrium strategy. Belief-Free Equilibrium (BFE) is therefore a strong refinement that requires strategic network interactions to be robust to all off-path beliefs.

Our first result shows that, for any radius of information $\rho \geq 2$, the same set of network outcomes can emerge under both PBE and BFE. In particular, regardless of the refinements imposed on off-path beliefs, private monitoring does not restrict the set of equilibrium networks relative to perfect information because, when a player deviates against one neighbor, all other neighbors can observe this deviation and punish immediately.

This strong equivalence breaks down when $\rho = 1$, which we call *local monitoring*. With local monitoring, players can deviate from cooperation on some ties without other neighbors observing this deviation. As a result, the network not only pools incentive constraints, but also plays a crucial role in the transmission of information. Players cannot monitor directly and so require contagion strategies that spread through the network to enforce cooperation. Since contagion takes time, the effectiveness of third-party punishment can depend on the distance between nodes, and tighter networks (with shorter paths between nodes) are needed to make cooperation incentive compatible. In such environments, we first characterize the network outcomes that can emerge in a PBE, and show how local monitoring restricts observable network outcomes relative to perfect monitoring by restricting network paths. Our characterization of PBE network outcomes thereby extends previous literature (e.g., Raub and Weesie [1990], Lippert and Spagnolo [2011]) to a framework where incentives in the formation of ties are in tension with cooperative behavior on the resulting network.

Requiring belief-free best-responses generates much sharper predictions on network outcomes. In a BFE, we show that network ties must satisfy a "triadic closure" property, which requires that, in many cases, two neighbors of player i also need to be direct neighbors with each other. Triadic closure is both necessary and sufficient for belief-free enforcement mechanisms because it significantly reduces the time taken for contagion to spread through the network. At the network level, triadic closure implies a "clustering" of network relationships that is commonly observed in real-world networks. Our model provides a strategic explanation for such patterns.

Overall, while BFE may be overly decisive in refining the set of equilibria in some applications, our results provide a useful boundary in distinguishing the network structures that can be rationalized under different assumptions about the extent of belief coordination.

Finally, our results address the relationship between heterogeneity in the formation technology, network structure and cooperative outcomes. In a setting with a homogeneous formation technology, a network can only facilitate cooperation when there are non-convexities in the technology. We provide a simple example of such non-convexities in Section 4.1, where a network of cooperative ties can emerge in equilibrium, and all heterogeneity in network ties are due to differential investments. When there are asymmetries in the formation technology, non-trivial networks can emerge in equilibrium even for convex formation technologies. We provide two such examples in Sections 4.2 and 4.3, where asymmetries in the formation technology seem natural, and relate them to applications in informal risk-sharing and information-sharing networks.

Our paper contributes to the literature by providing a novel economic rationale for the "small-world" property, which has been observed in a number of social and economic networks (Milgram [1967], Watts [1999]). Small-world networks have two main characteristics: (1) small diameter and average path lengths, and (2) high clustering relative to networks generated by an independent random process (Jackson and Rogers [2005]). There is a long-standing interest in rationalizing such network structures: while an early literature uses mechanical network formation processes to generate small-worlds (see, e.g., Jackson and Rogers [2007]), there has been more recent interest in providing economic explanations. Jackson and Rogers [2005] provide a simple formation model that generates small-worlds based on heterogeneity in costs and benefits of forming ties, without modeling continuation play on the network explicitly. Jackson et al. [2012] show that, under perfect information with infrequent bilateral interactions, favor exchange is supported by renegotiation-proof equilibria through high clustering of relationships in "social quilts," which generate a localized penalty to those who deviate. The clustering concept ("supported") is a homogenous relationships analogue of our version of triadic closure. Our model provides an alternative mechanism to rationalize the clustering observed by Jackson et al. [2012] based on the robust enforcement of cooperative norms under local monitoring.

Our paper also contributes to a literature on network enforcement under imperfect information, where players can choose heterogeneous actions on each tie.[1] Raub and Weesie [1990] model the intuition that tightly-connected networks shorten contagion time, without making specific predictions about network structure. Lippert and Spagnolo [2011] study a fixed network of asymmetric, bilateral Prisoner's Dilemma games, and consider the role of key nodes in sustaining cooperation with limited Word-of-Mouth communication. Ali and Miller [2013] study

[1]A complementary literature studies repeated games on networks in which each player takes a single action affecting all partners, e.g., Haag and Lagunoff [2006], Wolitzky [2013], Nava and Piccione [2014].

the role of local monitoring in a setting with exponentially-distributed matching, where each player is limited to at most d partnerships, and show that the socially optimal network is composed of disconnected cliques of $d + 1$ players. Balmaceda and Escobar [2017] study an optimal network design problem, and characterize networks under different assumptions about the flow of information through the network. Under limited information, cohesive networks tend to be efficient as they facilitate coordination and, similarly to Ali and Miller [2013], optimal networks have cliques of equally-sized sub-components.

We contribute to this literature in at least two ways. First, we do not start with an exogenous network, but explicitly model the formation and continuation of heterogeneous ties in a unified framework.[2] Our model therefore generates predictions about a different observable. In models with a fixed network, equilibrium predictions address the degree of cooperative behavior, which is often difficult to measure in practice. Our results make predictions directly about network structure, which has been a growing topic of empirical research (see, e.g., Newman et al. [2006]; Jackson [2009]; Latora et al. [2017]). Extending the analysis to incorporate network formation explicitly highlights potential trade-offs between incentives in forming versus cooperating on ties. As we show, these trade-offs generate an endogenous source of tie heterogeneity, which in turn affects the role of the network in pooling incentive constraints. Second, we identify the range of equilibrium network outcomes in environments with local monitoring. Our PBE characterization provides arguably the weakest conditions that an equilibrium network should satisfy, while our results for BFE identify a very robust class of equilibrium network outcomes. The BFE result also generates new predictions on network structure in a local monitoring environment. The prior literature has often used normative criteria (i.e., efficiency) to analyze a subset of the PBE (e.g., Ali and Miller [2013], Balmaceda and Escobar [2017]). Our result for BFE similarly highlights the importance of clustering and density for cooperative outcomes, but results from an entirely different mechanism that requires robust punishment strategies under local monitoring. This leads to a distinct structural prediction, which is more permissible than requiring completely-connected cliques or the complete network. Thus, our results provide a rationale for a network property (triadic closure) that has received long-standing attention in the broader networks literature (Simmel [1908]; Granovetter [1973]; Watts [1999]).

The remainder of the paper is organized as follows. In Section 2, we present our framework: a dynamic network game where players form heterogeneous ties and then decide whether to cooperate on ties in subsequent periods. Section 3 presents our main results. First, we consider the efficient outcomes that would emerge if players could perfectly commit to cooperate. We

[2] In Ali and Miller [2013], the network of cooperation gains can evolve over time because players can, in each period, choose the stakes at which to cooperate. However, the underlying structure that determines which players can interact, and which interactions a player can monitor is exogenous.

then assume that players are unable to commit to cooperate, but are able to partially monitor the behavior of other players. Finally, our main analysis considers network outcomes that can emerge in PBE and BFE under local monitoring. In Section 4, we illustrate some of the substantive restrictions of our results in the context of a risk-sharing and information-sharing examples. Section 5 concludes. Proofs are given in a separate Appendix.

2 Model

There is a finite set of players $\mathcal{I} = \{1, ..., I\}$, time is indexed $t = 0, 1, 2, ...$, and players have a common discount factor $\delta \in (0, 1)$. The model has two phases. In period $t = 0$, players form bilateral ties in a *formation phase*. After ties are formed, in periods $t \geq 0$, players decide whether to cooperate on their ties in a *continuation phase*.

The ties between players can be viewed as a *network*, where *nodes* are players and *ties* are bilateral relationships of varying *strength*. Tie strength is determined by investments during the formation phase, and indicates how costly it is for players to cooperate in the continuation phase. A tie between nodes i and j is indexed by $(c_{ij}, c_{ji}) \in [0, 1]^2$, where $c_{ij} = c_{ji} = 1$ indicates that there is no tie between i and j. When $c_{ij}c_{ji} < 1$, c_{ij} is the cost that i incurs to cooperate with j in the continuation phase, and so a lower c_{ij} corresponds to a stronger tie for player i. A network is an $I \times I$ matrix $c \equiv (c_{ij})_{i\in\mathcal{I},j\in\mathcal{I}} \in [0, 1]^{I\times I}$, which satisfies (i) $c_{ii} = 1$, and (ii) $c_{ij} = 1$ if and only if $c_{ji} = 1$ (i.e., if player i has a tie with player j, then player j also has a tie with player i). When $c_{ij} < 1$ we say that i and j are partners in network c and write $ij \in c$.

2.1 Formation phase

In the formation phase, players decide how much to invest in ties. Investments are made unilaterally and simultaneously in period $t = 0$. Player i can either choose not to form a tie with j, or choose to initiate a tie with a positive investment from $\left[\underline{F}, \bar{F}\right]$, where $0 \leq \underline{F} < 1 < \overline{F}$. A tie is formed by players i and j if and only if $F_{ij}, F_{ji} \neq \emptyset$. We denote player i's investment in a tie with player j by $F_{ij} \in \mathcal{F} \equiv \{\emptyset\} \cup \left[\underline{F}, \overline{F}\right]$, where $\emptyset$ corresponds to not forming a tie.[3] The cost of not forming a tie is 0, and we require that $F_{ii} = \emptyset$.[4]

[3]With some abuse of notation, we write $F_{ij} = 0$ when i does not initiate a tie with j, since the investment cost is 0 in this case. However, this action is distinct from $F_{ij} = \underline{F}$, even when $\underline{F} = 0$. When $\underline{F} > 0$, both partners must make a strictly positive invest to form a tie, which imposes a bilateral formation constraint.

[4]It is not essential that formation costs are incurred in period 0, but rather that players commit upfront to invest resources in a tie, whether or not cooperation actually occurs. For instance, microfinance institutions often rely on network interactions to enforce cooperative behavior, especially when it is not possible to contractually enforce cooperation. Typically participants need to commit to meet regularly with their partners, and the frequency of meetings (and penalties for not meeting) could be interpreted as formation costs, which determine the opportunities for cooperation on ties, but are distinct from the cooperative behavior that ensues after tie

Investments determine whether a tie is formed and the strength of the tie. The mapping from investment costs to tie-strengths is described by a *formation technology*: an $I \times I$ matrix of functions $\Gamma \equiv (\gamma_{ij})_{i,j\in\mathcal{I}}$, such that $\gamma_{ij} : \left[\underline{F}, \overline{F}\right] \to (0,1)$ is monotone non-increasing and lower-semicontinuous. The formation technology Γ has the following interpretation. Suppose players i and j invest $F_{ij}, F_{ji} \neq \emptyset$ in forming a tie. Then tie ij is formed, the strength of the tie from i's point of view is $\gamma_{ij}(F_{ij})$, and the strength from j's point of view is $\gamma_{ji}(F_{ji})$. For a given Γ, a profile of investments $F = (F_{ij})_{i,j\in\mathcal{I}}$ induces a network $c(F) \equiv (c_{ij}(F))_{i,j\in\mathcal{I}}$, defined by

$$c_{ij}(F) \equiv \begin{cases} \gamma_{ij}(F_{ij}) & \text{if } F_{ij}, F_{ji} \neq \emptyset \\ 1 & \text{otherwise} \end{cases}. \tag{1}$$

Example 1 (Formation technology)**.** (i) In a *homogeneous formation* model, the formation technology does not depend on the identity of either partner, and any variation in tie-strengths is attributable to differences in formation investments. This model is the special case of Γ where $\gamma_{i,j} \equiv \gamma$ for some $\gamma : \left[\underline{F}, \overline{F}\right] \to (0,1)$. (ii) In a *within-between group* model, the population is divided into groups and the formation technology depends on whether a tie is formed by members within a group, or members of different groups. This model is the special case of Γ where there is an equivalence relation $\approx$ on the set of players $\mathcal{I}$ and two functions $\gamma_W, \gamma_B : \left[\underline{F}, \overline{F}\right] \to (0,1)$ such that $\gamma_{ij} = \gamma_W$ if $i \approx j$ and $\gamma_{ij} = \gamma_B$ otherwise. (iii) In a *target-partner* model, the population is also divided into groups, but the formation technology depends on the group-membership of the partner. This model is the special case of Γ where there is a partition $(\mathcal{I}_1, ..., \mathcal{I}_n)$ of $\mathcal{I}$, and n functions $\gamma_k : \left[\underline{F}, \overline{F}\right] \to (0,1)$ for $k = 1, ..., n$, such that $\gamma_{ij} = \gamma_k$ if $j \in \mathcal{I}_k$. (iv) In a *constant formation* model, investments in the formation phase do not affect cooperation costs in the continuation phase. This model is the the special case of Γ where, for some exogenously given network c, $\gamma_{ij}(f) = c$ for all $f \in [\underline{F}, \bar{F}]$. When $\underline{F} = 0$, it is costless for players to form the network c and impossible for them to form any other network; our framework thereby encompasses a model with an exogenously given network, which is the special case with no trade-offs between the formation of ties and benefits of future cooperation.[5] □

formation. In a recent study, Feigenberg et al. [2013] exogenously vary the frequency with which participants in a microfinance program were required to meet (weekly or monthly), and measure the degree of cooperative behavior using measures from trust games, willingness to share real wealth, and performance on microfinance loans. They find that increasing the strength of ties by requiring more frequent meetings encourages greater cooperation, consistent with the idea that higher formation investments increases cooperation incentives. Similarly, in high-tech industries, firms often form strategic alliances that commit resources (e.g., staff, facilities, and even profit-sharing) to promote cooperative behavior with other firms (see, e.g., Stuart [1998]; Eisenhardt and Schoonhoven [1996]). The resources committed in the formation of such strategic alliances, contractual asset pooling or resource exchange agreements can be interpreted as formation costs, which impact the benefit-cost calculus of cooperation but are committed whether or not useful information-sharing occurs.

[5]For the constant formation model with $\underline{F} = 0$, the incentive conditions in Section 3 simplify because, in

We define a subclass of formation strategies that minimize the cost of forming a network. In particular, for network c, define the formation strategy-profile F^c as follows: for all $i, j \in \mathcal{I}$,

$$F_{ij}^c \equiv \begin{cases} \min\left\{f \in \left[\underline{F}, \bar{F}\right] \,:\, \gamma_{ij}(f) = c_{ij}\right\} & \text{if } ij \in c \\ \emptyset & \text{if } ij \notin c \end{cases}. \tag{2}$$

The formation profile F^c is well-defined because γ_{ij} is lower semi-continuous. Moreover, $c\left(F^c\right) = c$ by definition, and F^c minimizes the formation costs to form network c. In our equilibrium analysis, it is without loss of generality to focus on formation strategies that minimize formation costs because any cooperative network that can be supported in equilibrium for an arbitrary formation strategy can also be supported by the cost-minimizing formation strategy.

2.2 Continuation phase

In any period of the continuation phase $t \geq 0$, there is a current network c^t, and each player has an opportunity to cooperate (e.g., share information, exchange favors, insure risks) with their partners. We denote by $A_{ij}^t \in \{C, D\}$ the action of player i in the tie with player j, where C denotes that i *cooperates* and D denotes that i *doesn't cooperate*. Cooperation decisions are unilateral, simultaneous, and independent across ties.

When player i cooperates with j, player j receives a *cooperation benefit* normalized to 1 and i incurs *cooperation cost* c_{ij}. The cooperation cost therefore links the formation and continuation phase of the game, because it is less costly for i to cooperate the stronger the tie between i and j. When i cooperates with j, and j does not reciprocate, then i incurs the cooperation cost without receiving the benefit, and j receives the benefit without incurring the cost. The one period interaction between i and j therefore has the structure of a Prisoner's Dilemma, where non-cooperation is strictly dominant but mutual cooperation is Pareto improving.

Given a network c^t, when players i and j both cooperate their tie continues to period $t+1$; otherwise the tie is discontinued. The actions C and D therefore also denote *continue* and *discontinue*. We follow Jackson et al. [2012] in assuming that ties are discontinued after a non-cooperative action. This assumption simplifies the analysis by facilitating backward-induction arguments. The sufficient conditions we provide for existence of equilibrium would also be sufficient if we modeled bilateral ties as infinitely repeated Prisoner's Dilemma games, where cooperation can be restored after a defection. However, we view it as more significant to determine potential restrictions on network structure and, for this, we require the assumption

equilibrium, $F_{ij} = \underline{F} = 0$ for all $i, j \in \mathcal{I}$. Our results then provide predictions about the exogenously given networks on which full cooperation can be sustained under different monitoring restrictions and assumptions about off-path beliefs.

that ties are discontinued permanently following a defection.[6] A decision by each player about whether to cooperate with each partner determines a new network c^{t+1}, which is a subset of the network in period t, and where players again decide whether to cooperate on remaining ties.

2.3 Bilateral incentive constraints

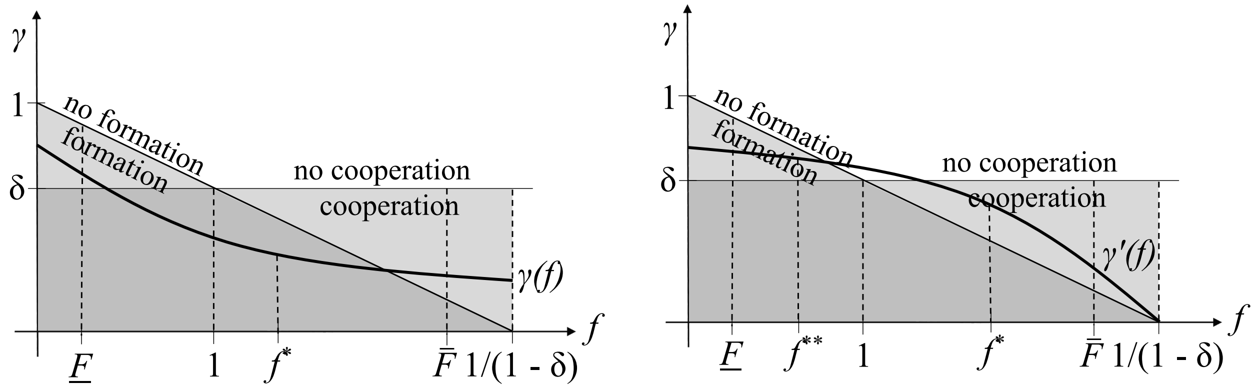
We are primarily interested in ties that can be formed only when they are embedded in a network of other ties. To illustrate, first consider a strategically isolated bilateral tie $c = (c_{ij}, c_{ji})$ between players i and j. To continue cooperating on this tie, player i must believe that j will cooperate. Moreover, the discounted value of cooperation must exceed the one-off benefit of defection. Hence, a *bilateral cooperation constraint* $\left(\frac{1-c_{ij}}{1-\delta}\right) \geq 1$ must be satisfied. In addition, in order for player i to form the tie in the first place, the discounted value of cooperation must exceed the formation cost, and so a *bilateral formation constraint* $\left(\frac{1-c_{ij}}{1-\delta}\right) \geq F^c_{ij}$ must also be satisfied. Combining cooperation and formation constraints, a strategically isolated tie $c = (c_{ij}, c_{ji})$ must satisfy the following *bilateral incentive constraint (BIC)*:

$$c_{ij} \leq \min\left\{\delta, 1 - (1-\delta)F^c_{ij}\right\}. \tag{3}$$

Whether i and j can form a strategically isolated bilateral tie therefore depends on the formation technology $(\gamma_{ij}, \gamma_{ji})$. In particular, it is not possible to form a bilateral tie if $\gamma_{ij}(f) > \min\{\delta, 1-(1-\delta)f\}$ for all $f \in \left[\underline{F}, \overline{F}\right]$.

Since γ_{ij} is decreasing, there is a trade-off between investments in the formation of ties and future cooperation incentives. Figure 1(a) depicts a technology where, despite the trade-off, the BIC can be satisfied: at the investment level f^* both the formation constraint and the continuation constraint are satisfied simultaneously. However, Figure 1(b) depicts a technology where the BIC cannot be satisfied for any investment $f \in [\underline{F}, \bar{F}]$. With investment f^*, the bilateral cooperation constraint is satisfied because cooperation costs $\gamma(f^*)$ are sufficiently low, but the bilateral formation constraint is not satisfied because the formation cost is too high. To meet the formation constraint, investments would need to be lower, but this increases cooperation costs so that the cooperation constraint is no longer satisfied (e.g., at the lower investment f^{**}). However, even when the BIC cannot be satisfied, it may be possible for a network to "pool" incentive constraints across ties.

[6] In the literature on community enforcement (e.g., Kandori [1992]; Ellison [1994]), Folk-theorems often require that deviations are punished for some time, and then forgiven. Such strategies generally require players to have information about their opponents' past actions. In our context, players never learn past actions outside their radius of information. In principle, to overcome information asymmetries, players could use patterns of partial cooperation to communicate about the structure of the network. Inference problems for such strategies would be highly complex, and we have not been able to make progress on how and when they could arise in equilibrium.



(a) BIC can be satisfied

(b) BIC cannot be satisfied

Figure 1: Bilateral incentive constraints

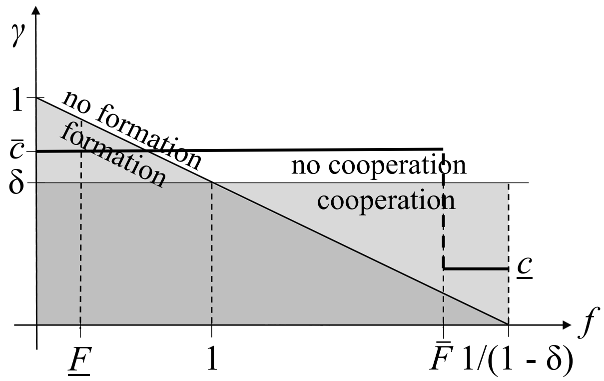
Example 2 (Binary formation model)**.** Suppose players face a binary formation choice $f \in \{\underline{F}, \bar{F}\}$, where the low investment $\underline{F}$ leads to a high cooperation cost $\bar{c}$, and the high investment $\bar{F}$ leads to a low cooperation cost $\underline{c}$.[7] We use this model as an on-going example, with the following restriction on parameters:

$$1-(1-\delta)\bar{F} < \underline{c} < \delta < \bar{c} < 1-(1-\delta)\underline{F}. \tag{4}$$

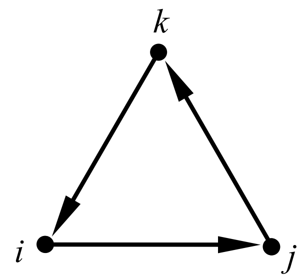
Condition (4) illustrates the trade-off between formation and cooperation incentives on bilateral ties. The inequality $\underline{c} < \delta$ is equivalent to $1 < \frac{1-\underline{c}}{1-\delta}$, i.e., the one-off payoff from defecting on a tie with low cooperation cost $\underline{c}$ is less than discounted benefit of infinite cooperation. As such, for a player who incurs the high formation cost $\bar{F}$, cooperation is incentive compatible. However, the inequality $1-(1-\bar{F}) < \underline{c}$ is equivalent to $\frac{1-\underline{c}}{1-\delta} - \bar{F} < 0$, i.e., the discounted benefit of infinite cooperation net of the high formation cost is less than the cost of not forming the tie. As such, it is not incentive compatible to form a bilateral tie with the with $\bar{F}$, even though it would be incentive compatible to cooperate on the tie after it is formed. Similarly, the inequalities $\delta < \bar{c} < 1-(1-\delta)\underline{F}$ imply that (i) it is not incentive compatible to cooperate on a tie with the high cooperation cost $\bar{c}$ (because $\delta < \bar{c}$ is equivalent to $\frac{1-\bar{c}}{1-\delta} < 1$), but (ii) it would be incentive compatible to form a tie with low formation cost $\underline{F}$ if cooperation could be sustained on that tie (because $\bar{c} < 1-(1-\delta)\underline{F}$ is equivalent to $0 < \frac{1-\bar{c}}{1-\delta} - \underline{F}$).

Condition (4) therefore implies that the BIC cannot be satisfied. However, it is possible for a network to pool incentive constraint. To illustrate, consider a setting with three players $\mathcal{I} = \{i, j, k\}$, $\delta = \frac{8}{10}$, $\bar{F} = \frac{16}{10}$, $\underline{F} = \frac{4}{10}$, $\gamma\left(\overline{F}\right) = \frac{7}{10}$, and $\gamma(\underline{F}) = \frac{9}{10}$, which satisfies Condition (4).

[7] This binary formation model is strategically equivalent to a homogenous formation model (Example 1) where $\gamma_{ij} = \gamma$ for all players i and j and, for some $\underline{c} < \bar{c}$, $\gamma(\bar{F}) = \underline{c}$ and $\gamma(f) = \bar{c}$ for $f \in [\underline{F}, \bar{F})$, depicted in Figure 2(a).



(a) Binary formation technology



(b) Triangle network

Figure 2: Technology and network in Example 2

Figure 2(b) illustrates the triangle network in Example 2. Dots represent nodes (or players); there is a line connecting nodes iff both players form a tie; arrows represent where a player makes the high formation investment. The network in Figure 2(b) therefore consists of three nodes, there is a tie between each pair of nodes, and all ties are asymmetric. For instance, the arrow pointing from i to j indicates that player $i$makes the high formation investment on tie ij, but j makes the low investment.

Suppose there is perfect monitoring and consider the following strategies: in the formation phase, $F_{ij} = F_{jk} = F_{ki} = \bar{F}$ and $F_{ik} = F_{kj} = F_{ji} = \underline{F}$; in the continuation phase, players cooperate on ties as long as all other players cooperate, and otherwise cease cooperating on all remaining ties. This strategy profile induces network c in Figure 2(b), on which players cooperate indefinitely. Moreover, the strategy profile is a subgame perfect equilibrium. In the continuation phase, the discounted value of cooperation for player i is $\frac{1-\underline{c}}{1-\delta} + \frac{1-\bar{c}}{1-\delta} = 2$. The optimal deviation would be to discontinue both ties simultaneously, also yielding a payoff of 2. In the formation phase, the cost of forming ties is $\underline{F} + \bar{F} = 2$, so the net benefit of forming the ties is 0. The optimal deviation in the formation phase would be to form no ties, also yielding a payoff of 0. Hence, there is an equilibrium in which a cooperative network is formed.[8] $\square$

2.4 Network game

We study networks that can be formed in equilibrium, and where players cooperate indefinitely. We view such networks as the long-term stable and observable outcomes of the game. To study equilibrium network outcomes, we focus on pure strategies because, given that a defection discontinues a tie, mixing makes the network unstable. In other words, a network is formed and continued indefinitely only if, on the equilibrium path, players cooperate with probability

[8] If $\bar{F} = \frac{16}{10} - \varepsilon$ and $\underline{F} = \frac{4}{10} - \varepsilon$ for sufficiently small $\varepsilon > 0$, the BIC still cannot be satisfied. However, the cooperative network c remains a subgame perfect equilibrium, and now yields strictly positive net payoffs.

1. Below, we provide an overview of the dynamic game and solution concepts, with additional details given in Appendix A.1.

In every period, a *history* is a record of all actions that have been taken in the preceding periods. We introduce monitoring restrictions via a *radius of information* $\rho \in \mathbb{N} \cup \{\infty\}$, which specifies the longest network path on which a player can observe actions. When $\rho = 1$, players can only observe actions on their own ties; when $\rho = 2$, players can observe actions also on their neighbor's ties; when $\rho = 3$, players can observe their neighbors' neighbors' ties, and so on. When $\rho = \infty$, players can observe the entire history and the game therefore has perfect information. We refer to $\rho = 1$ as *local monitoring*; and cases where $\rho \geq 2$ as *partial monitoring.*

In the formation phase, player i chooses an investment $F_{ij} \in \left[\underline{F}, \bar{F}\right]$ for each potential partner $j \in \mathcal{I} - \{i\}$, and the profile of investment decisions $F = (F_{ij})$ determines a network. In the continuation phase, a strategy of player i, denoted σ_i, is mapping from the current history that player i can observe (given the radius of information) to a cooperation decision on each of her remaining ties. Recursively, a strategy (F, σ) thereby induces a unique a history in every period, which in turn specifies a network of ties that were formed during the formation phase, and then continued in every period of the continuation phase. We write $c(F, \sigma) = c$ if the strategy profile (F, σ) induces network c in every period.

When players cannot observe the entire history, expected payoffs depend on players' *beliefs* about actions outside their radius of information. As a minimal requirement, such beliefs should be *Bayesian consistent* with an equilibrium strategy. Bayesian consistency with a strategy profile (F, σ) requires player i to believe (with probability 1) that other players have followed strategy (F_{-i}, σ_{-i}) unless i observes a history that is inconsistent with (F_{-i}, σ_{-i}).

We focus on two equilibrium concepts that incorporate Bayesian consistent beliefs. In a *perfect Bayesian equilibrium (PBE)*, player i's strategy is a best response for *some* Bayesian consistent beliefs. A PBE therefore ensures that strategies are sequentially rational given beliefs, while beliefs are consistent with strategies. However, equilibrium predictions can be highly sensitive to beliefs, which are not generally observable. We therefore also study network outcomes in a *belief-free equilibrium (BFE)*, where player i's strategy must be a best response for *all* beliefs that are Bayesian consistent with the strategy profile. Under perfect monitoring (i.e., $\rho = \infty$), players can observe the entire history in each period, and so each strategy profile has only one Bayesian consistent system of beliefs. In that case, PBE and BFE coincide, and both are equivalent to sub-game perfection. With private monitoring, on the other hand, there are generally multiple Bayesian consistent beliefs off-the path of play, and BFE then refines PBE by requiring best responses that are not sensitive to beliefs.

BFE was introduced for imperfect monitoring games in Piccione [2002], Ely and Välimäki [2002], and Ely et al. [2005], and games with incomplete information in Hörner and Lovo [2009] and Hörner et al. [2011]. In the typical setup, payoffs depend stochastically on the action profile, and BFE requires best responses for all Bayesian consistent beliefs, both on and off the equilibrium path. We adapt BFE to a network setting with a more structured set of relationships between players, and an information structure where players receive no information at all about actions outside their radius of information. Under our adaptation, BFE is a refinement of PBE that ensures that cooperation is not enforced by non-credible threats (as in PBE), and punishment threats are robust to off-path beliefs (unlike in PBE).

2.5 Network notation

For our results, we require some additional notation to describe network structures.

Subnetworks: $\tilde{c}$ is a *subnetwork* of c (denoted $\tilde{c} \subseteq c$) if (i) $ij \in \tilde{c}$ implies $ij \in c$, and (ii) $\tilde{c}_{ij} = c_{ij}$ whenever $ij \in \tilde{c}$; $c - \tilde{c}$ is the complementary subnetwork defined by $ij \in c - \tilde{c}$ if and only if $ij \notin \tilde{c}$ and $ij \in c$; $\mathcal{N}(\tilde{c})$ are the players who are a partner on some tie in $\tilde{c}$.

Paths: There is a *path* of length $n+1$ between nodes i and j in network c whenever there are distinct nodes $k_1, ..., k_n$ such that $ik_1 \in c$, $k_z k_{z+1} \in c$ for all $z = 1, ..., n-1$, and $k_n j \in c$; the distance $d_c(i,j)$ is the length of the shortest path between i and j (where $d_c(i,j) = 1$ if $ij \in c$, and $d_c(i,j) = \infty$ if there is no path). For $\tilde{c} \subseteq c$, $d_c(i,j|\neg\tilde{c}) \equiv d_{c-\tilde{c}}(i,j)$ is the distance between i and j in c when paths are restricted not to pass through ties in $\tilde{c}$.

Categories of ties: We say that i is a *sponsor* of tie ij if $0 < c_{ij} \leq \delta$; a *strict sponsor* if $0 < c_{ij} < \delta$; a *recipient* if $\delta \leq c_{ij}$; and a *strict recipient* if $\delta < c_{ij}$. A tie is *strong* if both partners are strict sponsors (i.e., $\max\{c_{ij}, c_{ji}\} < \delta$); *weak* if both partners are strict recipients (i.e., $\min\{c_{ij}, c_{ji}\} > \delta$); and *asymmetric* if one partner is a strict sponsor and one is a strict recipient (i.e., $\min\{c_{ij}, c_{ji}\} < \delta < \max\{c_{ij}, c_{ji}\}$). Categories of ties are relevant to our analysis because the discounted value of infinite cooperation on tie ij (strictly) exceeds the one-off payoff from a defection on tie ij for player i if and only if i (strictly) sponsors the tie (see Example 2). As such, on strong ties the bilateral cooperation constraint is satisfied for both partners; on weak ties the bilateral cooperation constraint is not satisfied for either partner; and on asymmetric ties the bilateral cooperation constraint is satisfied for one partner but not the other. Off the equilibrium path of play, these bilateral incentives affect how quickly a player would like to discontinue cooperation on a tie: a sponsor will want to maintain cooperation as long as possible, while a strict recipient will want to defect as soon as possible.

Neighborhood: $\mathcal{N}_i(c)$ is set of players who are a partner of i in network c; $\mathcal{S}_i(c)$ are the players with whom i sponsors a tie; $\bar{\mathcal{S}}_i(c)$ are the players with whom i strictly sponsors a tie; and $\mathcal{R}_i(c)$ are the players with whom i is a recipient of a tie. The *ratio* of player i, $\mathcal{Q}_i(c) \equiv \frac{|\mathcal{S}_i(c)|}{|\mathcal{N}_i(c)|}$, measures the proportion of ties that i sponsors.

3 Equilibrium network outcomes

In this section, we analyze equilibrium network outcomes. Our focus is on environments where players cannot commit to cooperative behavior, and efficient networks are then generally not incentive compatible. For partial monitoring we give a strong equivalence result for PBE and BFE. We then study environments with local monitoring, where PBE and BFE identify a range of equilibrium networks outcomes.

3.1 Efficiency

As a benchmark, we first consider an environment where players can commit to cooperative behavior, and are able to make side-payments. Players can then achieve efficient outcomes by treating every bilateral interaction independently. In particular, because γ_{ij} is lower-semicontinuous, there exists some $F^*_{ij} \in \arg\max_{f\in[\underline{F},\bar{F}]} \frac{1-\gamma_{ij}(f)}{1-\delta} - f$, which maximizes player i's net-payoff from a bilateral interaction with player j, and yields a maximum net-payoff $\pi^*_{ij} = \frac{1-\gamma_{ij}(F^*_{ij})}{1-\delta} - F^*_{ij}$. An efficient network outcome is one where any pair i and j form tie $(\gamma_{ij}(F^*_{ij}), \gamma_{ji}(F^*_{ji}))$ if and only if $\pi^*_{ij} + \pi^*_{ji} \geq 0$, and all players commit to cooperate. However, the efficient network outcome is generally not incentive compatible when players cannot commit to cooperative behavior, especially when the BIC cannot be satisfied. In that case, whenever tie ij provides a positive net-payoff for player i there is also a strict incentive for player i to defect in the continuation phase. The tension between formation and continuation incentives therefore implies that efficient outcomes cannot be achieved, as the following example illustrates.

Example 3. [Efficient network with binary formation] In the binary formation model (see Example 2), the efficient network is completely-connected: every players invest $\underline{F}$ in ties with every other player, and cooperates indefinitely. However, since $\gamma(\underline{F}) = \bar{c} > \delta$, each player has a strict incentive to discontinue all of their ties in the continuation phase. As a result, the efficient network is not an equilibrium outcome. At best, players can form heterogeneous ties that allow them to cooperate by pooling incentive constraints. □

3.2 Partial monitoring

Without commitment, players must enforce cooperative behavior with the threat of future punishment. On ties where the BIC (Condition 3) cannot be satisfied, bilateral punishment is not sufficient to enforce cooperation whenever the tie offers a strictly positive net-payoff. As such, players cannot treat bilateral interactions independently, and these ties can only be formed when there are credible threats of third-party punishment.

A network can overcome bilateral incentive constraints by pooling incentives. To illustrate, consider a network c where player i has at least one partner. To cooperate on all ties in network c, the discounted value of cooperation for player i must exceed the one-off benefit of discontinuing all ties, which is always a feasible deviation in the continuation phase. Hence, the following *pooled cooperation constraint (PCC)* must be satisfied:

$$\left(\frac{1}{1-\delta}\right)\sum_{j\in\mathcal{N}_i(c)}(1-c_{ij}) \geq |\mathcal{N}_i(c)|. \tag{5}$$

In addition, the discounted value of cooperating on the ties in network c must exceed the formation cost player i incurs to form these ties, because not forming any ties is also a feasible deviation. Hence, the following *pooled formation constraint (PFC)* must also be satisfied:

$$\left(\frac{1}{1-\delta}\right)\sum_{j\in\mathcal{N}_i(c)}(1-c_{ij}) \geq \sum_{j\in\mathcal{N}_i(c)} F_{ij}^c. \tag{6}$$

Combining (5) and (6) yields the following pooled incentive constraint.

Definition 1. Network c satisfies a *pooled incentive constraint (PIC)* if, for all $i \in \mathcal{N}(c)$,

$$\frac{\sum_{j\in\mathcal{N}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \leq \min\left\{\delta, 1-(1-\delta)\frac{\sum_{j\in\mathcal{N}_i(c)} F_{ij}^c}{|\mathcal{N}_i(c)|}\right\}. \tag{7}$$

The PIC allows player i to form some ties where $1-(1-\delta)F_{ij}^c < c_{ij} < \delta$, and other ties where $\delta < c_{ij} < 1-(1-\delta)F_{ij}^c$. On ties where $c_{ij} < \delta$, there is slack in the bilateral cooperation constraint, while on ties where $c_{ij} < 1-(1-\delta)F_{ij}^c$, there is slack in the bilateral formation constraint. Network c must ensure that the slack in cooperation constraints on ties where $c_{ij} < \delta$ is sufficient to overcome the incentive to defect on ties where $\delta < c_{ij}$, and the slack in formation constraints on ties where $c_{ij} < 1-(1-\delta)F_{ij}^c$ is sufficient to overcome the negative net-payoff on ties where $1-(1-\delta)F_{ij}^c < c_{ij}$.

For any radius of information, the PIC is clearly necessary for network c to emerge in an equilibrium. The following proposition shows that, for any radius of information $\rho \geq 2$, the PIC is also sufficient: when network c satisfies the PIC, there is a BFE where c is formed and players cooperate indefinitely. A fortiori, this means that c is also a network outcome in a PBE. In terms of network outcomes, BFE and PBE are therefore indistinguishable.

Proposition 1. *There is a BFE (F, σ) such that $c(F, \sigma) = c$ if and only if c satisfies the PIC.*

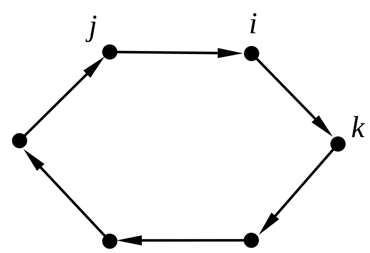
When $\rho \geq 2$, the PIC is sufficient because any deviation can be punished immediately. For instance, if player i deviates from a strategy profile that induces network c on the path of play, all of player i's neighbors observe this deviation. The neighbors can punish i by discontinuing all ties with i in the next period. In that case, player i's optimal deviation is to either not form any ties during the formation phase, or to defect on all ties simultaneously in the cooperation phase; the PIC ensures that these two deviations are not profitable.

To see why, when the PIC is satisfied by network c, enforcement mechanism do not depend on beliefs, suppose $\rho = 2$. The equilibrium strategy profile specifies that a player should discontinue all ties if they observe any deviation inside their radius of information. Now suppose that player i observes a deviation in period t by a direct neighbor. Since $\rho = 2$, this deviation is also observed by all of player i's other neighbors, who will therefore discontinue their ties with i in period $t+1$. As a result, it is optimal for player i to discontinue all ties in $t+1$, regardless of i's beliefs about the history of play outside their radius of information. Alternatively, suppose player i observes a deviation by a neighbor's neighbor. The strategy specifies that i's neighbor discontinue all ties in $t+1$. At that stage, all of i's other neighbors will observe a deviation, and therefore discontinue their ties with i in $t+2$. It is therefore still optimal for i to discontinue all ties in $t+1$ (pre-empting the defection by her neighbors in $t+2$), regardless of i's beliefs about the history. A similar argument holds for any radius $\rho \geq 2$. While the radius of information will determine how quickly cooperation collapses in a network following a deviation, it does not affect the incentive constraints that are relevant on the path of play, as long as neighbors are able to punish deviations immediately.

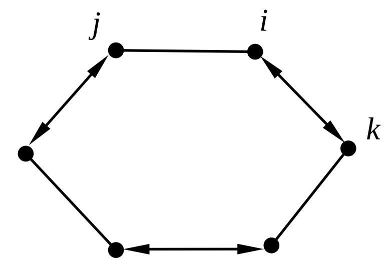
Example 4. [Binary formation with partial monitoring] In the binary formation model, a network c satisfies PIC if and only if, for every player $i \in \mathcal{N}(c)$,

$$\frac{\bar{c}-\delta}{\bar{c}-\underline{c}} \leq \mathcal{Q}_i(c) \leq \frac{\left(\frac{1-\bar{c}}{1-\delta} - \underline{F}\right)}{\left(\frac{1-\bar{c}}{1-\delta} - \underline{F}\right) - \left(\frac{1-\underline{c}}{1-\delta} - \bar{F}\right)} \equiv \frac{\bar{C}}{\bar{C}-\underline{C}}. \tag{8}$$

Condition (8) is consistent with Condition (4) for a wide range of parameters $(\underline{c}, \bar{c}, \underline{F}, \bar{F}, \delta)$. For instance, *any* network where each player sponsors and receives at least one tie satisfies the PIC for some parameters consistent with Condition (4). For Condition (8) to be satisfied, it is necessary for $0 < \mathcal{Q}_i(c) < 1$ and so player i must sponsor and receive at least one tie. In that case, condition (8) is satisfied whenever $\bar{c}$ is sufficiently close to δ and $\bar{C}$ is sufficiently close to $\underline{C}$, which holds for a range of parameters consistent with Condition (4). Conversely, given *any* parameters such that $\frac{\bar{c}-\delta}{\bar{c}-\underline{c}} < \frac{\bar{C}}{\bar{C}-\underline{C}}$, Condition (4) is satisfied and yet there is a completely-connected network (where all players are partners) that is an equilibrium outcome for a sufficiently large population. The reason is that, for a sufficiently large population, it is possible to find a complete network where $\mathcal{Q}_i(c)$ takes any value in $(0, \infty)$ for every player i in the population (see Section 4.1). As a result, a variety of network outcomes can arise in a BFE with partial monitoring. $\square$



(a) Asymmetric ties



(b) Alternating ties

Figure 3: Cycle networks of length $n = 6$

3.3 Local monitoring

With local monitoring ($\rho = 1$), in addition to pooling incentive constraints, the network plays an important role in transmitting information. When players cannot monitor the actions taken by neighbors on other ties, defective behavior must spread through the network to inform third-parties that a deviation has occurred. As a result, incentive constraints depend also on the distance between nodes that a contagion of punishment must travel. How distance affects incentives to cooperate can be illustrated with the class of cycle networks.

Example 5. [Cycle networks with binary formation] In the binary formation model, we can compare when a cycle network of length $n \geq 3$ (Figure 3) is incentive compatible under local versus partial monitoring. Every player must sponsor and receive at least one tie. This is possible, for instance, when all ties are asymmetric (Figure *3(a)*) or ties alternate between weak and strong (Figure *3(b)*). In both cases, the cycle network satisfies the pooled cooperation constraint if and only if

$$\frac{\bar{c} - \delta}{\delta - \underline{c}} \leq 1. \tag{9}$$

When condition (9) is satisfied, a cycle network formed during the formation phase can be continued indefinitely for $\rho \geq 2$. Players cooperate unless they observe a single defection, and then defect on all remaining ties. In particular, the pooled cooperation constraint does not depend on the size of the cycle. On the other hand, with local monitoring, when player i defects on tie ij, player k does not observe the defection. Instead, third-party punishment by player k may be delayed until defection can spread through the network, which takes at least $n - 2$ periods. As a result, player i can defect on tie ij in period 0 and continue cooperating on ik until period $n - 2$, benefiting from the higher cooperative payoff on tie ik in the meantime. Comparing the value of this deviation to the discounted value of cooperating indefinitely leads to the following private cooperation constraint on a cycle network of length n:

$$\frac{\bar{c} - \delta}{\delta - \underline{c}} \leq \delta^{n-2}. \tag{10}$$

The left-hand side in Condition (10) is the same as under partial monitoring. However, the right-hand side is smaller (since $\delta < 1$), and depends on the length of the cycle. In particular, the right-hand side equals δ when $n = 3$ (the smallest cycle), is decreasing in n, and converges to 0 as $n \to \infty$. As a result, there are parameters such that a cycle network of any length can emerge under partial monitoring, but a cycle of sufficient length n is not incentive compatible under local monitoring. □

Equilibrium predictions under local monitoring can be sensitive to beliefs about the history of play. For instance, consider a strategy profile that prescribes the cycle network outcome c in Figure 3(a). Now suppose that, in period t, player j defects on tie ij, but player k cooperates on ik. What is the optimal response for player i? There are (at least) two histories consistent with i's observations. First, it could be that only tie ij has been discontinued, and so it will take at least $n-1$ periods before defection spreads through the network and reaches k. Since player i would like to cooperate on tie ik as long as possible, it is optimal for player i to continue cooperating until period $t+n-2$. Alternatively, all ties other than ik may have been discontinued. In that case, player k will defect on ik in $t+1$, and it is optimal to defect on ik immediately. Both of these histories, which lead to different best-responses, are in the support of Bayesian consistent beliefs but, with local monitoring, there is no way for i to distinguish between the histories. Player i's response to j's deviation is optimal given some beliefs, but not optimal for other beliefs. This dependence on beliefs raises questions about the robustness of network predictions under local monitoring. We therefore consider two extremes, PBE and BFE, which identify a range of reasonable equilibrium outcomes.

3.3.1 Perfect Bayesian equilibrium

To describe the incentive constraints that must be satisfied for a network c to emerge in a PBE, we fix a player i and let all opponents follow a simple "grim" strategy. Suppose every player $j \neq i$ invests F^c_{jk} with player $k \in \mathcal{I}$ in the formation phase, cooperates on every tie when their current network neighborhood is c_j, and otherwise defects on all remaining ties. This grim strategy is not sequentially rational for player j, but allows us to formulate the incentive constraints for player i. In particular, given the grim strategy followed by her opponents, we can consider possible *defection plans* for player i. A defection plan is a vector $\alpha^i = \left(\alpha^i_j\right)_{j \neq i} \in (\{-1\} \cup \mathbb{N})^{I-1}$, where α^i_j represents the first period in which player i plans to discontinue cooperation with player j, and $\alpha^i_j = -1$ if i does not plan to form a tie with player j.

For network c, a defection plan for player i should satisfy at least the following two conditions: (i) $\alpha_j^i = -1$ if $ij \notin c$, and (ii) $\alpha_j^i \leq \min_{k \in \mathcal{N}_i(c)-\{j\}} \{\alpha_k^i + d_c[i,j|\neg\{ik\}] - 1\}$ if $ij \in c$. The first condition states that i does not form ties that are not in network c, because it would be sub-optimal to do so. The second condition can be interpreted as follows. When i discontinues a tie with player k in period α_k^i, the grim strategy prescribes that player k will discontinue all their ties in period $\alpha_k^i + 1$; player k's neighbors discontinue their remaining ties in $\alpha_k^i + 2$, and so on. Player i's defection with k in period α_k^i thereby creates a contagion of defection in the network, and $d_c[i,j|\neg\{ik\}] - 1$ is the maximum number of periods for this contagion to reach player j.[9] Hence, player j will defect on ij in at most period $\alpha_k^i + d_c[i,j|\neg\{ik\}]$. Preempting the defection by j, player i should defect on ij latest in period $\alpha_k^i + d_c[i,j|\neg\{ik\}] - 1$. We denote by $\mathcal{A}_i^c$ the set of defection plans for player i that satisfy conditions (i) and (ii) for a network c.

Given the set of defection plans $\mathcal{A}_i^c$, we can describe the maximum continuation payoff player i can achieve by deviating from a strategy that prescribes network outcome c. If $\alpha_j^i = -1$, player i does not plan to form a tie with player j, and therefore receives a payoff of 0 in the interaction with j. If $\alpha_j^i \geq 0$, player i plans to form the tie $ij \in c$, incurring at least the formation cost F_{ij}^c, then receives the cooperative payoff $(1 - c_{ij})$ for at most $\alpha_j^i - 1$ periods, and finally at best receives the defection payoff 1 in period α_j^i. The maximum continuation payoff player i achieves in the interaction with player j is therefore

$$\Pi_j^i\left(\alpha_j^i\right) \equiv \begin{cases} \left(\frac{1-\delta^{\alpha_j^i}}{1-\delta}\right)(1-c_{ij}) + \delta^{\alpha_j^i} - F_{ij}^c & \text{if } \alpha_j^i \geq 0 \\ 0 & \text{if } \alpha_j^i = -1 \end{cases}. \tag{11}$$

We say that a network c satisfies a local incentive constraint if, for all defection plans $\alpha^i \in \mathcal{A}_i^c$, the maximum continuation payoff, $\sum_{j \neq i} \Pi_j^i(\alpha^i)$, is no greater than the continuation payoff from network outcome c.

Definition 2. Network c satisfies a *local incentive constraint* (LIC) if, for all $i \in \mathcal{N}(c)$,

$$\sum_{j \in \mathcal{I}-\{i\}} \Pi_j^i(\alpha_j^i) \leq \sum_{j \in \mathcal{N}_i(c)} \left(\frac{1-c_{ij}}{1-\delta} - F_{ij}^c\right) \qquad \forall \alpha^i \in \mathcal{A}_i^c. \tag{12}$$

The LIC compares the payoff player i receives from network outcome c with all possible continuation payoffs from deviations when the other players follow a grim strategy. The grim strategy defines the harshest punishment regime for player i because any alternative strategy profile by opponents can only increase the time it takes for player i to be punished for a deviation. As such, when player i has a profitable deviation from network outcome c under the

[9] Note that $d_c[i,j|\neg\{ik\}] - 1$ is the maximum time needed for contagion to reach j because, if i deviates on other ties in the meantime, defection could spread via network paths that are shorter than the path from k.

grim strategy, there is also a profitable deviation for any alternative strategy profile. As a result, the LIC is necessary for network outcome c to emerge in a PBE. The following proposition shows that the LIC is also sufficient.

Proposition 2. *There is a PBE (F, σ) such that $c(F, \sigma) = c$ if and only if c satisfies the LIC.*

To provide intuition for the sufficiency argument, consider a network c where all players have at least two partners.[10] Then the LIC is sufficient for a PBE because there are Bayesian consistent beliefs such that it is optimal for player i to discontinue all remaining ties when i observes a defection by j on tie ij. In particular, suppose that $c(F, \sigma) = c$ and, when i observes a deviation from j, she believes (with probability 1) that all ties outside her neighborhood have also been discontinued. These beliefs are Bayesian consistent with the strategy-profile because the history is consistent with i's observations, and play is off the equilibrium path. Moreover, if the strategy-profile σ prescribes other players to discontinue all ties when they observe a deviation, then it is a best-response for i to discontinue all ties because player i believes that her neighbors have observed a deviation, and will therefore defect on i in the next period.

Proposition 2 highlights the distinction between an environment with partial and local monitoring. To see how the LIC is related to the PIC, consider a network c with $i \in \mathcal{N}(c)$. One defection plan that is always feasible for player i is to form no ties, i.e., $\tilde{\alpha}^i = (-1, ..., -1)$. In that case, $\Pi^i_j(\tilde{\alpha}^i_j) = 0$ for all $j \in \mathcal{I} - \{i\}$, and so the inequality in (12) reduces to

$$0 \leq \sum_{j \in \mathcal{N}_i(c)} (1 - c_{ij}) - (1 - \delta) \sum_{j \in \mathcal{N}_i(c)} F^c_{ij},$$

which is exactly the pooled formation constraint (6). Another defection plan that is always feasible is to form all ties in c_i but then defect on these ties immediately, i.e., $\hat{\alpha}^i_j = -1$ when $ij \notin c$ and $\hat{\alpha}^i_j = 0$ when $ij \in c$. In that case, $\Pi^i_j(\hat{\alpha}^i_j) = 0$ for $ij \notin c$ and $\Pi^i_j(\hat{\alpha}^i_j) = 1 - F^c_{ij}$ for $ij \in c$, and so the inequality in (12) reduces to

$$(1 - \delta) \sum_{j \in \mathcal{N}_i(c)} (1 - F^c_{ij}) \leq \sum_{j \in \mathcal{N}_i(c)} (1 - c_{ij}) + (1 - \delta) \sum_{j \in \mathcal{N}_i(c)} F^c_{ij},$$

which is exactly the pooled cooperation constraint (5). As a result, the LIC implies the PIC. On the other hand, the PIC does not imply the LIC because, in addition to the defection plans $\tilde{\alpha}^i$ and $\hat{\alpha}^i$, the set of defection plans $\mathcal{A}^c_i$ generally contains defection plans that exploit local monitoring. For instance, for the cycle network in Example 5, one feasible defection plan for player i is to choose $\alpha^i_l = -1$ for $l \notin \{j, k\}$, $\alpha^i_j = 0$ and $\alpha^i_k = n - 2$. This defection plan

[10] When player j has only one partner i in network c, the LIC implies that tie ij must satisfy the BIC and there is no reason for i or j to ever discontinue tie ij. As a result, such ties can be treated as strategically independent of the remaining network.

is feasible because the distance between nodes i and k on the path that do not use tie ij is $d_c\left[i,k|\neg\{ij\}\right]=n-1$. For this defection plan, inequality (12) in the LIC is equivalent to

$$1-\bar{F}+\frac{(1-\underline{c})(1-\delta^{n-2})}{1-\delta}+\delta^{n-2}-\underline{F}\leq\frac{1-\underline{c}}{1-\delta}+\frac{1-\bar{c}}{1-\delta}-\bar{F}-\underline{F}, \tag{13}$$

which is exactly inequality (10).

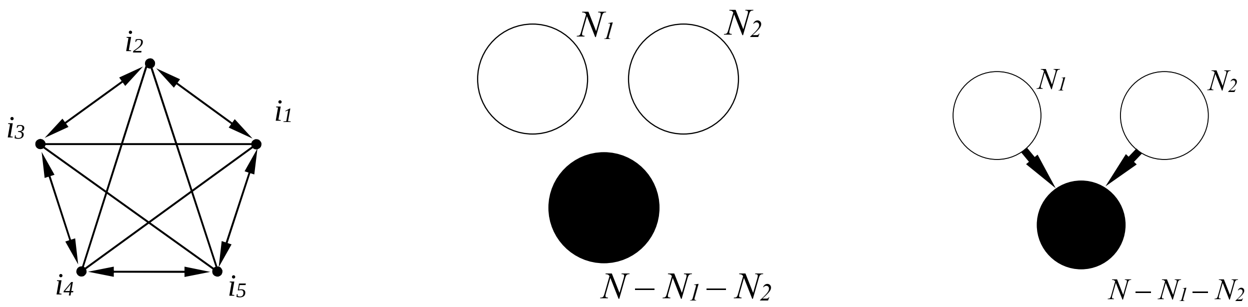
3.3.2 Belief-free equilibrium

In a BFE, a strategy must be sequentially rational for *all* Bayesian consistent beliefs. As a result, equilibrium network outcomes involve only enforcement mechanisms that are robust to local monitoring. For instance, even when a cycle network of length n satisfies the LIC, it is generally not a BFE outcome because, as the discussion following Example 5 indicates, enforcement mechanisms are sensitive to off-path beliefs. There is one important exception: the smallest cycle with length $n=3$ (Figure *2(b)*). On a cycle of length $n=3$, when i observes a deviation by j, there are essentially two histories consistent with i's observations: either the tie jk has also been discontinued, or the tie jk is still cooperative. However, for both of these histories, it is optimal for i to defect on player k in period $t+1$, because k will discontinue the tie latest in $t+2$. More generally, whatever strategy profile players follow to induce a cycle network of length $n=3$, when player i observes a non-cooperative deviation, it is always optimal to defect on all remaining ties in the following period for *every* Bayesian consistent belief. In that sense, enforcement mechanisms on a cycle of length $n=3$ are robust to uncertainty about the history in parts of the network that players cannot monitor.

To generalize the idea that the smallest cycle network can be enforced more robustly than larger cycle networks, we require one further graph-theoretic definition.

Definition 3. Network c satisfies *triadic closure* if $j\in\bar{\mathcal{S}}_i(c)$ and $k\in\mathcal{N}_i(c)-\{j\}$ implies $kj\in c$, i.e., if player i sponsors a tie with player j and is also a partner with player k, then players j and k must be partners as well.

First introduced by Simmel [1908], the theory that strong ties tend to be triadically closed has occupied a prominent role in sociology. Our framework rationalizes triadic closure in terms of belief-free strategic interaction under local monitoring. Our notion of triadic closure is somewhat stronger than the standard definition, where a network is triadically closed if there is a tie between any two nodes that have strong ties to the same partner (see, e.g., Easley and Kleinberg [2010, Chapter 3]). Our property requires that, whenever player i strictly sponsors a tie with player j, any other partner k must also be a partner of player j. Triadic closure therefore implies dense clustering around all strong and asymmetric ties.



(a) Complete network (b) Disjoint clusters (c) Connected clusters

Figure 4: Triadic closure.

Figure 4(a) depicts a completely-connected network. In Figure 4(b), the empty circle indicates a cluster of nodes that are connected to each other by weak ties, and the filled circle indicates a cluster of nodes that are connected to each other by strong ties. In Figure 4(c), the circles indicate completely-connected clusters, as in 4(b). The wide arrows indicate that each player in a weak cluster has an asymmetric tie to each player in the strong cluster, where the tie is sponsored by the player in the weak cluster.

Example 6. A completely-connected network (Figure *4a*) satisfies triadic closure. There are also incomplete networks that satisfy the property. For instance, suppose that a subset of nodes $\mathcal{N}_1$ are all connected to each other by weak ties, a separate subset of nodes $\mathcal{N}_2$ are also connected to each other by weak ties, and all remaining nodes are connected to each other by strong ties (Figure *4b*). This network consists of disjoint clusters, and therefore satisfies triadic closure. Moreover, suppose that, in addition, for any nodes $i \in \mathcal{N}_1$, $j \in \mathcal{N}_2$, and $k \in \mathcal{N} - \mathcal{N}_1 - \mathcal{N}_2$, there is a tie ik strictly sponsored by i, and a tie jk strictly sponsored by j (Figure *4c*). The resulting network has two weak clusters connected to a strong cluster by asymmetric ties. While it does not consist of separate cliques, this network also satisfies triadic closure. $\square$

The LIC simplifies when a network satisfies triadic closure. Consider the possible defection plans for player i given network c, which satisfies triadic closure. First, to ensure that the defection plan $\hat{\alpha}^i = (-1, ..., -1)$, where i forms no ties, is not a profitable deviation, network c must satisfy the pooled formation constraint (6). Another feasible defection plan is

$$\bar{\alpha}^i_j \equiv \begin{cases} -1 & \text{if } j \notin \mathcal{N}_i(c) \\ 0 & \text{if } j \in \mathcal{S}_i(c) \\ 1 & \text{if } j \in \mathcal{N}_i(c) - \mathcal{S}_i(c) \end{cases},$$

where player i defects on ties she receives in period 0, and defects on remaining ties in period 1.

Since $\bar{\alpha}^i$ is feasible, it must be the case that

$$\sum_{j \neq i} \Pi_j^i(\bar{\alpha}_j^i) \leq \sum_{j \in \mathcal{N}_i(c)} \left(\frac{1 - c_{ij}}{1 - \delta} - F_{ij}^c \right).$$

Re-arranging terms, the feasibility of the plans $\hat{\alpha}^i$ and $\bar{\alpha}^i$ yields a robust incentive constraint that is necessary for a BFE network outcome.

Definition 4. Network c satisfies a *robust incentive constraint* (RIC) if, for all $i \in \mathcal{N}(c)$,

$$\frac{\sum_{j \in \mathcal{N}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \leq 1 - (1 - \delta) \frac{\sum_{j \in \mathcal{N}_i(c)} F_{ij}^c}{|\mathcal{N}_i(c)|} \tag{14}$$

and

$$\frac{\frac{1}{\delta} \sum_{j \in \mathcal{R}_i(c)} c_{ij} + \sum_{j \in \mathcal{S}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \leq 1 - (1 - \delta) \mathcal{Q}_i(c). \tag{15}$$

To illustrate how RIC is related to LIC, consider a network c with $i \in \mathcal{N}_i(c)$. The LIC is satisfied when $\sum_{j \neq i} \Pi_j^i(\alpha^i)$ is less than the continuation value of network outcome c for player i in every period, and for every defection plan $\alpha^i \in \mathcal{A}_c^i$. As $\mathcal{A}_c^i$ contains the defection plans $\hat{\alpha}^i$ and $\bar{\alpha}^i$, the LIC implies the RIC. On the other hand, the RIC is defined only in terms of the defection plans $\hat{\alpha}^i$ and $\bar{\alpha}^i$, and so a network can satisfy the RIC even when it does not satisfy the LIC. In particular, the RIC does not depend on the distance between nodes in the network. Indeed, the RIC is more closely related to the incentive constraints under partial monitoring. Condition (14) in the RIC is exactly the pooled formation constraint (6). Moreover, it is easily verified that the pooled cooperation constraint (5) is satisfied for any network for which the robust cooperation constraint (15) is satisfied. Hence, a network that satisfies the RIC always satisfies the PIC, and both conditions–unlike the PIC–are independent of network distance.

The RIC is more restrictive than the PIC for the following reason. Under partial monitoring, if player i deviates from cooperation on a tie she receives in period t, i's neighbors can punish i in period $t+1$ on all ties. With local monitoring, such third-party punishments are infeasible because player i's deviation is not observed by her other neighbors. Third-party punishment is therefore delayed for at least two additional periods. If the deviation induces third-party punishment on ties in $t+2$, the optimal deviation by player i is to discontinue all ties she receives in period t, and discontinue ties she sponsors in $t+1$ (preempting punishment in period $t+2$). As with the LIC, the difference between the RIC and the PIC is therefore due to the delay in third-party punishment, which player i discounts with δ.

The preceding intuition is incomplete because, while third-party punishment must occur with a delay of at least two periods, there is no guarantee that it will occur in exactly two periods. In general networks, it may take much longer for non-cooperative deviations to spread.

It is triadic closure that ensures non-cooperative deviations can reach player i in exactly two periods, because any partner of player i is a partner of every other player with whom i sponsors a tie. For a network that satisfies triadic closure, the RIC is therefore equivalent to the LIC. The RIC characterizes incentive constraints in a BFE because, as the following proposition shows, BFE network outcomes must satisfy triadic closure.

Proposition 3. *(i) If network c satisfies triadic closure and the RIC, then there is a BFE (F, σ) such that $c(F, \sigma) = c$.*

(ii) If there is a BFE (F, σ) such that $c(F, \sigma) = c$, then there is a subnetwork $\tilde{c} \subseteq c$, containing all weak and asymmetric ties, and satisfying triadic closure and the RIC.

Part (i) provides sufficient conditions for BFE network outcomes under local monitoring. For instance, in the binary formation model (Example 2), given *any* parameters such that $\frac{1}{\delta}\frac{\bar{c}-\delta}{\bar{c}-\underline{c}} < \frac{\bar{C}}{\bar{C}-\underline{C}}$ there is a complete network that can be realized as a BFE outcome for a sufficiently large population. A complete network satisfies the triadic closure property and, for a sufficiently large population, the RIC can be satisfied as the LIC can (Example 4). However, it is no longer the case that any network where every player strictly receives and sponsors at least one tie can emerge for some parameters. The reason is that, by part (ii), triadic closure imposes a significant restriction on global network structure. The subnetwork $\tilde{c}$, which must satisfy triadic closure, must contain all weak and asymmetric ties. In many cases, these ties are essential to make network formation incentive compatible (Condition 14). In addition, the network $\tilde{c}$ must satisfy the RIC: it must contain enough strong or asymmetric ties to ensure that each player's cooperation constraint (Condition 15) holds. Strong and asymmetric ties must therefore be clustered with other ties. As such, any BFE network must consist of dense clusters of ties, exhibiting high clustering coefficients and short path lengths, consistent with the "small-worlds" observed in many social and economic networks (Watts and Strogatz [1998]).

3.3.3 Discussion of local monitoring

Our results identify two key differences between local versus partial monitoring. First, players have greater opportunity for deviations with local monitoring because neighbors cannot observe deviations directly. The delay for contagion to spread through a network makes deviations more profitable, and therefore refines the set of networks that satisfy incentive constraints. Second, under local monitoring, when a player observes a deviation, the best-response generally depends on the player's beliefs about the history of play outside their radius of information. Equilibrium predictions about network outcomes are therefore sensitive to assumptions about beliefs, which are generally not observable.

Our results establish the scope for equilibrium network outcomes under local monitoring. PBE imposes (arguably) the weakest restrictions on off-path beliefs consistent with an equilibrium notion of sequential rationality. As Proposition 2 shows, PBE network outcomes are characterized by the LIC, which illustrates how local monitoring can refine network outcomes because of the additional opportunities for unmonitored deviations. In particular, comparing the LIC with the PIC indicates how the opportunity for network enforcement diminishes when the distance between nodes increases, and so local monitoring necessitates the formation of tighter networks with shorter paths and higher clustering. On the other hand, BFE reflects (arguably) the strongest restrictions on network structure, by imposing that best-responses are independent of beliefs. As Proposition 3 shows, networks that can emerge in a BFE must satisfy a triadic closure property, which illustrates how sensitive behavior is to beliefs under local monitoring.

For the range of network outcomes under local monitoring, we can compare the predictions under PBE and BFE. For instance, for the cycle networks in Example 5, arbitrary lengths can emerge in a PBE for the right parameters, while only length $n = 3$ can emerge in a BFE. However, for fixed parameters, the LIC imposes restriction on length. For instance, when

$$1 + (1 - \underline{c}) + \delta < \frac{(1 - \underline{c}) + (1 - \bar{c})}{1 - \delta} < 1 + (1 - \underline{c})(1 + \delta) + \delta^2,$$

a cycle of length $n \geq 4$ does not satisfy the LIC, and PBE and BFE therefore make the same prediction about the length of a cycle network (i.e., $n = 3$). For parameters where the network outcomes under PBE diverge significantly from BFE, there is wider scope for equilibrium predictions under local monitoring. When the predictions under PBE and BFE diverge, alternative equilibrium refinements can impose more structure on equilibrium beliefs than a PBE. In Appendix A.4, we provide an example to illustrate how sequential equilibrium (SE) refines the set of PBE network outcomes. Since a BFE is sequential, Proposition 3 provides sufficient conditions for network outcomes in a SE. On the other hand, since a SE is a PBE, Proposition 2 provides necessary conditions for SE network outcomes. The example illustrates that neither of these conditions fully characterizes SE network outcomes, which can depend on fine details of network structure.

4 Examples

In our framework, the role of a network is most pronounced when no pair of players can form a tie that satisfies the BIC. Such environments are characterized by the following condition:

Assumption 1. *For all $i, j \in \mathcal{I}$, either $\gamma_{ij}(f) > \min\{\delta, 1 - (1 - \delta)f\}$ for all $f \in \left[\underline{F}, \bar{F}\right]$ or $\gamma_{ji}(f) > \min\{\delta, 1 - (1 - \delta)f\}$ for all $f \in \left[\underline{F}, \bar{F}\right]$.*

Under Assumption 1, any bilateral tie must be embedded in a network of other ties because, in a strategic isolation, either the BIC is violated for player i or for player j (or both). For the network to pool incentives, the ties cannot be homogeneous. In a network where all ties are weak ($c_{ij} \geq \delta$ for all i and j), players would not have an incentive to cooperate; if all ties are strong ($c_{ij} \leq \delta$ for all i and j), players would receive strictly negative payoffs in the formation phase. On the other hand, when ties are heterogeneous, players can use the network to pool incentives, and Propositions 1–3 establish necessary and sufficient conditions on network outcomes. There are two potential sources of heterogeneity in our setting: (i) primitive asymmetries in the formation technology Γ, and (ii) endogenous asymmetries from differing investments during the forming phase.[11] In this section, we consider some examples of the formation technology Γ, and illustrate the type of network outcomes that can arise under Assumption 1.

4.1 Homogeneous formation model

In the homogeneous formation model, the formation technology is described by a single function $\gamma : [\underline{F}, \bar{F}] \to (0, 1)$, such that $\gamma_{ij} \equiv \gamma$ for all players i and j. As a result, there are no primitive asymmetries, and any heterogeneity in ties results from differences in formation investments. Non-empty networks can emerge in this setting only when there are non-convexities in the formation technology. With a homogeneous and convex formation technology γ, networks cannot pool incentive constraints, and the empty network is the only equilibrium outcome.[12]

When there are non-convexities in the formation technology, non-empty networks can emerge in equilibrium even when the formation technology is homogeneous. A simple illustration is the binary formation model, where $\gamma(f) = \bar{c}$ for $f \in [\underline{F}, \bar{F})$ and $\gamma(\bar{F}) = \underline{c}$, for some $0 < \underline{c} < \bar{c} < 1$. In this binary model, the formation technology is homogeneous but non-convex, and the unique efficient outcome is a complete network of weak ties (Example 3). The efficient network outcome is not feasible, but network outcomes can emerge where players form cooperative ties to achieve

[11] An increasingly studied area in the multidisciplinary literature on networks where asymmetries can naturally arise for both exogenous and endogenous reasons are multidimensional networks, also called multiplex or multilayer networks (Mucha et al. [2010]; De Domenico et al. [2013]; Kivelä et al. [2014]). In such networks, there can be several, qualitatively different ties between any two nodes. For instance, countries, organizations, or people may interact in multiple spheres, which could be modeled as separate ties, with varying incentives. The analysis of multidimensional networks potentially raises a number of new challenges, but a significant literature has studied scenarios under which the interactions on ties are "reducible" (e.g., De Domenico et al. [2015]). When ties are reducible, it is possible to collapse multi-link relationships between nodes into a single "net-value," but the ties will generally exhibit significant heterogeneity.

[12] To illustrate, consider an environment with partial monitoring, where the PIC characterizes equilibrium network outcomes. Let c be a non-empty network with $i \in \mathcal{N}(c)$, and let $\bar{f} \equiv |\mathcal{N}_i(c)|^{-1} \sum_{j \in \mathcal{N}_i(c)} F^c_{ij}$. When γ is convex, $|\mathcal{N}_i(c)|^{-1} \sum_{j \in \mathcal{N}_i(c)} c_{ij} \geq \gamma\left(\bar{f}\right)$. Since the PIC requires $|\mathcal{N}_i(c)|^{-1} \sum_{j \in \mathcal{N}_i(c)} c_{ij} \leq \min\{\delta, 1 - (1-\delta)\bar{f}\}$, in order for network c to be an equilibrium outcome it must be that $\gamma(\bar{f}) \leq \min\{\delta, 1 - (1-\delta)\bar{f}\}$, which violates Assumption 1 because $\bar{f} \in [\underline{F}, \bar{F}]$.

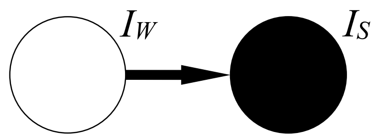


Figure 5: Endogenous asymmetries network

Nodes in cluster $\mathcal{I}_W$ are connected by weak ties; nodes in cluster $\mathcal{I}_S$ are connected by strong ties; each node in cluster $\mathcal{I}_W$ is connected by an asymmetric tie to each node in cluster $\mathcal{I}_S$, which is sponsored by the node in $\mathcal{I}_W$. This network is completely-connected and therefore satisfies triadic closure. By adjusting the absolute size of the population, as well as the relative size of cluster $\mathcal{I}_S$, the maximal and minimal ratio in the network can take any value in $(0, \infty)$ to ensure that the RIC is satisfied .

strictly positive net-payoffs. For instance, a complete network outcome c, in which all players are partners, can emerge in a BFE under local monitoring if and only if

$$\frac{1}{\delta}\frac{\bar{c}-\delta}{\bar{c}-\delta} \leq \min_{i\in\mathcal{I}} \mathcal{Q}_i(c) \leq \max_{i\in\mathcal{I}} \mathcal{Q}_i(c) \leq \frac{\bar{C}}{\bar{C}-\underline{C}}, \tag{16}$$

i.e., every player sponsors enough ties to ensure that the private cooperation constraint $\frac{1}{\delta}\frac{\bar{c}-\delta}{\bar{c}-\delta} \leq \min_{i\in\mathcal{I}} \mathcal{Q}_i(c)$ is satisfied, and receives enough ties to ensure that formation constraint $\max_{i\in\mathcal{I}} \mathcal{Q}_i(c) \leq \frac{\bar{C}}{\bar{C}-\underline{C}}$ is satisfied. Similar to the partial monitoring case (Example 4), as long as $\frac{1}{\delta}\frac{\bar{c}-\delta}{\bar{c}-\delta} < \frac{\bar{C}}{\bar{C}-\underline{C}}$, a complete network can therefore be formed for any sufficiently large population. For instance, partition the set of players into two groups: $\mathcal{I}_W$ and $\mathcal{I}_S$. Consider a network c where all players in $\mathcal{I}_W$ are connected by weak ties, all players in $\mathcal{I}_S$ are connected by strong ties, and every player in $\mathcal{I}_W$ has an asymmetric tie with every player in $\mathcal{I}_S$ that is sponsored by the player in $\mathcal{I}_W$ (Figure 5). Then $\mathcal{Q}_i(c) = \frac{|\mathcal{I}_S|}{I-1}$ for $i \in \mathcal{I}_W$ and $\mathcal{Q}_j(c) = \frac{|\mathcal{I}_S|-1}{I-1}$ for $j \in \mathcal{I}_S$. For a large population, $\min \mathcal{Q}_i(c) \approx \max \mathcal{Q}_i(c)$, and these ratios can be chosen so that Condition (16) is satisfied by adjusting the relative size of group $\mathcal{I}_S$. As a result, a BFE network outcome can emerge where all players are connected and receive strictly positive net-payoffs.

4.2 Within-between group model

With asymmetries in the formation technology, non-empty equilibrium network outcomes can emerge also when the formation technology is convex. A simple example is the within-between group model, where the population is divided into groups and the technology is described by two functions, γ_W and γ_B, which represent the technology for forming ties within versus between groups. A stylized example are informal risk-sharing arrangements in rural communities, where farmers can interact with other farmers in their own village or farmers in other villages (see. e.g. Cox and Fafchamps [2007]; or Fafchamps [2008]).

Suppose there are two villages of equal size, and two linear formation technologies. The

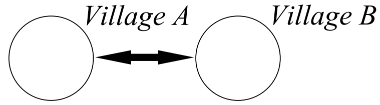


Figure 6: Risk-sharing network

Within villages, farmers form weak ties; between villages, farmers form strong ties. The network is completely-connected and therefore satisfies triadic closure.

function γ_W describes the technology for forming ties within a village and has both a lower intercept and smaller slope than the function γ_B, which describes the technology for forming ties between villages. These assumptions reflect the idea that (i) it is relatively easy to form ties within a village (e.g., because of geographic or cultural proximity), and (ii) investing more in a tie between villages has a greater marginal benefit in terms of the cooperative payoff (e.g., because the benefits of risk-sharing are greater when a partner is exposed to less correlated shocks). These asymmetries generate opportunities for villagers to form ties where they make low investments $\underline{F}$ within a village, and high investments $\bar{F}$ between villages. Specifically, when farmers i and j are in the same village, assume that $\gamma_W(\underline{F}) \in (\delta, 1-(1-\delta)\underline{F})$. Since $\gamma_W(\underline{F}) < 1-(1-\delta)\underline{F}$, it would be beneficial for i and j to form a tie with the low investment $\underline{F}$ if they could commit to cooperate. However, $\gamma_W(\underline{F}) > \delta$ implies that cooperation is not bilaterally incentive compatible for the investment $\underline{F}$. Forming a weak tie is therefore relatively easy within a village, but the benefits of cooperation are relatively low. On the other hand, when farmers i and j are in different villages, let $\gamma_B(\bar{F}) \in (1-(1-\delta)\bar{F}, \delta)$. Since $\gamma_B(\bar{F}) < \delta$, farmers i and j could enforce cooperation bilaterally on a tie with high investment $\bar{F}$. However, $\gamma_B(\bar{F}) > 1-(1-\delta)\bar{F}$ implies that the large investment $\bar{F}$ needed to form such a tie generates a negative net-payoff.

Assumption 1 is satisfied when $\delta < \min\{\gamma_W(1), \gamma_B(1)\}$, in which case it is not possible for villagers to form strategically independent ties. However, it is possible to form cooperative ties embedded in a network. Suppose that all farmers receive ties with other farmers in their own village, and sponsor ties with farmers in the other village. These investments induce a complete network c (Figure 6) that satisfies triadic closure. Moreover, receiving a tie from a farmer in the same village is relatively cheap and, as long as cooperative behavior can be enforced, ties within a village can generate a strictly positive net payoff. On the other hand, once ties are formed, the high value of cooperation between villages can generate a strictly positive net payoff from cooperation, overcoming incentives to defect within-village ties. Setting $\bar{c} \equiv \gamma_W(\underline{F})$ and $\underline{c} \equiv \gamma_B(\bar{F})$, Condition (16) from the binary formation model again characterizes the RIC, where now $\min \mathcal{Q}_i(c) = \frac{I-1}{2(I-1)}$ and $\max \mathcal{Q}_i(c) = \frac{I}{2(I-1)}$. For a sufficiently large population, $\min \mathcal{Q}_i(c) \approx \max \mathcal{Q}_i(c) \approx \frac{1}{2}$ and the RIC is satisfied when $\frac{2-\underline{c}-\bar{c}}{1-\delta} > \max\left\{\underline{F}+\bar{F}, 2-\underline{c}+\delta\right\}$, which holds for $\delta \to 1$.

This within-between group model can be used to illustrate a stylized mechanism by which formal institutions crowd-out informal risk sharing arrangements (see, e.g., Dercon and Krishnan [2003]; Klohn and Strupat [2013]). For instance, suppose that formal insurance instruments are introduced in village A. Farmers offered formal insurance may prefer these to informal risk-sharing and, ceteris paribus, the insurance program may increase welfare. However, the program also has a spill-over effect on village B. When farmers in village A have formal insurance, they no longer need the strong ties with farmers in village B. At first, this reduces the amount of informal insurance that farmers in village B have access to as the between-village ties with farmers in A are discontinued. However, there is a secondary equilibrium effect. The strong between-village ties do not only provide opportunities for bilateral cooperation, but are also essential for third-party punishment that enforces cooperation on the weak ties within village B. Hence, when farmers in village A receive access to formal insurance products, the farmers village B are unable to enforce informal risk-sharing arrangements among themselves. As a result, the introduction of a partial formal insurance program leads to a "crowding-out" of informal risk-sharing arrangements in the village not directly targeted by the intervention.

4.3 Target partner model

The target partner model, where the population is divided into groups and the technology depends on the group membership of the partner, is another asymmetric formation technology. For example, consider a setting where ties represent informal information-sharing arrangements between firms. Empirical research on the patterns of information-sharing has identified a number of structural properties that are common across industries and time periods (Gulati [1995]). For example, in a large-scale study of R&D collaborations over 20 years, Tomasello et al. [2013, p. 25] conclude that "across sectors, firms show the tendency to organize their R&D collaborations in a core of densely connected companies and a periphery of companies that are linked to the core, but only weakly interconnected among themselves."

The core-periphery networks observed in empirical studies can emerge in a BFE of the target partner model. Suppose that firms can be divided in two groups: small firms ($\mathcal{I}_S$) and large firms ($\mathcal{I}_L$). It is more costly to initiate a tie with a large firm, but the marginal benefit of investing more heavily in the tie is also higher. This can be represented by a technology described in terms of two functions, γ_S and γ_L, which represent the technology for forming ties with small versus large firms. For simplicity, let γ_L and γ_S be linear and suppose that $\gamma_S(\underline{F}) \in (\delta, 1-(1-\delta)\underline{F})$ and $\gamma_L(\bar{F}) \in (1-(1-\delta)\bar{F}, \delta)$. The restriction $1 \leq \min\{\gamma_S(1), \gamma_L(1)\}$ ensure that Assumption 1 is satisfied, yet non-empty network outcomes can emerge in a BFE. For instance, suppose that there are $\ell \equiv |\mathcal{I}_L|$ large firms, and the set of small firms is divided

(a) Core-sponsored network (b) Periphery-sponsored network

Figure 7: Information-sharing network

Figure 7(a) illustrates a core-sponsored network. Large firms occupy the core, and all firms sponsor ties to the core. Small firms occupy the periphery and are divided into clusters. Within a cluster, small firms have weak ties. Between clusters, small firms are not directly connected. Large firms, in the core, receive the ties from small firms in the periphery. Figure 7(b) illustrates a periphery-sponsored network. Firms sponsor ties to the periphery, and receive ties from the core. The core-sponsored network satisfies triadic closure, but the periphery-sponsored network does not. As a result, BFE predicts where large firms must be positioned in a core-periphery network.

into n clusters with an equal number $s \equiv \frac{|\mathcal{I}_S|}{n}$ of firms in each cluster. In network c, every firm sponsors a tie with every large firm, large firms receive the ties from the small firms, within a given cluster the small firms receive ties from each other, and small firms do not form ties between clusters. This induces a network with a core-periphery structure (Figure 7a).

The network c is not complete, but satisfies the triadic closure property. While two small firms that are not in the same cluster are not partnered, they also do not receive ties from a common partner unless those partners also share a tie. As a result, the small firms in the periphery of network c are much more loosely connected than the large firms at the center. When $\bar{c} \equiv \gamma_S(\underline{F})$ and $\underline{c} \equiv \gamma_L(\bar{F})$, Condition (16) from the binary formation model again characterizes the RIC. In this example, a large firm sponsors $\ell - 1$ ties and receives ns ties, while a small firms sponsor ℓ ties and receive $s-1$ ties. As such, $\min_i \mathcal{Q}_i(c) = \frac{\ell}{\ell+s-1}$ and $\max \mathcal{Q}_i(c) = \frac{\ell-1}{I-1}$. With a sufficiently large population and $n = 1$, the minimal and maximal ratio are approximately equal to $\frac{1}{2}$, and the RIC is satisfied for $\delta \to 1$. With multiple clusters, the network more closely resembles the core-periphery structure observed empirically, but the minimal and maximal ratios diverge. As a result, for any population size, there is an open set of parameters such that the RIC can be satisfied, but the set of parameters is decreasing in the degree of asymmetry between the two types of firms.

5 Conclusion

We develop a model to study the formation of bilateral ties and enforcement of cooperative behavior on the resulting network. Our analysis primarily focuses on an environment where (i) players cannot pre-commit to cooperation, (ii) the formation of strategically isolated bilateral ties is not incentive compatible, and (iii) players cannot observe the bilateral interactions of others. As a result, cooperative behavior can emerge only when ties are embedded in a network that can pool incentive constraints, and provide credible contagion mechanisms for third-party punishment. In this local monitoring environment, we characterize the network outcomes that can emerge in equilibrium, and provide examples to illustrate when a network can facilitate cooperation. While our characterization for Perfect Bayesian Equilibrium (PBE) highlights constraints on the extent to which sparse networks can sustain cooperation, a PBE allows for a wide range of contagion equilibria, many of which are fragile to beliefs about contagion processes that players will never be able to verify. Motivated by the fragility of the PBE result, we adapt a belief-free equilibrium (BFE) refinement to the network setting. Our main result, Proposition 3, provides necessary and sufficient conditions for a network outcome to emerge in BFE, and shows that the critical condition is an intuitive structural property at the level of relational triads: triadic closure. Our main result can therefore be seen as characterizing a “bound” on the extent to which network structures can be refined through enforcing robustness of equilibria to beliefs about unobserved behavior on the network. With respect to applications, our results provide simple structural predictions at the level of triads (triadic closure) in settings with observable value transfer, or in terms of the global network structure (small-worlds) in settings with observable network patterns. In addition, the flexible heterogeneity in our model, where both the primitive formation technology and endogenous formation investment can generate heterogeneity in ties, allows for a variety of applications to settings where economic institutions (e.g., risk- and information-sharing) depend on how relationships are “embedded” in a network.

References

Ali, S. N. and D. A. Miller (2013): “Enforcing Cooperation in Networked Societies,” *Mimeo University of California at San Diego.*

Balmaceda, F. and J. F. Escobar (2017): “Trust in Cohesive Communities,” *Journal of Economic Theory*, 170, 289–318.

Bernheim, B. and M. Whinston (1990): “Multimarket Contact and Collusive Behavior,” *RAND Journal of Economics*, 21, 1–26.

COX, D. AND M. FAFCHAMPS (2007): "Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions," in *The Handbook of Development Economics*, ed. by T. P. Schultz and J. Strauss, North-Holland, vol. 4, 3711–3784.

DE DOMENICO, M., V. NICOSIA, A. ARENAS, AND V. LATORA (2015): "Structural reducibility of multilayer networks," *Nature communications*, 6, 6864.

DE DOMENICO, M., A. SOLÉ-RIBALTA, E. COZZO, M. KIVELÄ, Y. MORENO, M. A. PORTER, S. GÓMEZ, AND A. ARENAS (2013): "Mathematical formulation of multilayer networks," *Physical Review X*, 3, 041022.

DERCON, S. AND P. KRISHNAN (2003): "Risk Sharing and Public Transfers," *The Economic Journal*, 113, C86–C94.

EASLEY, D. AND J. KLEINBERG (2010): *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press.

EISENHARDT, K. M. AND C. B. SCHOONHOVEN (1996): "Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms," *organization Science*, 7, 136–150.

ELLISON, G. (1994): "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *The Review of Economic Studies*, 61, 567–588.

ELY, J. C., J. HÖRNER, AND W. OLSZEWSKI (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73, 377–415.

ELY, J. C. AND J. VÄLIMÄKI (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102, 84 – 105.

FAFCHAMPS, M. (2008): "Risk Sharing Between Households," in *The Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, Elsevier, vol. 1A, 1255–1279.

FEIGENBERG, B., E. FIELD, AND R. PANDE (2013): "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance," *The Review of Economic Studies*, 80, 1459–1483.

GOYAL, S. (2012): *Connections: An Introduction to the Economics of Networks*, Princeton University Press.

GRANOVETTER, M. S. (1973): "The Strength of Weak Ties," *American Journal of Sociology*, 1360–1380.

Gulati, R. (1995): "Social Structure and Alliance Formation Patterns: A Longitudinal Analysis," *Administrative Science Quarterly*, 619–652.

Haag, M. and R. Lagunoff (2006): "Social Norms, Local Interaction, and Neighborhood Planning," *International Economic Review*, 47, 265–296.

Hörner, J., , S. Lovo, and T. Tomala (2011): "Belief-free equilibria in games with incomplete information: Characterization and existence," *Journal of Economic Theory*, 146, 1770–1795.

Hörner, J. and S. Lovo (2009): "Belief-Free Equilibria in Games With Incomplete Information," *Econometrica*, 77, 453–487.

Jackson, M. and B. Rogers (2005): "The Economics of Small Worlds," *Journal of the European Economic Association*, 3, 617–27.

——— (2007): "Meeting Strangers and Friends of Friends: How Random are Social Networks?" *American Economic Review*, 97, 890–915.

Jackson, M. O. (2009): "Networks and Economic Behavior," *Annual Review of Economics*, 1, 489–511.

Jackson, M. O., T. Rodriguez-Barraquer, and X. Tan (2012): "Social Capital and Social Quilts: Network Patterns of Favor Exchange," *The American Economic Review*, 102, 1857–1897.

Kandori, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, 59, 63–80.

Kivelä, M., A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter (2014): "Multilayer networks," *Journal of complex networks*, 2, 203–271.

Klohn, F. and C. Strupat (2013): "Crowding Out of Solidarity? Public Health Insurance versus Informal Transfer Networks in Ghana," *Ruhr Economic Paper*.

Latora, V., V. Nicosia, and G. Russo (2017): *Complex Networks: Principles, Methods and Applications*, Cambridge University Press.

Lippert, S. and G. Spagnolo (2011): "Networks of Relations and Word-of-Mouth Communication," *Games and Economic Behavior*, 72, 202–217.

Milgram, S. (1967): "The Small-World Problem," *Psychology Today*, 60–67.

Mucha, P. J., T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela (2010): “Community structure in time-dependent, multiscale, and multiplex networks,” *science*, 328, 876–878.

Nava, F. and M. Piccione (2014): “Efficiency in Repeated Games with Local Interaction and Uncertain Local Monitoring,” *Theoretical Economics*, 9, 279–312.

Newman, M., A.-L. Barabasi, and D. J. Watts (2006): *The Structure and Dynamics of Networks*, Princeton, NJ, USA: Princeton University Press.

Piccione, M. (2002): “The Repeated Prisoner’s Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory*, 102, 70 – 83.

Raub, W. and J. Weesie (1990): “Reputation and Efficiency in Social Interactions: An example of Network Effects,” *American Journal of Sociology*, 626–654.

Simmel, G. (1908): *Soziologie: Untersuchungen ueber die Formen der Vergesellschaftung*, Duncker and Humblot.

Stuart, T. E. (1998): “Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry,” *Administrative science quarterly*, 668–698.

Tomasello, M. V., M. Napoletano, A. Garas, and F. Schweitzer (2013): “The Rise and Fall of R&D Networks,” *arXiv preprint arXiv:1304.3623*.

Watts, D. (1999): *Small Worlds*, Princeton, NJ: Princeton University Press.

Watts, D. J. and S. H. Strogatz (1998): “Collective Dynamics of Small-World Networks,” *Nature*, 393, 440–442.

Wolitzky, A. (2013): “Cooperation With Network Monitoring,” *The Review of Economic Studies*, 80, 395–427.

A Proofs

A.1 Preliminaries

We first provide a more detailed description of the dynamic network game and our solution concepts.

An environment is defined by a population $\mathcal{I}$, a radius of information ρ, and parameters $\left(\bar{F}, \underline{F}, \delta, \Gamma\right)$. For a network c, c_i is i's neighborhood (i.e., the set of ties in c on which i is a partner) and $\rho_i(c)$ is the set of ties that are within i's radius of information. Let $\bar{\mathbb{N}} \equiv \mathbb{N} \cup \{\infty\}$; for $\delta \in (0,1)$, $\delta^\infty \equiv 0$.

We denote by $h^\emptyset$ the initial history when no actions have been taken. A 0-*history* h^0 is a record of actions in the formation phase. For $t \geq 1$, a t-*history* h^t is a record of actions by every player in periods $\tau < t$ (the formation phase and the continuation phase prior to period t). An ∞-*history* defines a complete path of play for the game. Denote by $\mathcal{H}^t$ the set of t-histories. We say that history $h^{t+\tau}$ is *consistent* with h^t, denoted $h^{t+\tau} \trianglerighteq h^t$, if all actions recorded in history $h^{t+\tau}$ prior to period t coincide with the actions recorded in h^t.

For radius of information ρ and a formation strategy-profile F, player $i \in \mathcal{I}$ can observe F_{jk} in the formation phase if and only if $\max\left\{d_{c(F)}(i,j), d_{c(F)}(j,k)\right\} \leq \rho$. Likewise, in the continuation phase on a network c, player i can observe A^t_{jk} if and only if $\max\{d_c(i,j), d_c(j,k)\} \leq \rho$. We formalize monitoring constraints with an equivalence relation on t-histories. Two t-histories h^t and $\tilde{h}^t$ are i-*equivalent*, denoted $h^t \sim_i \tilde{h}^t$, if player i cannot distinguish h^t and $\tilde{h}^t$ given the radius of information. We call an equivalence class of $\sim_i$ an it-*history*, and denote the set of it-histories by $\mathcal{H}^t_i$ with typical element h^t_i.

A t-history h^t induces a current network $c(h^t)$, which gives the ties that were formed during the formation phase of h^t and then continued in every period of the continuation phase prior to period t; $c_i(h^t)$ is player i's neighborhood in the current network; and $c_{ij}(h^t)$ denotes player i's tie with player j. For an it-history h^t_i, player i's network neighborhood is the same in every t-history $h^t \in h^t_i$. Hence, we also denote by $c_i(h^t_i)$ and $c_{ij}(h^t_i)$ the network neighborhood, and tie between players i and j, uniquely induced by the it-history h^t_i.

In the formation phase, the strategy of player i consists of unilateral investment decisions $F_i = (F_{ij})_{j\in\mathcal{I}}$ in ties with other players, where $F_{ii} \equiv \emptyset$ by default. In the continuation phase, a strategy of player i is a mapping, $\sigma_i : \mathcal{I} \times \bigcup_{t=0}^\infty \mathcal{H}^t \to \{C, D\}$, from opponents and t-histories into actions in period t. The continuation strategy must satisfy two conditions: (i) $ij \notin c\left(h^t\right)$ implies $\sigma_i\left(j, h^t\right) = D$, and (ii) $h^t \sim_i \tilde{h}^t$ implies $\sigma_i\left(., h^t\right) = \sigma_i\left(., \tilde{h}^t\right)$. Condition (i) means that players can cooperate only on ties that exist in the current network. Condition (ii) means that players can condition actions in period t only on the part of a t-history that is observable to them given the radius of information.

In the formation phase, the strategy profile (F, σ) induces a unique period 0-history h^0. For any t-history h^t, the strategy (F, σ) induces a unique $(t+1)$-history, denoted $h^{t+1}\left(F, \sigma | h^t\right)$. Recursively, the strategy (F, σ) therefore induces a unique $(t+\tau)$-history $h^{t+\tau}\left(F, \sigma | h^t\right)$. As a result, the strategy (F, σ) induces a *continuation path* for any t-history h^t, denoted $h^\infty\left(F, \sigma | h^t\right)$ (let $h^\infty\left(F, \sigma | h^\emptyset\right)$ denote the continuation path induced by (F, σ) prior to the formation phase).

For $t \geq \tau$, let $c^t(F,\sigma|h^\tau) \equiv c^t(h^t(F,\sigma|h^\tau))$ denote the network induced in period t by the continuation path of (F,σ) given history h^τ. In particular, $c(F,\sigma) = c$ if the continuation path of play $h^\infty\left(F,\sigma|h^\emptyset\right)$ induces the network c in every period $t \geq 0$.

In period $t \geq 0$ of the continuation phase, player i observes an it-history h_i^t and has beliefs about the t-history. Denote by $\Delta(\mathcal{H}^t)$ the set of all probability distributions on $\mathcal{H}^t$. Player i's *beliefs* are described by a mapping $\mu_i : \bigcup_{t=0}^\infty \mathcal{H}_i^t \to \bigcup_{t=0}^\infty \Delta(\mathcal{H}^t)$, where $\mu_i(h^t|h_i^t)$ is the probability that player i attaches to the t-history h^t given she observes the it-history h_i^t. A *system of beliefs* is a tuple $\mu = (\mu_i)_{i\in\mathcal{I}}$, specifying beliefs for every player. A system of beliefs μ is *Bayesian consistent* with strategy profile (F,σ) if, for all $i \in \mathcal{I}$, (i) player i does not believe a history has occurred that is inconsistent with the it-history she observes, and (ii) player i believes that other players have followed strategy (F_{-i},σ_{-i}) unless she observes an it-history that is inconsistent with (F_{-i},σ_{-i}). Formally, let $(\tilde{F}_i, F_{-i}, \tilde{\sigma}_i, \sigma_{-i})$ be the the strategy profile where player i follows $(\tilde{F}_i,\tilde{\sigma}_i)$ and players other than i follow (F_{-i},σ_{-i}). Then (i) means that $h^t \notin h_i^t$ implies $\mu_i(h^t|h_i^t) = 0$, and (ii) means that, for all strategies $\left(\tilde{F}_i,\tilde{\sigma}_i\right)$ of player i, $h^t\left(\tilde{F}_i, F_{-i}, \tilde{\sigma}_i, \sigma_{-i}|h^\emptyset\right) \in h_i^t$ implies that $\mu_i\left(h^t\left(\tilde{F}_i, F_{-i}, \tilde{\sigma}_i, \sigma_{-i}|h^\emptyset\right)|h_i^t\right) = 1$.

For player $i \in \mathcal{I}$, π_i maps any continuation path to player i's continuation payoff, i.e., $\pi_i(F,\sigma|h^t)$ is player i's continuation payoff under strategy profile (F,σ) following history h^t. Expected payoffs and best responses are defined with respect to a system of beliefs μ. For player $i \in \mathcal{I}$, the *expected continuation payoff* under strategy profile (F,σ) given $h^\emptyset$ is $\pi_i\left(F,\sigma|h^\emptyset\right)$, and given an it-history h_i^t is $E_\mu[\pi_i(F,\sigma|h_i^t)] = \sum_{h^t\in\mathcal{H}^t}\mu_i(h^t|h_i^t)\,\pi_i(F,\sigma|h^t)$. The strategy (F_i,σ_i) is a (μ,h_i^t)-*best response* to (F_{-i},σ_{-i}) if

$$E_\mu\left[\pi_i\left(F,\sigma \mid h_i^t\right)\right] \geq E_\mu\left[\pi_i\left(\tilde{F}_i, F_{-i}, \tilde{\sigma}_i, \sigma_{-i} \mid h_i^t\right)\right] \qquad \forall\,(\tilde{F}_i,\tilde{\sigma}_i). \tag{17}$$

The strategy profile (F_i,σ_i) is a μ-*best response* to (F_{-i},σ_{-i}) if, for every it-history h_i^t, (F_i,σ_i) is a (μ,h_i^t)-best response to (F_{-i},σ_{-i}). A strategy (F_i,σ_i) is therefore a best response if it maximizes player i's expected continuation payoff at every information set, given that other players follow (F_{-i},σ_{-i}). We can now define our solution concepts.

Definition 5 (Equilibrium)**.** A strategy profile (F,σ) is a PBE if, for every player i, (F_i,σ_i) is a μ-best response for some system of beliefs μ that is Bayesian consistent with (F,σ). A strategy profile (F,σ) is a BFE if, for every player i, (F_i,σ_i) is a μ-best response for every system of beliefs μ that is Bayesian consistent with (F,σ).

A.2 Continuation strategies

We start by defining some strategy-profiles for the continuation phase of the game, which we use in the proofs of Propositions 1--3.

We fix a "target" network $\tilde{c}$, and start by defining a grim-strategy $\sigma^{\tilde{c}}$ for a game with partial monitoring as follows: for all $i, j \in \mathcal{I}$ and any history h^t,

$$\sigma_{ij}^{0\tilde{c}}\left(h^t\right) = \begin{cases} C & \text{if } ij \in c_i(h^t) \text{ and } \rho_i(c(h^t)) = \rho_i(\tilde{c}) \\ D & \text{otherwise} \end{cases}.$$

In words, players cooperate with partners in any history h^t where there is a tie between these players and the current network they observe is the target network $\tilde{c}$, otherwise players do not cooperate with anyone. For the game with local monitoring, we adjust the grim-strategy so that, if j's only partner is player i, i continues to cooperates with j after observing a deviation:

$$\sigma_{ij}^{1\tilde{c}}\left(h^t\right) = \begin{cases} C & \text{if } ij \in c_i(h^t) \text{ and } \left(\rho_i(c(h^t)) = \rho_i(\tilde{c}) \text{ or } \tilde{c}_j = \{ij\}\right) \\ D & \text{otherwise} \end{cases}.$$

With local monitoring, regardless of incentives, the strategy-profile $\sigma^{1\tilde{c}}$ is generally not a PBE because it does not distinguish why player i's network neighborhood differs from the target network. First, it could be that a player other than i has deviated from a cooperative path of play. In that case, player i's beliefs about the history on ties she cannot monitor are unrestricted, and it is straightforward to find a Bayesian consistent system of beliefs that rationalizes the grim-strategy response. However, it is also possible that player i's network neighborhood differs from the target network because player i has deviated from a cooperative path of play. In that case, Bayesian consistency restricts player i's beliefs and, for a consistent set of beliefs, the grim-strategy is not generally a best-response. To formulate a strategy-profile for a PBE in Proposition 2, we need to adjust the strategy-profile $\sigma^{1\tilde{c}}$ to distinguish between these cases.

We use $\sigma^{1\tilde{c}}$ to define an alternative strategy profile $\sigma^{2\tilde{c}}$, which we describe from the perspective of player i. Fix some history h^t, and let h_i^t be the corresponding it-history. First, suppose that in h_i^t player i's network neighborhood coincides with the target network neighborhood $\tilde{c}_i$ (i.e., $c_i(h^t) = \tilde{c}_i$), then let $\sigma^{2\tilde{c}}(h^t) = \sigma^{1\tilde{c}}(h^t)$ (i.e., player i cooperates on all ties). Now suppose that in h_i^t player i's network neighborhood differs from the target network neighborhood (i.e., $c_i(h^t) \neq \tilde{c}_i$). We distinguish two cases. In the first case, the actions observed by player i in h_i^t are inconsistent with other players following the strategy profile $\left(F_{-i}^{\tilde{c}}, \sigma_{-i}^{1\tilde{c}}\right)$, given the actions chosen by player i. That means, given the actions chosen by player i in h_i^t, player i has observed some action (in the formation or continuation phase) by at least one other player that would not have been observed if other players were following the strategy profile $\left(F^{\tilde{c}}_{-i}, \sigma^{1\tilde{c}}_{-i}\right)$. In that case, player i defects on all of her ties in period t, except ties to a partner whose only tie is with i (i.e., $\sigma^{2\tilde{c}}_{ij}(h^t) = D$ for all $j \neq i$ such that $\tilde{c}_j \neq \{ij\}$). In the second case, the actions observed by player i in h^t_i are consistent with other players following the strategy profile $\left(F^{\tilde{c}}_{-i}, \sigma^{1\tilde{c}}_{-i}\right)$, given the actions chosen by player i. In that case, we describe player i's actions by means of a defection plan $\tilde{\alpha}^i = (\tilde{\alpha}^i_j)_{j\neq i} \in \bar{\mathbb{N}}^{I-1}$. As in Section 3.3.1, the interpretation is that $\tilde{\alpha}^i_j$ is the (first) period in which player i defects against player j in the continuation phase. Let $\mathcal{A}^{\tilde{c}}_i(h^t) \subseteq \bar{\mathbb{N}}^{I-1}$ be the set of defection plans α^i such that, for all $j \neq i$, (i) if ij is not formed in h^t, then $\alpha^i_j = 0$; (ii) if ij is formed in h^t, but the tie differs from $(\tilde{c}_{ij}, \tilde{c}_{ji})$, then $\alpha^i_j = 0$; (iii) if ij is formed in h^t and the tie coincides with $(\tilde{c}_{ij}, \tilde{c}_{ji})$, but player i defected on ij for the first time in some period $\tau < t$, then $\alpha^i_j = \tau$, and (iv) $\alpha^i_j \leq \min_{k\in\mathcal{N}_i(\tilde{c})-\{j\}} \{\alpha^i_k + d_{\tilde{c}}\,[i, j|\neg\{ik\}] - 1\}$. Hence, $\mathcal{A}_i(h^t)$ consists of defection plans where $\alpha^i_j \geq t$ if only if there is currently a tie between players i and j and that tie coincides with the tie that "should" exist between these players in the target network. The defection plan $\tilde{\alpha}^i$ is chosen from $\mathcal{A}^{\tilde{c}}_i(h^t)$ to solve the following problem:

$$\max_{\alpha^i \in \mathcal{A}^{\tilde{c}}_i(h^t)} \sum_{j\neq\mathcal{N}_i(c(h^t))} \left(\frac{(1-\delta^{\alpha^i_j})(1-c_{ij})}{1-\delta} + \delta^{\alpha^i_j} \right). \tag{18}$$

In words, player i considers all possible time periods in which they could defect against their remaining partners, taking into account how quickly defection spreads through the network under the strategy profile $\left(F^{c}_{-i}, \sigma^{1\tilde{c}}_{-i}\right)$ given the it-history h^t_i.

To see that the optimization problem in (18) is well-defined, partition the set of players $\mathcal{I} - \{i\}$ into $(\mathcal{I}_l)_{l=1,\dots,L}$ such that $j, k \in \mathcal{I}_l$ for some l if and only if there is a path from player j to player k in network $\tilde{c}$ that does not pass through player i (i.e., $d_{\tilde{c}}\,[j, k|\neg\tilde{c}_i] < \infty$). This partition is well-defined, and the optimal defection plan can treat these parts separately. It is sufficient for us to consider the parts $\mathcal{I}_l$ where some node has a tie with player i in period t. First, suppose player i defected on some node in $\mathcal{I}_l$ in a previous period. Suppose that all players in $\mathcal{I}_l$ follow the strategy profile $(F^{\tilde{c}}, \sigma^{1\tilde{c}})$, then in some (finite) period $\tau \geq t$, all ties between players in $\mathcal{I}_l$ have seen a defection. Hence, there is a finite number of possible choices for α^i_j when $j \in \mathcal{I}_l$, which can be used to determine an optimal defection plan relative to players in $\mathcal{I}_l$. Second, suppose that player i has not defected on any node in $\mathcal{I}_l$ in a previous period. Then player i could decide not to defect on any of the ties with players in $\mathcal{I}_l$ (yielding some finite continuation value). Or, player i could defect in period t in a tie with some players in $\mathcal{I}_l$. In the second case, there is again a finite time when all ties are discontinued, and so there is a finite number of choices for α^i_j. A comparison of the continuation value for continuing on all ties indefinitely, or the optimal way to defect on ties in $\mathcal{I}_l$ in period t, determines an optimal defection plan for player i relative to the players in $\mathcal{I}_l$. As a result, there is an optimal defection plan. If the plan is not unique, any optimal plan can define the strategy for player i.

The solution to the optimization problem also has the following property. Suppose $\tilde{\alpha}^i$ solves (18) given the history h^t. In period t, player i cooperates with player j if $\tilde{\alpha}^i_j > t$, and defects against player j if $\tilde{\alpha}^i_j \leq t$. If other players follow the strategy profile $(F^{\tilde{c}}, \sigma^{1\tilde{c}})$, this generates a history in period $t+1$, h^{t+1}, and $\tilde{\alpha}^i$ again solves

$$\max_{\alpha^i \in \mathcal{A}^{\tilde{c}}_i(h^{t+1})} \sum_{j \neq i} \left(\frac{(1-\delta^{\alpha^i_j})(1-c_{ij})}{1-\delta} + \delta^{\alpha^i_j} \right).$$

We can now complete the description of $\sigma^{2\tilde{c}}_i$. In a history h^t where the actions observed by player i in h^t_i are consistent with other players following the strategy profile $\left(F^{\tilde{c}}_{-i}, \sigma^1_{-i}\right)$, given the actions chosen by player i, player i cooperates with player j in period t if and only if $\tilde{\alpha}^i_j > t$ (i.e., for all $j \neq i$,$\sigma^{2\tilde{c}}_{ij}(h^t) = C$ if and only if $\tilde{\alpha}^i_j > t$ for the optimal defection plan $\tilde{\alpha}^i \in \mathcal{A}_i(h_t)$).

The following lemma follows from the definition of the strategy profiles σ^{0c}, σ^{1c}, and σ^{2c}.

Lemma 1. *Consider a network c. (i) In a game with partial monitoring $c\left(F^c, \sigma^{0c}\right) = c$. (ii) In a game with local monitoring, $c\left(F^c, \sigma^{1c}\right) = c$. (iii) In a game with local monitoring, $c\left(F^c, \sigma^{2c}\right) = c$.*

A.3 Proof of propositions in Section 3

Proof of Proposition 1

We first show that the PIC is sufficient for a BFE, and then show that the PIC is necessary for a PBE. Since BFE is a refinement of PBE, this establishes that the PIC characterizes the equilibrium network outcomes that can emerge under both solution concepts.

Part [i]: If c satisfies the PIC, then (F^c, σ^{0c}) is a BFE.

Consider the continuation phase. First, let h^t be any history such that $\rho_i(c(h^t)) \neq \rho_i(c)$ (i.e., player i has observed a deviation). Given that other players follow $(F^c_{-i}, \sigma^{0c}_{-i})$, all players other than i who have observed this deviation will discontinue all their ties in $t+1$. As a result, all of player i's neighbors will observe a deviation from the equilibrium path in $t+1$, and discontinue their ties with i in $t+2$ (if they have not already discontinued ties with i in $t+1$). It is therefore optimal for i to discontinue all ties in $t+1$. This means that, whenever any player i observes a deviation, (F^c_i, σ^{0c}_i) is a belief-free best-response to $(F^c_{-i}, \sigma^{0c}_{-i})$. Second, let h^t be any history such that $\rho_i(c(h^t)) = \rho_i(c)$. Given the strategy-profile (F^c, σ^{0c}), there is a unique Bayesian consistent belief for player i where $\mu_i(\tilde{h}^t | h^t_i) = 1$ for the unique history $\tilde{h}^t$ such that $c(\tilde{h}^t) = c(\tilde{h}^0) = c$. For the history $\tilde{h}^t$, the strategy profile (F^c, σ^{0c}) prescribes that all players cooperate on existing ties, and continue to cooperate unless they observe a defection. Player i's continuation payoff under this strategy profile is $\sum_{j\in\mathcal{N}_i(c)}\left(\frac{1-c_{ij}}{1-\delta}\right)$.

Given the strategy profile $\left(F^c_{-i},\sigma^{0c}_{-i}\right)$ of opponents, player i's optimal deviation is to defect on all ties in the network simultaneously: if she defects on any subset of her ties in t, all neighbors will discontinue their ties with i in $t+1$. Player i's continuation payoff from defecting on all ties is $|\mathcal{N}_i(c)|$, which is less than the continuation payoff from following the strategy profile (F^c_i,σ^{0c}_i) because c satisfies the PIC.

Now consider the formation phase. Under that strategy profile (F^c,σ^{0c}), player i's continuation payoff is $\sum_{j\in\mathcal{N}_i(c)}\left[\left(\frac{1-c_{ij}}{1-\delta}\right)-F^c_{ij}\right]$. It is not optimal for player i to choose an alternative strategy F'_i that, given F^c_{-i}, also induces the network c, since F^c_i is the least costly way to induce the target network. Hence, given the strategy profile $\left(F^c_{-i},\sigma^{0c}_{-i}\right)$, player i's optimal deviation is to form no ties at all (i.e., $F_{ij}=0$ for all $j\neq i$): if she chooses an alternative strategy F'_i in which $c\left(F'_i,F^c_{-i}\right)\neq c$, then all ties will be discontinued in the first round of the continuation phase; hence, her continuation payoff is bounded above by $-F'_i$ which is less than the payoff of forming no ties (i.e., 0). However, since c satisfies the PIC, the continuation payoff for player i under the strategy profile (F^c,σ^{0c}) is non-negative. As a result, there is no profitable deviation for player i in the formation phase. Since no player has a profitable deviation in any history h^t, (F^c,σ^{0c}) is a BFE. It follows from Lemma 1 that if c satisfies the PIC, there is a BFE (F,σ) such that $c(F,\sigma)=c$. □

Part [ii]: If there is a PBE (F,σ) such that $c(F,\sigma)=c$, then c satisfies the PIC.

By way of contradiction, suppose (F,σ) is a PBE where $c(F,\sigma)=c$, and yet c does not satisfy the PIC. Then, for some player $i\in\mathcal{I}$, either (i) $|\mathcal{N}_i(c)|>\sum_{j\in\mathcal{N}_i(c)}\left(\frac{1-c_{ij}}{1-\delta}\right)$, or (ii) $\sum_{j\in\mathcal{N}_i(c)}\left[\left(\frac{1-c_{ij}}{1-\delta}\right)-F^c_{ij}\right]<0$. In case (i), the strategy profile σ must prescribe for player i to cooperate on all ties in $\mathcal{N}_i(c)$ on the path of play, giving a maximum continuation payoff of $\sum_{j\in\mathcal{N}_i(c)}\left(\frac{1-c_{ij}}{1-\delta}\right)$. However, player i can guarantee a continuation payoff of $|\mathcal{N}_i(c)|$ by defecting on all ties. Hence, player i has a profitable deviation in the continuation phase of the game. In case (ii), the maximum continuation payoff for player i in the formation phase is

$$\sum_{j\in\mathcal{N}_i(c)}\left[\left(\frac{1-c_{ij}}{1-\delta}\right)-F_{ij}\right]\leq\sum_{j\in\mathcal{N}_i(c)}\left[\left(\frac{1-c_{ij}}{1-\delta}\right)-F^c_{ij}\right].$$

However, player i can guarantee a continuation payoff of 0 by forming no ties (i.e., $F_{ij}=0$ for all $j\neq i$). Hence, player i has a profitable deviation at the formation stage. Hence, at least one player has a profitable deviation from (F,σ) in either the formation or continuation phase of the game. □

Proof of Proposition 2

Part [i]: If c satisfies the LIC, then (F^c, σ^{2c}) is a PBE.

To establish that (F^c, σ^{2c}) is a PBE we must find a system of beliefs μ that is consistent with (F^c, σ^{2c}) such that each player is playing a best-response for each observable history. Consider the following system of beliefs μ. In any it-history h_i^t in which player i has observed actions by others that – given player i's own previous actions – are consistent with other players following the strategy profile $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$, player i believes with probability one that all other players have followed the actions prescribed by the strategy profile (F^c, σ^{2c}) (given player i's own actions). In any it-history h_i^t in which player i has observed an action by others that – given player i's own previous actions – is inconsistent with other players following the strategy profile $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$, player i believes with probability one that the history that has occurred is the one in which all actions she actually observed took place, and $F_{jk} = 0$ for all $j, k \neq i$. This system of beliefs is Bayesian consistent with the strategy profile (F^c, σ^{2c}).

To show that (F^c, σ^{2c}, μ) is a PBE, we consider possible deviations for a player $i \in \mathcal{I}$. First, consider the continuation phase. Let h_i^t be an it-history where $c_i(h_i^t) = c_i$. In this case, $\mu_i(h^t | h_i^t) = 1$ for the history h^t where $c(h^t) = c$. Hence, player i's optimal defection plan solves

$$\max_{\alpha^i \in \mathcal{A}(h^t)} \sum_{j \in \mathcal{N}_i(c)} \left(\frac{(1 - \delta^{\alpha_j^i})(1 - c_{ij})}{1 - \delta} + \delta^{\alpha_j^i} \right),$$

which has the unique solution [$\alpha_j^i = 0$ if $ij \notin c$, $\alpha_j^i = \infty$ if $ij \in c$], because c satisfies the LIC. Hence, $\sigma^{2c}(h_i^t)$ is a μ-best response to $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$ for any it-history h_i^t where $c_i(h_i^t) = c$.

Now let h_i^t be an it-history where $c_i(h_i^t) \neq c_i$. We distinguish two cases. In the first case, the actions observed by player i in h_i^t are consistent with other players following $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$ given the previous actions by player i. Then there is a unique history h^t such that $\mu(h^t | h_i^t) = 1$. Given this history, and the strategy $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$ followed by others, $\sigma_i^{2c}(h^t)$ is a μ-best response by definition (since it specifies for player i to follow a defection plan that is optimal when other players follow $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$ and the history is h^t).

In the second case, the actions observed by player i in h_i^t are inconsistent with other players following $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$, given the previous actions by player i. For the strategy profile (F^c, σ^{2c}), this is a probability 0 event, and $\mu(h^t | h_i^t) = 1$ for the history $h^t \in h_i^t$ where $F_{jk} = 0$ for all $j, k \neq i$. In this history h^t, (F^c, σ^{2c}) specifies that all players who have a tie with i and at least one other tie should defect on their ties with player i in period t (and all subsequent periods). Hence, it is optimal for player i to defect on all of these remaining ties in period t. If player j has only one tie, which is with player i, then j has not observed a deviation from (F^c, σ^{2c}) and continues the tie with i in period $t+1$; it is therefore optimal for i to also continue the tie because c satisfies the LIC (which implies that the tie ij satisfies the BIC). As a result, $\sigma_i^{2c}(h^t)$ is a μ-best response to $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$ for the it-history h_i^t. It follows that σ_i^{2c} is a μ-best response in every it-history h_i^t of the continuation phase.

Finally, consider the formation phase. Under that strategy profile (F^c, σ^{2c}), player i's continuation payoff is $\sum_{j \in \mathcal{N}_i(c)} \left[\left(\frac{1-c_{ij}}{1-\delta}\right) - F_{ij}^c\right]$. It is not optimal for player i to choose an alternative formation strategy F_i' that, given F_{-i}^c, also induces the network c, since F_i^c is the least costly way to induce the target network c. Hence, given the strategy profile $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$, player i's μ-best-response in history $h_i^t = \emptyset$ solves $\max_{\alpha^i \in \mathcal{A}_i^c(\emptyset)} \sum_{j \in \mathcal{I}-\{i\}} \Pi_j^i(\alpha_j^i)$, where $\mathcal{A}_i^c(\emptyset)$ is the set of defection plans in $\left(\{-1\} \cup \bar{\mathbb{N}}\right)^{I-1}$ that satisfy the constraints (i) and (ii) in the definition of $\mathcal{A}_c^i$ in Section 3.3.1. Since c satisfies the LIC, one optimal plan is α^i, where $\alpha_j^i = -1$ if $ij \notin c$ and $\alpha_j^i = \infty$ otherwise. Hence, player i does not have a profitable deviation from F_i^c in the formation phase. Since no player has a profitable deviation from (F^c, σ^{2c}) after any it-history given the beliefs μ, (F^c, σ^{2c}) is a PBE. It follows from Lemma 1 that if c satisfies the LIC, then there is a PBE (F, σ) such that $c(F, \sigma) = c$. □

Part [ii]: If there is a PBE (F, σ) such that $c(F, \sigma) = c$, then c satisfies the LIC.

By way of contradiction, suppose (F, σ) is a PBE where $c(F, \sigma) = c$, and yet c does not satisfy the LIC. Then on the path of play for the strategy profile (F, σ), the network c is formed in the formation phase, and all ties are continued indefinitely. For each player i, the continuation payoff at the formation phase is then

$$\sum_{j \in \mathcal{N}_i(c)} \left[\left(\frac{1-c_{ij}}{1-\delta}\right) - F_{ij}\right] \leq \sum_{j \in \mathcal{N}_i(c)} \left[\left(\frac{1-c_{ij}}{1-\delta}\right) - F_{ij}^c\right],$$

and for any system of beliefs μ that is Bayesian consistent with (F, σ) and in any history h^t consistent with the path of play, the continuation payoff in the continuation phase is $\sum_{j \in \mathcal{N}_i(c)} \left(\frac{1-c_{ij}}{1-\delta}\right)$. However, for some player $i \in \mathcal{I}$, there exists a defection plan $\alpha^i \in \mathcal{A}_i^c$ such that

$$\sum_{j \neq i} \Pi_j^i(\alpha_j^i) > \sum_{j \in \mathcal{N}_i(c)} \left[\left(\frac{1-c_{ij}}{1-\delta}\right) - F_{ij}^c\right].$$

But the defection plan α^i is feasible for player i, when others follow (F_{-i}, σ_{-i}), since the constraint set $\mathcal{A}_i^c$ takes into account the fastest way that defection can spread through network c under local monitoring. Moreover, the defection plan α^i specifies $\alpha_j^i < \infty$ for some $j \in \mathcal{N}_i(c)$, and therefore represents a deviation from the path of play for (F, σ). Hence, there exists a history h^t that is consistent with the path of play such that (i) for any belief μ that is Bayesian consistent with (F, σ), $\mu(h^t | h_i^t) = 1$, and (ii) player i has a profitable deviation given (F_{-i}, σ_{-i}) in history h^t. This contradicts that (F, σ) is a PBE. □

Proof of Proposition 3

Part [i]: If c satisfies triadic closure and the RIC, then (F^c, σ^{1c}) is a BFE.

We consider possible deviations by player $i \in \mathcal{N}(c)$ for a system of beliefs μ that is Bayesian consistent with (F^c, σ^{1c}). Start with the continuation phase, and consider an it-history h_i^t where $c_i(h_i^t) = c_i$. Then there exists a unique history $h^t \in h_i^t$ such that $\mu(h^t|h_i^t) = 1$, and $c(h^t) = c$. If player i follows $\sigma_i^{1c}(h^t)$, the expected continuation payoff is $\sum_{j \in \mathcal{N}_i(c)} \left(\frac{1-c_{ij}}{1-\delta}\right)$. We argue that if it is optimal for player i to discontinue any tie in period t, then it is optimal for player i to discontinue all ties where i is a recipient in period t, and all ties where she is a strict sponsor in period $t+1$. Note that on ties where player i is a recipient, the payoff from discontinuing the tie in t (i.e., 1) exceeds the benefit of continuing the tie for any period of time, while on ties where player i is a sponsor, it is optimal to continue the tie as long as possible before receiving the payoff 1 from an (unanticipated) defection. Now suppose player i discontinues a tie in period t with player $j \in \mathcal{N}_i(c)$. Since c satisfies triadic closure, $jk \in c(h^t)$ for any $k \in \bar{\mathcal{S}}_i(c)$. The strategy (F^c, σ^{1c}) prescribes for player j to discontinue all ties in period $t+1$ after observing player i's defection in period t. Hence, any player $k \in \bar{\mathcal{S}}_i(c)$ observes a deviation from the path of play in period $t+1$; (F^c, σ^{1c}) then prescribes for player k to discontinue the tie ik in period $t+2$. Anticipating this, player i should plan to discontinue all ties where she is a strict sponsor in period $t+1$. Given that player i will discontinue all ties where she is a strict sponsor in latest period $t+1$, it is optimal to discontinue all ties where player i is a recipient in period t. The expected continuation payoff from this defection plan is

$$|\mathcal{R}_i(c)| + \delta|\bar{\mathcal{S}}_i(c)| + \sum_{j \in \mathcal{S}_i(c)} (1 - c_{ij}) = |\mathcal{N}_i(c)| + \delta|\bar{\mathcal{S}}_i(c)| - \sum_{j \in \mathcal{S}_i(c)} c_{ij},$$

which is less than $(1-\delta)^{-1} \sum_{j \in \mathcal{N}_i(c)} (1 - c_{ij})$ because c satisfies the RIC.

Now consider an it-history h_i^t where $c_i(h_i^t) \neq c_i$. Then there exists a partner j of player i such that for any history $h^t \in h_i^t$, $c_j(h_j^t) \neq c_j$. The strategy (F^c, σ^{1c}) therefore prescribes for player j to discontinue all their ties in period t. Since c satisfies triadic closure, $jk \in c$ for any $k \in \bar{\mathcal{S}}_i(c)$. Any player $k \in \bar{\mathcal{S}}_i(c)$ will therefore observe a deviation from the path of play in period t, and (as prescribed by (F^c, σ^{1c})) will discontinue the tie ik latest in period $t+1$. Anticipating this, it is optimal for player i to discontinue all ties where she is a strict sponsor in period t, and since she is discontinuing all ties where she is a strict sponsor it is also optimal to discontinue all ties where she is a recipient.

As a result, there is no optimal deviation from (F^c, σ^{1c}) for player i in any it-history. The argument uses only the assumption that the system of beliefs μ is Bayesian consistent with (F^c, σ^{1c}), and so holds for any system of beliefs that is Bayesian consistent with (F^c, σ^{1c}). The strategy profile (F^c, σ^{1c}) is therefore a BFE. It follows from Lemma 1 that if c satisfies triadic closure and RIC, there is a BFE (F, σ) such that $c(F, \sigma) = c$. $\square$

Part [ii]: If there exists a BFE (F, σ) such that $c(F, \sigma) = c$, then there exists a subnetwork $\tilde{c} \subseteq c$ such that (i) $\tilde{c}$ contains all weak and asymmetric ties in c, and (ii) $\tilde{c}$ satisfies triadic closure and the RIC.

Let (F, σ) be a BFE such that $c(F, \sigma) = c$. We will construct the subnetwork $\tilde{c}$ with the desired properties. We say that a tie $ij \in c$ is *eventually strategically independent* of $c - \{ij\}$ if there exists a history h^t that is consistent with the equilibrium path of play such that, for all $h^\tau \trianglerighteq h^t$, $\sigma_{ij}(h^\tau) = \sigma_{ji}(h^\tau) = C$ as long as $ij \in c(h^\tau)$. That is, for every history following h^t, players i and j continue tie ij independently of what they have observe other players doing on other ties. Denote by $\bar{c}$ the subset of ties $ij \in c$ that are *not* eventually strategically independent of $c - \{ij\}$. A tie can be eventually strategically independent only if both partners sponsor the tie, and so $\bar{c}$ contains all weak and asymmetric ties in c. The following claim establishes that in a BFE, players must discontinue any tie in $\bar{c}$ in any period when they observe a deviation from the equilibrium path of play.

Claim 1. If $ij \in \bar{c}$, then in any t-history h^t where player i has not deviated from (F_i, σ_i) but $c_i(h^t) \neq c_i$, $\sigma_{ij}(h^t) = D$.

Proof. If $ij \notin c_i(h^t)$ then $\sigma_{ij}(h^t) = D$ is a restriction on the strategy space, so assume that $ij \in c_i(h^t)$. If player i has followed (F_i, σ_i) up to period t and $c_i(h^t) \neq c_i$, then player i has observed a deviation from the equilibrium path of play. As a result, there is a system of beliefs μ that is Bayesian consistent with (F, σ) where $\mu_i\left(h^t | h_i^t\right) = 1$ for a history $h^t \in h_i^t$ such that $c(h^t) = c_i(h^t)$ (i.e., there are no ties in the network, other than the ties in i's network neighborhood). If the network in period t is $c_i(h_i^t)$, player i should discontinue any ties on which she is a strict recipient. Moreover, any partner of player i who is a strict recipient should discontinue their tie with player i. Hence, player i should also discontinue any asymmetric ties on which she is the sponsor. Hence, if ij is a weak or asymmetric tie, then $\sigma_{ij}(h_i^t) = D$. So it remains to consider the case where both partners sponsor ij. If the current network is $c_i(h^t)$ player i should discontinue the tie ij in period t if player j's strategy prescribes for her to discontinue ij in either period t or $t+1$, otherwise i should continue ij in period t. If player i believes (with probability 1) the network in period t is $c_i(h^t)$, then player j has also observed a deviation from the path of play, and so her position is symmetric. There are two possibilities: (1) the strategy prescribes for both players to continue ij indefinitely, (2) the strategy prescribes for one player to discontinue the tie in finite time. In case (2), it follows by a simple backward induction argument that player i should discontinue tie ij in period t. Now consider case (1). There is also a Bayesian consistent system of beliefs where $\mu_i(h'^t | h_i^t) = 1$ for a history h'^t such that $c(h'^t) \supseteq c_j$, and player j has observed no deviation from the path of play. Moreover, for any history $h'^\tau \trianglerighteq h'^t$, there is a Bayesian consistent system of beliefs where $\mu_i(h'^\tau|h_i'^\tau) = 1$. As a result, since (F, σ) is a BFE, it must be the case that for any history $h'^\tau \trianglerighteq h'^t$, $\sigma_{ij}(h'^\tau) = C$; otherwise, by backward induction, player i should discontinue tie ij in $\tau - 1$, contradicting that both players continue ij indefinitely. In the history h^t, the situation is symmetric for player j. As a result, for the history $\tilde{h}^t$ which is consistent with the path of play (up to period t), it follows that for all $\tilde{h}^\tau \trianglerighteq \tilde{h}^t$, players i and j cooperate on ij in period τ (as long as $ij \in c\left(\tilde{h}^\tau\right)$). The tie ij is therefore eventually strategically independent, and is therefore not in $\bar{c}$. $\square$

The following claims show that there exists a subnetwork $\tilde{c}$ such that (i) $\bar{c} \subseteq \tilde{c} \subseteq c$, (ii) $\tilde{c}$ satisfies triadic closure, and (iii) $\tilde{c}$ satisfies the RIC. Since $\bar{c}$ contains all weak and asymmetric ties, this completes the proof.

Claim 2. Any subnetwork c' such that $\bar{c} \subseteq c' \subseteq c$ satisfies the RIC.

Proof. To show that, for each $i \in \mathcal{N}(c)$,

$$\frac{\frac{1}{\delta}\sum_{j\in\mathcal{R}_i(c')} c_{ij} + \sum_{j\in\mathcal{S}_i(c')} c_{ij}}{|\mathcal{N}_i(c')|} \leq 1 - (1-\delta)\mathcal{Q}_i(c'), \tag{19}$$

consider a t-history h^t such that $c_i(h^t) = c_i$, and such that all ties in $c - \bar{c}$ are now strategically independent. For any Bayesian consistent system of beliefs, $\mu_i(h^t|h_i^t) = 1$ for a history $h^t \in h_i^t$ which is consistent with the path of play. The strategy (F, σ) prescribes that all players continue all ties, unless they observe a deviation from the path of play (since $c(F, \sigma) = c$). Moreover, for each tie $ij \in c - \bar{c}$, the strategy (F, σ) prescribes that i and j continue the tie indefinitely. Now consider the following deviation from player i: player i discontinues all ties in which she is recipient in period t, discontinues all ties in c' in which she is a sponsor in period $t+1$, and continues all of her ties in $c - c'$ indefinitely. Under local monitoring, the deviation on ties where she is a recipient can be transmitted to other players earliest in period $t+1$, and so other partners can respond earliest in $t+2$. Given the strategy profile (F_{-i}, σ_{-i}), the expected continuation value of this deviation is therefore

$$\begin{aligned}
\pi_1 \quad &\equiv \quad |\mathcal{R}_i(c)| + \sum_{j\in\mathcal{S}_i(c')} ((1-c_{ij}) + \delta) + \sum_{j\in\mathcal{N}_i(c-c')} \left(\frac{1-c_{ij}}{1-\delta}\right) \\
&= \quad |\mathcal{N}_i(c')| + \delta|\mathcal{S}_i(c')| - \sum_{j\in\mathcal{S}_i(c')} c_{ij} + \sum_{j\in\mathcal{N}_i(c-c')} \left(\frac{1-c_{ij}}{1-\delta}\right).
\end{aligned}$$

The expected continuation value for following the strategy profile (F, σ) at history h^t is

$$\pi_2 \equiv \sum_{j \in \mathcal{N}_i(c')} \left(\frac{1 - c_{ij}}{1 - \delta} \right) + \sum_{j \in \mathcal{N}_i(c - c')} \left(\frac{1 - c_{ij}}{1 - \delta} \right).$$

In order for (F, σ) to be a BFE, it must therefore be the case that $\pi_1 \leq \pi_2$, i.e.,

$$|\mathcal{N}_i(c')| + \delta |\mathcal{S}_i(c')| - \sum_{j \in \mathcal{S}_i(c')} c_{ij} \leq \sum_{j \in \mathcal{N}_i(c')} \left(\frac{1 - c_{ij}}{1 - \delta} \right),$$

which is equivalent to inequality (19).

To show that, for each $i \in \mathcal{N}(c)$,

$$\frac{\sum_{i \in \mathcal{N}_i(c')} c_{ij}}{|\mathcal{N}_i(c')|} \leq 1 - (1 - \delta) \frac{\sum_{j \in \mathcal{N}_i(c')} F^c_{ij}}{|\mathcal{N}_i(c')|} \tag{20}$$

suppose, by way of contradiction, that the inequality is not satisfied for some player $i \in \mathcal{N}(c)$. Since $c - c'$ contains only ties where the discounted value of indefinite cooperation is (weakly) lower than the formation cost, and since $F_{ij} \geq F^c_{ij}$, this implies

$$\frac{\sum_{i \in \mathcal{N}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} > 1 - (1 - \delta) \frac{\sum_{j \in \mathcal{N}_i(c)} F_{ij}}{|\mathcal{N}_i(c)|},$$

and so player i's expected continuation value for the initial history $h^\emptyset$ is strictly negative. But player i can guarantee a non-negative continuation payoff by not investing in any ties. As a result, if (F, σ) is a BFE, the inequality (20) must also be satisfied. □

To complete the proof, the following claim establishes that there exists a subnetwork $\tilde{c}$ such that $\bar{c} \subseteq \tilde{c} \subseteq c$ that satisfies triadic closure. From the previous claim, this subnetwork satisfies the RIC. Moreover, it contains all weak and asymmetric ties.

Claim 3. There exists a subnetwork $\tilde{c}$ such that $\bar{c} \subseteq \tilde{c} \subseteq c$, such that $\tilde{c}$ satisfies triadic closure.

Proof. If the network $\bar{c}$ satisfies triadic closure, the claim is trivial. So assume that $\bar{c}$ does not satisfy triadic closure. That means there is a player i with a tie $ij \in \bar{\mathcal{S}}_i(\bar{c})$ and another tie $ik \in \bar{c}$ such that $jk \notin \bar{c}$. There are two possibilities. Either (i) we can choose i, j and k such that $jk \notin c$, or (ii) $jk \in c$ for all possible choices of i, j and k. In case (ii), it is possible to augment $\bar{c}$ with strong ties from $c - \bar{c}$ to form a subnetwork $\tilde{c}$ that satisfies triadic closure, and so the claim is established. Without loss of generality, we therefore assume case (i), where we can choose i, j and k such that $jk \notin c$. To complete the proof, we show a contradiction.

Consider a history h^t where player i observed actions consistent with the path of play for (F, σ) in all periods prior to t, and in t player k discontinued the tie with i while all other players continued their ties with i. Since $ij \in \bar{c}$, by Claim 1, $\sigma_{ij}(h^t) = D$. However, there are consistent beliefs under which player i has a profitable deviation. Specifically, since h_i^t is not consistent with the path of play for (F, σ), there is a Bayesian consistent system of beliefs μ where $\mu_i(h^t | h_i^t) = 1$ for a history $h^t \in h_i^t$ such that $c(h^t) = c - \{ik\}$. Under these beliefs, if player i continues ij in period $t+1$, and other players follow σ_{-i}, player j observes the network neighborhood c_j in period t and $t+1$, and observes a deviation from the equilibrium path of play earliest in period $t+2$. Hence, σ_j prescribes for player j to discontinue tie ij at the earliest in period $t+3$. As a result, we can construct the following deviation strategy for player i: on all ties $il \neq ij$, player i continues as under σ_i, but player i discontinues ij in $t+2$ instead of $t+1$. Given the system of beliefs μ, the difference in the expected continuation payoff under this deviation and the expected continuation payoff under $\sigma_i(h_i^t)$ is $[(1 - c_{ij}) + \delta] - [1] = \delta - c_{ij}$. This payoff difference is strictly positive since i is a strict sponsor on the tie ij, contradicting the assumption that (F, σ) is a BFE. It follows that there exists a subnetwork $\tilde{c}$ such that $\bar{c} \subseteq \tilde{c} \subseteq c$, where $\tilde{c}$ satisfies triadic closure. □

A.4 A sequential equilibrium example

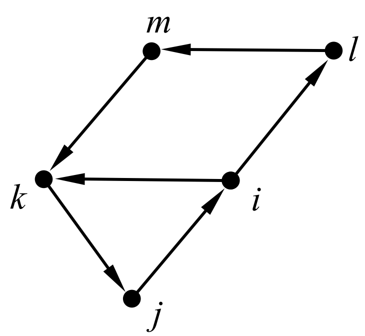
Various solution concepts impose structure on off-path beliefs in dynamic games with imperfect information. We view our results as providing bounds on the set of equilibrium network outcomes that can emerge for such solution concepts. To show how alternative solution concepts can refine the set of PBE network outcomes, we provide an example using Sequential Equilibrium (SE). A PBE (F, σ, μ) is a SE if there is a sequence of totally mixed strategies $(F^k, \sigma^k, \mu^k)_{k=1}^{\infty}$ that converges to (F, σ, μ), such that μ^k is Bayesian consistent with (F^k, σ^k) for all k. Since the set of equilibrium network outcomes under PBE and BFE are equivalent for $\rho \geq 2$, SE can only refine equilibrium network outcomes in an environment with local monitoring. For the example, we therefore focus on local monitoring in a binary formation model $(\underline{F}, \bar{F}, \underline{c}, \bar{c}, \delta)$ satisfying Condition (4).

Figure 8 illustrates two network structures that cannot emerge in a BFE (because the networks do not satisfy the necessary triadic closure property), but can emerge in a PBE for a range of parameters. The following conditions are sufficient to satisfy the LIC:

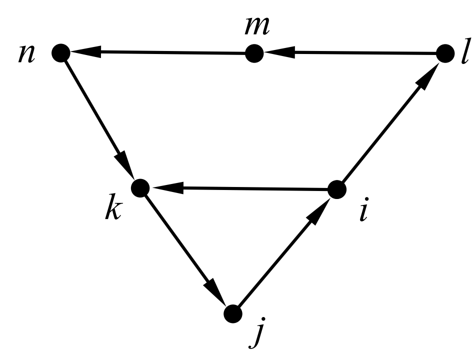
$$\frac{(1-\bar{c}) + 2(1-\underline{c})}{1-\delta} \geq 2\bar{F} + \underline{F} \tag{21}$$

$$\frac{2(1-\bar{c}) + (1-\underline{c})}{1-\delta} \geq 1 + (1-\bar{c})\left(\frac{1-\delta^3}{1-\delta}\right) + \delta^3 + (1-\underline{c})\left(\frac{1-\delta^4}{1-\delta}\right) + \delta^4, \tag{22}$$

$$\frac{(1-\bar{c}) + (1-\underline{c})}{1-\delta} \geq 1 + (1-\underline{c})\left(\frac{1-\delta^3}{1-\delta}\right) + \delta^3, \tag{23}$$



(a) SE network



(b) Not SE network

Figure 8: PBE network outcomes

Condition (21) ensures that player i does not have an incentive to deviate in the formation phase; given Condition (4) other players then also have no incentive to deviate. Conditions (22) and (23) ensure that players k and n do not have an incentive to deviate in the continuation phase; given Condition (4), other players then also have no incentive to deviate. All three conditions can be satisfied by a range of parameters, e.g., $(\underline{F}, \bar{F}, \underline{c}, \bar{c}, \delta) = (0.1, 4.1, 0.2, 0.9, 0.8)$, and Proposition 2 then implies that the networks are PBE outcomes.

Under Conditions (21)–(23), the network in Figure 8(a) is also a SE outcome and therefore provides an example of a SE network outcome that is not a BFE.[13] However, the network in Figure 8(b) is not a SE outcome, and therefore provides an example of a PBE network outcome that is not a SE. In particular, for the network in Figure 8(b), even when the LIC is satisfied, there is a "hold-up" problem that gives player l an opportunity for a profitable deviation. To illustrate, let (F, σ) be any strategy-profile that induces network outcome c in Figure 8(b) on the path of play. Player l forms ties li and lm, but then deviates by defecting on li in period 0, continuing cooperation on lm. Under Condition (4), this deviation is profitable for player l unless there is a credible threat that player m will discontinue tie lm in a future period. Since m does not observe l's deviation, defection must spread through the network, which requires i to discontinue either tie ik or tie ij (since i is the only player to observe l's deviation).

Relative to PBE, the key difference in a SE is that, when i observes l's deviation, i must believe with probability 1 that only l has deviated.[14] Assume i believes that l deviated on both

[13]Players use formation strategy F^c and continuation strategy σ, where they cooperate on all ties unless they observe a deviation, and discontinue all ties following a deviation. Off-path beliefs μ are as follows: when player r observes a deviation by player s, r believes (with probability 1) that s deviated on all their ties, and all other players continued to follow the equilibrium strategy. These beliefs ensure that (F^c, σ) is a best-response at every it-history. Moreover, if players follow the strategy-profile (F^c, σ) but deviate with probability $\varepsilon_k > 0$ on all their ties in each period, the Bayesian consistent beliefs for the sequence of totally mixed strategies converges to μ.

[14]Since randomizations in a mixed strategy are independent, for the Bayesian consistent beliefs of any sequence

il and lm (the argument is similar when i believes that l only deviated on il). Player i therefore believes that only i, l and m have observed deviations. On the corresponding network, the continuation strategy must be sequentially rational for players i and m, given that they both believe that only l deviated. Player m has an incentive to cooperate on tie mn until m believes that n will defect in the next period. Given that the defection must originate from i, the earliest m should defect on mn is period $\tau + 1$, where τ is the period when i defects. Player i also has an incentive to continue cooperating on ties ij and ik as long as possible. In particular, if m cooperates on tie mn until period τ, and then defects in $\tau + 1$, then i should continue to cooperate on ij at least until period $\tau + 1$ and cooperate on ik until $\tau + 2$. As a result, there is a "hold-up" problem: neither i nor m have an incentive to defect on their ties until the period after the other has defected. However, if player i does not defect on a remaining tie, then l cannot receive the third-party punishment required to deter l from the initial deviation.[15]

---

of totally mixed strategy profiles (F^k, σ^k) converging to (F, σ), it is an order of magnitude more likely that only l deviated in $t = 0$ than that other players also deviated.

[15]Beyond the bounds provided by our BFE and PBE results, we do not have further general results for SE network outcomes. Comparing the networks in Figures 8(a) and 8(b) illustrates that SE outcomes can depend on fine details of the network structure. In general, the strategic hold-up problems illustrated by Figure 8(b) interact with incentives in the continuation game, and providing general results when network incentive constraints interact with complex inference problems is a challenging problem, which we leave for future work.