# Cooperative Networks with Robust Private Monitoring*

Maximilian Mihm† Russell Toth‡

August 1, 2018

**Abstract**

Social networks support cooperative behavior in a variety of social and economic settings. We study the cooperative networks that can be formed when payoff and information asymmetries imply that cooperation relies on unverifiable third-party punishments. We focus on equilibrium network outcomes where individual strategies are robust to beliefs about unobservable behavior. Network structures with robust private monitoring 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 often rely on informal (non-contractual) 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 (Bernheim and Whinston [1990]). Decentralized community

*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.

enforcement mechanisms typically leverage contagion strategies, triggering a spread of non-cooperative behavior following deviations (Kandori [1992], Ellison [1994]). A recent strand of research has sought 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 network structures that can support cooperative behavior under local 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 of our game, 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 higher one-off benefit. 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. A player's investment in a tie during the formation phase has an inverse relationship to their subsequent cost of cooperation, reflected in a "formation technology" that can be different in each potential bilateral relationship. This generates a trade-off between the effort cost in forming ties, and the effort cost required to cooperate. We focus on the range of parameters where this trade-off implies 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 across heterogeneous ties.

Under perfect monitoring, virtually any network structure can emerge under the right configuration of parameters. In particular, full cooperation can be supported by equilibria that dictate defection on every active relationship for any observed deviation. 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 where players can only observe the history of play in their own network neighborhood.

In this local monitoring environment, we first characterize the network outcomes that can emerge in a Perfect Bayesian Equilibrium (PBE). Under local monitoring, the network pools incentive constraints and also plays an important 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 depends on the distance between nodes, and tighter networks (with shorter paths between nodes) are generally needed to make cooperation incentive compatible. Our characterization of PBE network outcomes extends previous literature (e.g., Raub and Weesie [1990], Lippert and Spagnolo [2011]) to a framework where incentives in the formation of ties is in tension with cooperative behavior on the resulting network.

A concern in the literature on network interactions with local monitoring is that, while a PBE restricts beliefs on the equilibrium path of play, it imposes very few restrictions on beliefs off the equilibrium path. Under local monitoring, a single deviation from the strategy profile implies that beliefs about behavior outside a player's own network neighborhood are essentially unrestricted. As a result, a PBE can be very sensitive to the beliefs following a deviation, and this raises questions about the fragility of the equilibrium predictions. Motivated by these concerns, we take a positive approach to equilibrium refinement by adapting Belief Free Equilibrium (Piccione [2002], Ely and Välimäki [2002], Ely et al. [2005], Hörner and Lovo [2009], Hörner et al. [2011]) to the network setting. In a belief-free equilibrium (BFE), a strategy must be sequentially rational for *all* beliefs that are Bayesian consistent with the strategy profile, and so individual strategies are robust to beliefs about unobservable behavior. It is natural that this equilibrium refinement generates sharper predictions on network outcomes. Indeed, we show that network ties must satisfy a "triadic closure" property.[1] At the network level this 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.

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

[1] Interest in clustering and specifically the concept of triadic closure in sociology dates to Simmel [1908], and was popularized by Granovetter [1973]. For a more recent discussion see also Watts [1999].

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 of their ties.[2] Raub and Weesie [1990] provide a model that captures the intuition that tightly-connected networks shorten the travel time for contagion to spread through a network, 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 the role of local monitoring in a setting with 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 network ties in unified framework.[3] Our approach thereby emphasizes a trade-off between investments in the formation of relationships and the returns to cooperation. Second, we propose a positive criterion for refining PBE in terms of BFE. The prior literature has often used a normative criteria (i.e., efficiency) to analyze a subset of the

[2] 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].

[3] 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.

PBE (e.g., Ali and Miller [2013], Balmaceda and Escobar [2017]). Our main result for BFE similarly highlights the importance of clustering and density for cooperative outcomes, but results from a very different mechanism that requires punishment strategies that are robust to 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 based on strategic play under limited information 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. We analyze the game by restricting the environment in stages. First, we consider the efficient outcomes that would emerge if players could perfectly commit to cooperate. We then assume that players are unable to commit to cooperate, but are able to perfectly monitor the behavior of all other players. Finally, our main analysis considers network outcomes that can emerge (in PBE and BFE) under local monitoring, where players cannot commit to cooperate and can monitor behavior only in their own network neighborhood. 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 an 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 $i$).[4] When $c_{ij} < 1$ we say that $i$ and $j$ are partners in network $c$ and write $ij \in c$.

[4]It is important for our formulation of the continuation game that both players either have the option to

## 2.1 Formation phase

In the formation phase, players decide how much to invest in ties. Investments are made unilaterally and simultaneously in period $t = 0$. We impose a bilateral formation constraint by assuming that the formation of a tie requires that each partner incurs some minimal investment cost $\underline{F} > 0$.[5] Player $i$ can either invest 0 with $j$ (so no tie is formed), or choose a strictly positive investment from $\left[\underline{F}, \bar{F}\right]$, where $\underline{F} < 1 < \overline{F}$. We denote player $i$'s investment in a tie with player $j$ by $F_{ij} \in \mathcal{F} \equiv \{0\} \cup \left[\underline{F}, \overline{F}\right]$, and require that $F_{ii} = 0$.

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 the function $\gamma_{ij} : \left[\underline{F}, \overline{F}\right] \to (0, 1)$ is monotone non-increasing and lower-semicontinuous for all $i, j \in \mathcal{I}$. The formation technology $\Gamma$ has the following interpretation. Suppose players $i$ and $j$ invest $F_{ij} > 0$ and $F_{ji} > 0$ in forming a tie. Then a tie is formed between players $i$ and $j$, the strength of the tie from $i$'s point of view is $\gamma_{ij}(F_{ij})$, and the strength of the tie from $j$'s point of view is $\gamma_{ji}(F_{ji})$. For a given $\Gamma$, a profile of investments $F = (F_{ij})_{i,j\in\mathcal{I}}$ therefore induces a network $c(F) \equiv (c_{ij}(F))_{i,j\in\mathcal{I}}$, defined by

$$c_{ij}(F) \equiv \begin{cases} \gamma_{ij}(F_{ij}) & \text{if } F_{ij}F_{ji} > 0 \\ 1 & \text{otherwise} \end{cases}.$$

The following examples, discussed in more detail in Section 4, illustrate the flexibility of the formation model:

**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 $\Gamma$ where there is a function $\gamma : \left[\underline{F}, \overline{F}\right] \to (0, 1)$ such that $\gamma_{ij} = \gamma$ for all $i, j \in \mathcal{I}$. (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 the same group, or members of different groups. This model is the special case of $\Gamma$ 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 with whom a tie is

cooperate or both players do not have the option to cooperate.

[5]For our results, it is not necessary that $\underline{F}$ is strictly positive. If we allowed $\underline{F} = 0$, what matters is that there is still a qualitative difference between not forming a tie and forming a tie without incurring a cost. In the former case, both players cannot cooperate in the future, while in the latter they will face a cooperation decision already in the same period. It simplifies exposition to assume that $\underline{F} > 0$ so that we do not need to distinguish between these two cases.

formed. This model is the special case of $\Gamma$ 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 $j \in \mathcal{I}_k$. □

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

$$F_{ij}^c \equiv \begin{cases} \min \left\{ f \in \left[\underline{F}, \bar{F}\right] \; : \; \gamma_{ij}(f) = c_{ij} \right\} & \text{if } c_{ij} > 0 \\ 0 & \text{if } c_{ij} = 0 \end{cases}. \tag{1}$$

The formation profile $F^c$ is well-defined because $\gamma_{ij}$ 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 the 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 corresponding formation strategy that minimizes formation costs.

## 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 in the network. We denote by $A_{ij}^t \in \{C, D\}$ the action of player $i$ in the tie with player $j$ in period $t$, where $C$ denotes that $i$ *cooperates* with $j$ and $D$ denotes that $i$ *doesn't cooperate* with $j$. Cooperation decision are unilateral, simultaneous, and independent across ties.

When player $i$ cooperates with player $j$, player $j$ receives a *cooperation benefit* normalized to 1 and $i$ incurs the *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 with $j$ 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-cooperative behavior is a strictly dominant strategy but mutual cooperation is strictly Pareto improving.

Given a network $c^t$ in period $t \geq 0$, when players $i$ and $j$ both cooperate their tie continues to period $t+1$; otherwise the tie is discontinued. The actions $C$ and $D$ could 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 significantly because it facilitates 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 could 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 that ties are discontinued permanently following a defection.[6] A decision by each player about whether to cooperate with each partner then determines a new network $c^{t+1}$ in period $t+1$, which is (weakly) a subset of the network in period $t$, and where players again decide whether to cooperate on each of their remaining ties.

## 2.3 Bilateral incentive constraints

We are interested in model parameters where bilateral ties can be formed only when they are strategically embedded in a network of other ties. We therefore impose an assumption on parameters to ensure that isolated bilateral ties are not incentive compatible.

To illustrate, 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 player $j$ will cooperate. Moreover, the discounted value of cooperation must exceed the one-off benefit of discontinuing the tie. Hence, the following *bilateral cooperation constraint* must be satisfied:

$$\left(\frac{1-c_{ij}}{1-\delta}\right) \geq 1 \;\Leftrightarrow\; c_{ij} \leq \delta. \tag{2}$$

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 player $i$ incurs, and so the following *bilateral formation constraint* must be satisfied:

$$\left(\frac{1-c_{ij}}{1-\delta}\right) \geq F^c_{ij} \;\Leftrightarrow\; c_{ij} \leq 1-(1-\delta)F^c_{ij}. \tag{3}$$

Combining the cooperation and formation constraints, a strategically isolated tie $c = (c_{ij}, c_{ji})$ must satisfy the following *bilateral incentive constraint (BIC)*:

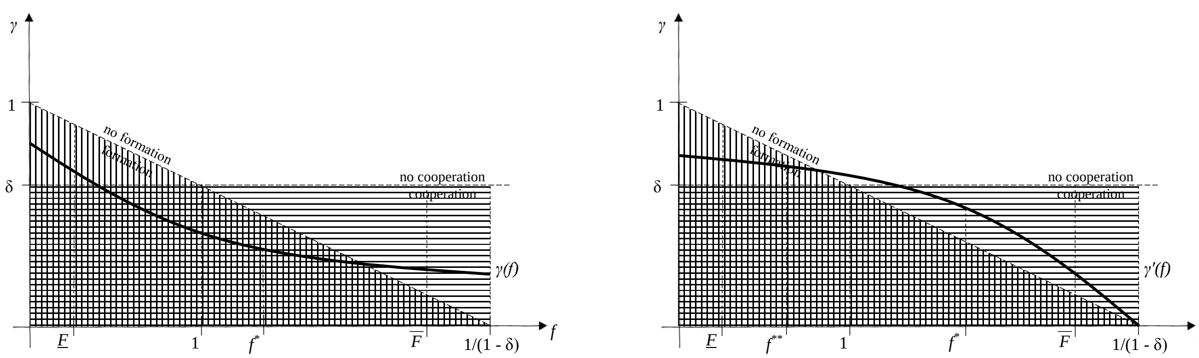
$$c_{ij} \leq \min\left\{\delta, 1-(1-\delta)F^c_{ij}\right\}.$$

[6] In general, private monitoring games are challenging to analyze because there many possible deviations, which generate complex inference problems for the agents. For example, in the literature on community enforcement (e.g., Kandori [1992]; Ellison [1994]), Folk-theorem type results often require that players who observe a deviation punish for a short period, and then forgive the deviation and resume cooperation. However, in the typical community enforcement set-up, players receive information about their opponents' past actions. In our context, there is the additional challenge that there are significant information asymmetries, because players never learn the past actions of their opponents in ties they are not party to. In principle, players could use patterns of partial cooperation to communicate information about the structure of the network, to overcome information asymmetries. Such strategies would be very complex, and we have not been able to make progress on how and when they could arise in equilibrium if we allowed players to forgive defections. We therefore focus in the following on the simpler environment where defections lead to a permanent discontinuation of a tie.

Our main assumption on parameters implies that, for any pair of players $i$ and $j$, the BIC cannot be satisfied.

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

Assumption 1 is a joint restriction on the model parameters $(\delta, \underline{F}, \bar{F}, \Gamma)$, illustrated in figure 1. The assumption implies that, for any choice of formation investment $f \in [\underline{F}, \bar{F}]$, either the bilateral continuation constraint or the bilateral formation constraint (or both) is not satisfied. As a result, a strategically isolated bilateral tie is not incentive compatible for any pair of players, and any cooperative behavior that emerges must be embedded in a network.




((a)) BIC satisfied

((b)) BIC never satisfied

Figure 1: Bilateral incentive constraints

*Figure 1(a) illustrates a formation technology $\gamma$ that violates Assumption 1 because at investment $f^*$ the BIC. is satisfied. Figure 1(b) illustrates a formation technology $\gamma'$ that is consistent with Assumption 1. At investment $f^*$, the bilateral cooperation constraint is satisfied, but the bilateral formation constraint is violated; at investment $f^{**}$, the bilateral formation constraint is satisfied, but the bilateral cooperation constraint is violated; there is no investment $f \in [\underline{F}, \bar{F}]$ such that the bilateral cooperation and formation constraints are satisfied simultaneously.*

While bilateral incentive constraints cannot be satisfied under Assumption 1, it is possible for a network to "pool" incentive constraints across ties. The following example provides an illustration.

**Example 2** (Binary formation model)**.** Consider the *binary formation* model, which is strategically equivalent to a special case of the homogenous formation technology in Example 1, where players face a binary formation choice $f \in \{\underline{F}, \bar{F}\}$.[7] This binary model satisfies Assumption 1 if and only if

$$\underline{c} > 1-(1-\delta)\bar{F} \text{ and } \bar{c} > \delta.$$

[7]Formally, this is a version of the homogenous formation technology model from Example 1, where $\gamma_{ij} = \gamma$ for all players $i$ and $j$. In addition, suppose that, for some $\underline{c} < \bar{c}$, $\gamma(\bar{F}) = \underline{c}$ and $\gamma(f) = \bar{c}$ for $f \in [\underline{F}, \bar{F})$, and the function $\gamma$ is monotone non-increasing and lower semi-continuous.

In order for a network of ties to be incentive compatible it is necessary that $\underline{c} \leq \delta$ and $\bar{c} \leq 1-(1-\delta)\underline{F}$. Hence, the following is a necessary condition for a network to pool constraints:

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

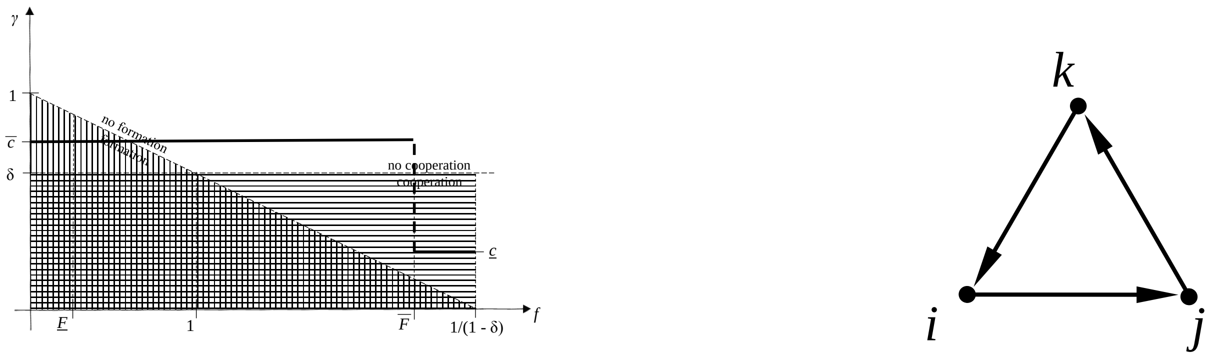
When condition (4) is satisfied, it is possible for some ties to ensure that continuation is incentive compatible, while other ties ensure that formation is incentive compatible. To illustrate, suppose there are three players $\mathcal{I}=\{i,j,k\}$, $\delta=\frac{8}{10}$, $\bar{F}=\frac{16}{10}$, $\underline{F}=\frac{4}{10}$, $\gamma\left(\bar{F}\right)=\frac{7}{10}$, and $\gamma(\underline{F})=\frac{9}{10}$. Then condition (4) is satisfied with strict inequalities, and so Assumption 1 holds. Now suppose there is perfect monitoring and consider the following strategy profile: 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 their ties as long as all other players cooperate, and otherwise stop cooperating on all remaining ties. On the path of play, this strategy profile induces the network $c$ in figure 2, on which players cooperate indefinitely. Moreover, the strategy profile is a subgame perfect Nash 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 both 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, while strategically isolated bilateral ties are not incentive compatible, there is an equilibrium in which a network of ties is formed and cooperation continues indefinitely.[8] □

## 2.4 Network game

We are interested in the networks that can be formed in an equilibrium, and on which players continue to cooperate indefinitely. We view such networks as the long-term stable and observable outcomes of the game. To study such equilibrium network outcomes we can focus on pure strategies because, given our assumption that a defection by one of the partners on a tie leads to a permanent discontinuation of the tie, any mixing by players makes the network unstable. In other words, a network is formed and continued indefinitely if and only if, on the equilibrium path of play, all players cooperate with probability 1 on all the ties that are formed. Focusing on pure strategies also simplifies the formal exposition of the strategic environment.

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 all actions by

[8]It is also easily verified that if, instead, we assume $\bar{F}=\frac{16}{10}-\varepsilon$, $\underline{F}=\frac{4}{10}-\varepsilon$ for sufficiently small $\varepsilon>0$ (holding other parameters fixed), Assumption 1 is still satisfied. Moreover, the strategy profile whereby the cooperative network $c$ is formed remains a subgame perfect equilibrium outcome, and now yields strictly positive net payoffs for every player.




((a)) Binary formation technology ((b)) Triangle network

Figure 2: Example 2

*Figure 2(a) illustrates the binary formation technology in Example 2, which is consistent with Assumption 1. With this technology, strategically isolated ties are not incentive compatible. Figure 2(b) illustrates a triangle network consisting of three asymmetric ties, which can emerge in a subgame perfect equilibrium. Dots represent nodes (or players), arrows represent ties, and the direction indicates where a player sponsors a tie.*

every player in periods $\tau < t$ (the formation phase and the continuation phase prior to period $t$). An $\infty$-*history* defines a complete path of play for the game. Denote by $\mathcal{H}^t$ the set of $t$-histories. 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 $t$.

A *monitoring structure* describes the actions a player is able to observe. We consider two cases: (i) *perfect* monitoring, and (ii) *local* monitoring. Under perfect monitoring, a player can observe the full history of play in every period. Under local monitoring, player $i$ can only observe player $j$'s direct interactions with player $i$, not $j$'s interaction with player $k \neq i$. That means, in the formation phase, player $i$ can observe $F_{ji}$ but cannot observe $F_{jk}$ for $k \neq i$. Likewise, in period $t \geq 1$ of the continuation phase, player $i$ can observe $A_{ji}^{\tau}$ for $\tau < t$, but cannot observe $A_{jk}^{\tau}$ for $k \neq i$. As a result, player $i$ is unable to distinguish between two $t$-histories $h^t$ and $\tilde{h}^t$ that differ only in terms of actions that do not directly involve player $i$. 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 $\tilde{h}^t$. We call an equivalence class of $\sim_i$ an *it-history*, and denote the set of *it*-histories by $\mathcal{H}_i^t$ with typical element $h_i^t$.

**Definition 1.** The game has *perfect monitoring* if, for all $i \in \mathcal{I}$ and $t \geq 0$, $h^t \sim_i \tilde{h}^t$ if only if $h^t = \tilde{h}^t$. The game has *local monitoring* if, for all $i \in \mathcal{I}$ and $t \geq 0$, $h^t \sim_i \tilde{h}^t$ if only if (i) $F_{ij} = \tilde{F}_{ij}$ and $F_{ji} = \tilde{F}_{ji}$ for all $j \neq i$, and (ii) $A_{ij}^{\tau} = \tilde{A}_{ij}^{\tau}$ and $A_{ji}^{\tau} = \tilde{A}_{ji}^{\tau}$ for all $j \neq i$ and $\tau = 0, ..., t-1$.

In the continuation phase, a strategy of player $i$ is 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) $ij \notin c(h^t)$ implies $\sigma_i(j, h^t) = D$, and (ii) $h^t \sim_i \tilde{h}^t$ implies $\sigma_i(., h^t) = \sigma_i\left(., \tilde{h}^t\right)$. 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.

In the formation phase, the strategy profile $(F, \sigma)$ induces a unique period 0-history $h^0$. For any $t$-history $h^t$, the strategy $(F, \sigma)$ induces a unique $(t+1)$-history, denoted $h^{t+1}(F, \sigma | h^t)$. Recursively, the strategy $(F, \sigma)$ therefore induces a unique $(t+\tau)$-history $h^{t+\tau}(F, \sigma | h^t)$. As a result, the strategy $(F, \sigma)$ induces a *continuation path* for any $t$-history $h^t$, denoted $h^\infty(F, \sigma | h^t)$ (let $h^\infty\left(F, \sigma | h^\emptyset\right)$ denote the continuation path induced by $(F, \sigma)$ 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 $(F, \sigma)$ given history $h^\tau$. We are interested in the networks that can be formed and continued for an equilibrium strategy profile $(F, \sigma)$. With some abuse of notation, we write $c(F, \sigma) = c$ 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 $\mu$ is *Bayesian consistent* with strategy profile $(F, \sigma)$ 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 $(F_{-i}, \sigma_{-i})$ unless she observes an $it$-history that is inconsistent with $(F_{-i}, \sigma_{-i})$.[9] Conditions (i) and (ii) ensure that, whenever possible, beliefs at an information set are Bayesian updated from the strategy profile $(F, \sigma)$.

For player $i \in \mathcal{I}$, $\pi_i$ 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 $(F, \sigma)$ following history $h^t$. Expected payoffs and best responses are defined with respect to a system of beliefs $\mu$. For player $i \in \mathcal{I}$, the *expected continuation payoff* under strategy profile $(F, \sigma)$ 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 $(F_i, \sigma_i)$ is a $(\mu, h_i^t)$-*best response* to $(F_{-i}, \sigma_{-i})$ if

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

[9] Formally, let $(\tilde{F}_i F_{-i}, \tilde{\sigma}_i \sigma_{-i})$ be the the strategy profile where player $i$ follows $(\tilde{F}_i, \sigma_i)$ and players other than $i$ follow $(F_{-i}, \sigma_{-i})$. 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 $\mu_i\left(h^t\left(\tilde{F}_i F_{-i}, \tilde{\sigma}_i \sigma_{-i} | h^\emptyset\right) | h_i^t\right) = 1$.

The strategy profile $(F_i, \sigma_i)$ is a *$\mu$-best response* to $(F_{-i}, \sigma_{-i})$ if, for every *it*-history $h_i^t$, $(F_i, \sigma_i)$ is a $(\mu, h_i^t)$-best response to $(F_{-i}, \sigma_{-i})$. A strategy $(F_i, \sigma_i)$ is therefore a best response if it maximizes player $i$'s expected continuation payoff at every information set, given that other players follow $(F_{-i}, \sigma_{-i})$.

## 2.5 Equilibrium

We are interested in equilibrium *network outcomes*: the set of networks that can be formed and continued indefinitely. Under perfect monitoring our equilibrium concept is *sub-game perfect Nash equilibrium* (SPNE). Under local monitoring there are no proper sub-games, and a belief-based refinement of Nash equilibrium is required to exclude non-credible punishments. We consider two solution concepts. In a *perfect Bayesian equilibrium (PBE)*, player $i$'s strategy is a best response for *some* system of beliefs that is Bayesian consistent with the equilibrium strategy profile. PBE ensures that play is sequentially rational and punishments are credible given a specific system of beliefs. However, after a single deviation from the equilibrium strategy profile, beliefs about behavior outside a player's own network neighborhood are essentially unrestricted in a PBE. As such, equilibrium predictions are very sensitive to the beliefs players hold about cooperative behavior following a deviation. We therefore also study network outcomes in a *belief-free equilibrium (BFE)*, which requires sequential rationality for *all* beliefs that are Bayesian consistent with the equilibrium strategy profile.

**Definition 2** (Equilibrium)**.** A strategy profile $(F, \sigma)$ is a PBE if, for all players $i$, $(F_i, \sigma_i)$ is a $\mu$-best response for some system of beliefs $\mu$ that is Bayesian consistent with $(F, \sigma)$. A strategy profile $(F, \sigma)$ is a BFE if, for all players $i$, $(F_i, \sigma_i)$ is a $\mu$-best response for every system of beliefs $\mu$ that is Bayesian consistent with $(F, \sigma)$.

The BFE solution concept was introduced for games of imperfect monitoring in Piccione [2002], Ely and Välimäki [2002], and Ely et al. [2005], and extended to games with incomplete information in Hörner and Lovo [2009] and Hörner et al. [2011]. In the typical setup in this prior literature, a player's payoff in a given period depends stochastically on the action profile chosen by them and others, and an equilibrium is belief-free if, at every information set, each player's response is optimal for all Bayesian consistent beliefs, both on and off the equilibrium path of play. We adapt BFE to a network setting with a more structured set of relationships between players, and a corresponding information structure whereby players receive no information at all about the actions chosen on ties on which they are not one of the partners. Under our adaptation, BFE is a refinement of PBE that ensures that cooperative behavior is (i) not enforced by non-credible threats (as in a PBE), and (ii) that punishment threats are robust to beliefs *off the equilibrium path* (unlike in a PBE). Enforcing belief-freeness on the equilibrium path would guarantee that only the trivial equilibrium exists under which no players makes any investments during the formation phase. However, this equilibrium is still trivially a BFE, and so a BFE always exists.

## 2.6 Network notation

As our results make equilibrium predictions about network outcomes, we require some additional notation to describe network structures.

**Subnetworks:** A network $\tilde{c}$ is a *subnetwork* of network $c$ (denoted $\tilde{c} \subseteq c$) if (i) $ij \in \tilde{c}$ implies $ij \in c$, and (ii) $\tilde{c}_{ij} = c_{ij}$ whenever $ij \in \tilde{c}$. When $\tilde{c}$ is a subnetwork of $c$, we use $c - \tilde{c}$ to denote the complementary subnetwork of $c$ defined by the condition that $ij \in c - \tilde{c}$ if and only if $ij \notin \tilde{c}$ and $ij \in c$. For a subnetwork $\tilde{c} \subseteq c$, $\mathcal{N}(\tilde{c})$ 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 between nodes $i$ and $j$ in network $c$, denoted $d_c(i,j)$, is the length of the shortest path between these nodes; in particular, $d_c(i,j) = 1$ if $ij \in c$, and $d_c(i,j) = \infty$ if there is no path from $i$ to $j$ in network $c$. For a subnetwork $\tilde{c} \subseteq c$, let $d_c(i,j|\neg\tilde{c}) \equiv d_{c-\tilde{c}}(i,j)$ be the distance between nodes $i$ and $j$ when paths are restricted not to pass through ties in $\tilde{c}$.

**Categories of ties:** For network $c$, we say that $i$ is a *sponsor* of the tie with $j$ 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 partner is a strict recipient (i.e., $\min\{c_{ij}, c_{ji}\} < \delta < \max\{c_{ij}, c_{ji}\}$).

**Neighborhood:** We denote player $i$'s *neighborhood* in network $c$ by $c_i$ , i.e., $c_i$ is the set of ties in $c$ where $i$ is one of the partners. Let $\mathcal{N}_i(c)$ be the set of players who are a partner of $i$ in $c$; $\mathcal{S}_i(c)$ be the set of players $j \in \mathcal{N}_i(c)$ such that $i$ sponsors a tie with player $j$; $\bar{\mathcal{S}}_i(c)$ the players with whom $i$ strictly sponsors a tie; and $\mathcal{R}_i(c)$ the players $j \in \mathcal{N}_i(c)$ such that player $i$ receives a tie from player $j$. The *ratio* of player $i$ in network $c$, $\mathcal{Q}_i(c) \equiv \frac{|\mathcal{S}_i(c)|}{|\mathcal{N}_i(c)|}$, measures the proportion of ties in her neighborhood on which player $i$ is a sponsor.

# 3 Network outcomes

## 3.1 Efficiency

As a benchmark, we first consider an environment where players can enforce contracts that commit them to cooperative behavior, and are able to make side-payments. Players can then achieve efficient outcomes by treating every bilateral interaction independently. To describe the network outcomes that can emerge, fix two players $i$ and $j$. Because $\gamma_{ij}$ 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 $j$. Let $\pi^*_{ij} = \frac{1-\gamma_{ij}(F^*_{ij})}{1-\delta} - F^*_{ij}$ be the maximum net-payoff player $i$ can achieve from a tie with player $j$. An efficient network outcome is one where players $i$ 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.

While it is efficient for players to treat bilateral interactions independently, the efficient network outcome is not incentive compatible when players cannot commit to cooperative behavior. When Assumption 1 holds, $\pi^*_{ij} \geq 0$ if and only if $\gamma_{ij}(F^*_{ij}) > \delta$. Hence, when a tie $ij$ is viewed independently of other ties, any time the tie provides a positive net-payoff for player $i$ there is also a strict incentive for player $i$ to discontinue the tie after it is formed (and likewise for player $j$). The strategic tension between formation and continuation incentives therefore implies that efficient outcomes cannot be achieved. The following example provides a simple illustration.

**Example 3.** [Efficient network with binary formation] In the binary formation model (Example 2), the efficient network outcome is a complete network of weak ties: players invest $\underline{F}$ in ties with every other player and cooperate in every period. However, since $\gamma(\underline{F}) = \bar{c} > \delta$, a player who has only weak ties has a strict incentive to discontinue all of their ties in period 0 of the continuation phase. As a result, there is no way to achieve the efficient outcome in an equilibrium. At best, players can hope to form heterogeneous ties (as in Example 2) that allow them to achieve cooperative outcomes by pooling incentive constraints via the network. □

## 3.2 Perfect monitoring

Without commitment, players must enforce cooperative behavior with the threat of future punishment. When Assumption 1 holds, bilateral punishment is not sufficient to enforce cooperative behavior on any tie that provides a strictly positive net-payoff. As a result, players cannot treat bilateral interactions independently, and ties can only be formed when there are credible threats of third-party punishment. A network can overcome bilateral incentive constraints by pooling incentives for formation and cooperation across ties.

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 a feasible deviation. Hence, the following *public cooperation constraint (PubCC)* 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 the ties in their own neighborhood $c_i$, because not forming any ties is also a feasible deviation. Hence, the following *public formation constraint (PubFC)* must be satisfied:

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

Combining the cooperation and formation constraints leads to the following incentive constraint for a game with perfect monitoring.

**Definition 3.** Network $c$ satisfies a *public incentive constraint (PubIC)* if, for all players $i \in \mathcal{N}(c)$,

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

The PubIC is clearly necessary for a cooperative network $c$ to emerge in a SPNE. If a network $c$ does not satisfy the PubIC, at least one player has an incentive to deviate from a strategy that prescribes network outcome $c$. The following proposition shows that the PubIC is also sufficient: when a network $c$ satisfies the PubIC, there is SPNE under which the network $c$ is formed during the formation phase, and players cooperate on the network indefinitely.

**Proposition 1.** *There exists a SPNE $(F, \sigma)$ such that $c(F, \sigma) = c$ if and only if network $c$ satisfies the PubIC.*

Under Assumption 1, player $i$ must form some ties where $1 - (1-\delta)F^c_{ij} < c_{ij} < \delta$ and other ties where $\delta < c_{ij} < 1 - (1-\delta)F^c_{ij}$. On ties where $c_{ij} < \delta$, there is slack in the continuation constraint, while on ties where $c_{ij} < 1 - (1-\delta)F^c_{ij}$, there is slack in the formation constraint. Network $c$ must ensures that slack in continuation 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 $c_{ij} < 1 - (1-\delta)F^c_{ij}$ is sufficient to overcome the negative net-payoff on ties where $1 - (1-\delta)F^c_{ij} < c_{ij}$. As a result, network $c$ satisfies the PubIC when it is relatively cheap for player $i$ to form ties with costly cooperation ($c_{ij} > \delta$), and relatively cheap to cooperate on ties with costly formation ($c_{ij} > 1 - (1-\delta)F^c_{ij}$).

**Example 4.** [Binary formation with perfect monitoring] In the binary formation model (Example 2), a network $c$ satisfies the PubIC 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 Assumption 1 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


((a)) Asymmetric ties ((b)) Alternating ties

Figure 3: Cycle networks

*Figures 3(a) and (b) illustrate cycle networks of length $n = 6$. In Figure 3(a), the cycle consists of asymmetric ties, arranged such that each player sponsors and receives one tie. In Figure 3(b), the cycle consists of weak and strong ties, which alternate such that each player sponsors and receives one tie.*

PubIC for some parameters consistent with condition (4). To see this, recall that $\mathcal{Q}_i(c)$ is the proportion of ties where player $i$ is a sponsor. For condition (8) to be satisfied, it is necessary for $0 < \mathcal{Q}_i < 1$ 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 $\delta$ and $\bar{C}$ is sufficiently close to $\underline{C}$, which holds for a range of parameters consistent with Assumption 1. 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 large variety of network outcomes can arise in SPNE with perfect monitoring. □

## 3.3 Local monitoring

### 3.3.1 Perfect Bayesian equilibrium

With local monitoring, in addition to pooling incentive constraints, the network plays an important role in transmitting information. Since third parties cannot observe deviations directly, 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 in the class of cycle networks.

**Example 5.** [Cycle networks with binary formation] In the binary formation model (Example 2), we can compare when a cycle network of length $n \geq 3$ (figure 3) is incentive compatible under perfect versus local 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 public 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 during the continuation phase. Players cooperate unless they observe a single defection, and then defect on all remaining ties. In particular, the public cooperation constraint does not depend on the size of the cycle. With local monitoring, when player $i$ defects on tie $ij$, player $k$ does not observe the defection. Instead, third-party punishment by player $k$ is 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 defer defection 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 perfect monitoring. However, the right-hand side is smaller, and depends on the length of the cycle. In particular, the right-hand side equals $\delta$ 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 perfect monitoring, but a cycle of sufficient length $n$ is not incentive compatible under local monitoring. $\square$

To characterize the incentive constraints that must be satisfied for a general network $c$, 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 in any period where their current network neighborhood is equal to $c_j$, 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 $\alpha^i_j$ represents the first period in which player $i$ plans to cease cooperation with player $j$, and $\alpha^i_j = -1$ if player $i$ does not plan to form a tie with player $j$ in the formation phase.

For network $c$, a defection plan for player $i$ should satisfy at least the following two conditions: (i) $\alpha^i_j = -1$ if $ij \notin c$, and (ii) $\alpha^i_j \leq \min_{k \in \mathcal{N}_i(c) - \{j\}} \{\alpha^i_k + d_c[i, j | \neg\{ik\}] - 1\}$ if $ij \in c$. The first condition states that player $i$ 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 player $i$ discontinues a tie with player $k$ in period $\alpha_k^i$, the grim strategy prescribes that player $k$ will discontinue all their ties in period $\alpha_k^i + 1$; player $k$'s partners discontinue their remaining ties in $\alpha_k^i + 2$, and so on. Player $i$'s non-cooperative action with player $k$ in period $\alpha_k^i$ thereby creates a contagion of non-cooperative actions in the network, and $d_c\left[i, j|\neg\{ik\}\right] - 1$ is the maximum number of periods needed for this contagion to reach player $j$.[10] Hence, player $j$ will defect on tie $ij$ in period $\alpha_k^i + d_c\left[i, j|\neg\{ik\}\right]$. Preempting the defection by $j$, player $i$ should defect on tie $ij$ latest in period $\alpha_k^i + d_c\left[i, j|\neg\{ik\}\right] - 1$. 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 $\alpha_j^i$. The maximum continuation payoff player $i$ can receive in the interaction with player $j$ is therefore be summarized by the following function:

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

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

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

$$\sum_{j \in \mathcal{I} - \{i\}} \Pi_j^i(\alpha_j^i) \leq \sum_{j \in \mathcal{N}_i(c)} \left(\frac{1 - c_{ij}}{1 - \delta} - F_{ij}^c\right) \qquad \forall \alpha^i \in \mathcal{A}_i^c. \tag{11}$$

The PrivIC 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 non-cooperative deviation. As such, when player $i$ has a profitable deviation from network

[10] The number of periods $d_c\left[i, j|\neg\{ik\}\right] - 1$ is the maximum time needed for contagion to reach player $j$ because, if player $i$ were to deviate on other ties in the meantime, defection could spread to player $j$ via other paths in the network that are shorter than the path via player $k$.

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

**Proposition 2.** *There exists a PBE $(F, \sigma)$ such that $c(F, \sigma) = c$ if and only if the network $c$ satisfies the PrivIC.*

Proposition 2 highlights the distinction between an environment with perfect and local monitoring. To see how the PrivIC is related to the PubIC, consider a network $c$ with $i \in \mathcal{N}(c)$. One defection plan that is always feasible for player $i$ is $\tilde{\alpha}^i = (-1, ..., -1)$, i.e., player $i$ does not form any ties. In that case, $\Pi^i_j(\tilde{\alpha}^i_j) = 0$ for all $j \in \mathcal{I} - \{i\}$, and so the inequality in (11) 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 public formation constraint (6). Another defection plan $\hat{\alpha}^i$ that is always feasible for network $c$ is one where $\hat{\alpha}^i_j = -1$ when $ij \notin c$, and $\hat{\alpha}^i_j = 0$ when $ij \in c$, i.e., player $i$ forms all ties in $c$ but then plans to defect on all ties immediately. 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 (11) 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 public cooperation constraint (5). As a result, the PrivIC implies the PubIC. On the other hand, the PubIC does not imply PrivIC 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, in the cycle network of 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 is feasible because the distance between nodes $i$ and $k$ in network $c$ on the path that do not use tie $ij$ is $d_c [i, k|\neg\{ij\}] = n - 1$. For this defection plan, inequality (11) in the PrivIC is equivalent to

$$1 - \bar{F} + \frac{(1 - \underline{c})(1 - \delta^{n-2})}{1 - \delta} + \delta^{n-2} - \underline{F} \leq \frac{1 - \underline{c}}{1 - \delta} + \frac{1 - \bar{c}}{1 - \delta} - \bar{F} - \underline{F}, \tag{12}$$

which is exactly inequality (10).

In general, the private cooperation constraint can depend in complex ways on global properties of the network structure, such as the distance that contagion must travel for third-party punishment to become effective. The opportunity for network enforcement diminishes when the distance between nodes in the network increases, and so local monitoring necessitates the formation of tighter networks with shorter paths, reflected in the distance constraints on defection plans in the PrivIC.

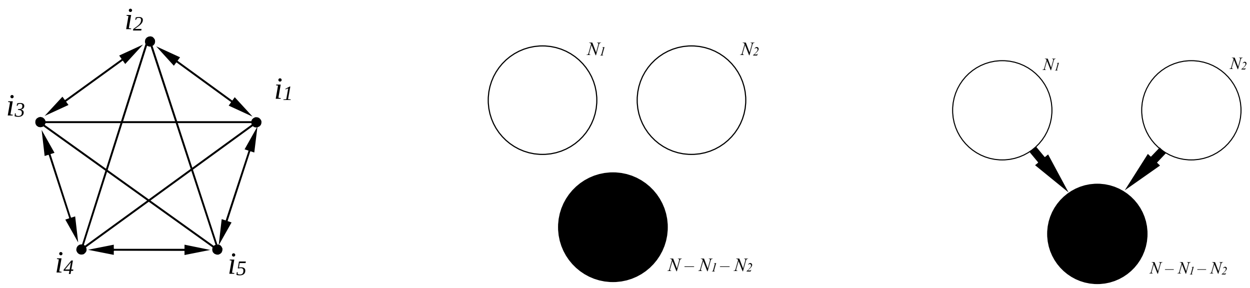
### 3.3.2 Belief-free equilibrium

The equilibrium on networks where punishment must travel long paths are sensitive to beliefs about the history of play on ties players are unable to monitor. This raises questions about the robustness of such network outcomes. We therefore consider a refinement of PBE that yields sharper predictions about network outcomes that are more robust under local monitoring. In a BFE, a strategy must be sequentially rational for *all* beliefs that are Bayesian consistent with the strategy profile. As a result, when players condition their behavior on the strategy profile, their own strategy is sequentially rational off the equilibrium path regardless of the history, and the network outcomes that can emerge involve only enforcement mechanisms that are robust to the local monitoring environment.

**Example 6.** Consider a strategy profile that prescribes the cycle network outcome $c$ in figure 3(a). Now suppose that, in some period $t$, player $j$ defects on tie $ij \in c$, but player $k$ cooperates on $ik$. What is the optimal way for player $i$ to respond? Given the strategy profile, there are (at least) two histories consistent with the observed $it$-history. 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 at least period $t+n-2$. Second, it is possible that all ties other than $ik$ have been discontinued. In that case, player $k$ will defect on $ik$ in the next period, and it is optimal to defect on $ik$ in period $t+1$. Both of these histories, which lead to different optimal responses, are in the support of Bayesian consistent beliefs. Moreover, with local monitoring, there is no way for player $i$ to distinguish between the histories: player $i$ will never know which history occurred. Whatever action player $i$'s strategy prescribes in the $it$-history after player $j$'s defection is therefore optimal given some beliefs, but is not optimal given other beliefs that are also consistent with the strategy profile and their own private history.

There is one exception: the smallest cycle with $n = 3$ (figure 2). On a cycle of length $n = 3$, the optimal period for player $i$ to defect on player $k$ is $t + 1$ for *both* of the histories described above. Moreover, 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* belief that is Bayesian consistent with the strategy profile and the private history. In that sense, the enforcement strategies that players follow on a cycle network of length $n = 3$ are robust to beliefs regarding the history in parts of the network that they 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.




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

Figure 4: Triadic closure.

*Figures 4(a)–(c) illustrate networks that satisfy triadic closure. Figure 4(a) illustrates a completely-connected network, which always satisfy triadic closure. In Figure 4(b), the empty circle indicates a cluster of nodes, all of which are connected to each other by weak ties, and the filled circle indicates a cluster of nodes, all of which are connected to each other by strong ties. In Figure 4(c), the circles indicate completely-connected clusters. The two wider arrows indicate that each player in one of the weak clusters has an asymmetric tie to each player in the strong cluster, where the tie is sponsored by the player in the weak cluster.*

**Definition 5.** 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. As we show below, our framework rationalizes triadic closure in terms of belief-free strategic interaction in a network setting with local monitoring. Our notion of triadic closure is somewhat stronger than the standard definition. In the standard definition, a network is triadically closed if there is a tie between any two nodes that have strong ties to the same partner, implying dense clustering around all strong ties (see, e.g., Easley and Kleinberg [2010, Chapter 3]). Our property requires that, whenever a player $i$ strictly sponsors a tie with player $j$ any partner $k$ of player $i$ must also be a partner of player $j$, and therefore implies dense clustering around all strong *and* asymmetric ties.

**Example 7.** A complete network (figure 4(a)) satisfies triadic closure. There are also incomplete networks that satisfy the property. For example, suppose that a subset of nodes $\mathcal{N}_1$ in a network 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 4(b)). This network consists of disjoint clusters (disjoint, completely-connected subnetworks), and therefore satisfies triadic closure. Moreover, suppose that, in addition, for any nodes $i \in \mathcal{N}_1$, $j \in \mathcal{N}_2$, and $k \in \mathcal{N} - \mathcal{N}_1 - \mathcal{N}_2$, we have that $ik$ is strictly sponsored by $i$, and $jk$ is strictly sponsored by $j$ (figure 4(c)). Then the resulting network consists of two weak clusters connected to a strong cluster by asymmetric ties. While it does not consist of independent, completely-connected cliques, this network also satisfies triadic closure. □

The PrivIC simplifies when a network satisfies triadic closure. Consider the possible defection plans for player $i$ given a network $c$. One defection plan that is always feasible is $\hat{\alpha}^i = (-1, ..., -1)$, where player $i$ forms no ties. To ensure that this defection plan is not a best-response when opponents follow the grim-strategy, network $c$ must satisfy the public formation constraints (6). The following is another defection plan that is always feasible for player $i$ under local monitoring:

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

i.e., player $i$ defects on all ties she receives in period 0, and defects on ties she sponsors in period 1. Since $\bar{\alpha}^i \in \mathcal{A}^i_c$ is a feasible defection plan, it must be the case that

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

Re-arranging terms, the feasibility of the plans $\hat{\alpha}^i$ and $\bar{\alpha}^i$ imply a robust private incentive constraint for the formation and continuation phases, respectively, that are necessary for a network outcome $c$ to emerge in a belief-free equilibrium.

**Definition 6.** Network $c$ satisfies a *robust private incentive constraint* (RPrivIC) if, for all $i \in \mathcal{N}(c)$,

$$\frac{\sum_{j \in \mathcal{N}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \leq 1 - (1 - \delta) \frac{\sum_{j \in \mathcal{N}_i(c)} F^c_{ij}}{|\mathcal{N}_i(c)|} \tag{13}$$

and

$$\frac{\frac{1}{\delta} \sum_{j \in \mathcal{R}_i(c)} c_{ij} + \sum_{j \in \mathcal{S}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \leq 1 - (1 - \delta) \mathcal{Q}_i(c). \tag{14}$$

To see how the RPrivIC is related to the PrivIC, consider a network $c$ with $i \in \mathcal{N}_i(c)$. The PrivIC is satisfied when $\sum_{j \neq i} \Pi^i_j(\alpha^i)$ is less than the continuation value of network outcome $c$ for player $i$ in every period, and for every defection plan $\alpha^i \in \mathcal{A}^i_c$. As $\mathcal{A}^i_c$ contains the defection plans $\hat{\alpha}^i$ and $\bar{\alpha}^i$, the PrivIC implies the RPrivIC. On the other hand, the RPrivIC is defined only in terms of the defection plans $\hat{\alpha}^i$ and $\bar{\alpha}^i$, and so a network can satisfy the RPrivIC even when it does not satisfy the PrivIC. In particular, the RPrivIC does not depend on the distance between nodes in the network. Indeed, the RPrivIC is more closely related to the incentive constraints under perfect monitoring. Condition (13) in the RPrivIC is exactly the public formation constraint (6). Moreover, it is easily verified that the public cooperation constraint (5) is satisfied for any network for which the robust private cooperation constraint (14) is satisfied. Hence, a network that satisfies the RPrivIC always satisfies the PubIC, but both incentive constraints–unlike the PrivIC–are independent of the distance between nodes.

The RPrivIC is more restrictive than the PubIC for the following reason. Under perfect monitoring, if player $i$ deviates from cooperation on a tie she receives in period $t$, third-parties can punish player $i$ in period $t+1$ on all ties that player $i$ sponsors. Under local monitoring, such third-party punishments are infeasible because player $i$'s deviation in period $t$ is not observed by other partners. Third-party punishment is therefore delayed for at least two additional periods. If the deviation induces third-party punishment on ties in period $t+2$, the optimal deviation by player $i$ would be to discontinue all ties she has received in period $t$, and discontinue all ties she has sponsored in period $t+1$ (preempting punishment in period $t+2$). As with the PrivIC, the difference between the RPrivIC and the PubIC is therefore due to the delay in third-party punishment, which player $i$ discounts with $\delta$.

However, as the PrivIC illustrates, the preceding intuition is incomplete. 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 the triadic closure property 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 RPrivIC is therefore equivalent to the PrivIC. The RPrivIC is sufficient to characterize incentive constraints in a BFE under local monitoring because, as the following proposition shows, belief-free equilibrium network outcomes must satisfy triadic closure.

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

*(ii) If there exists a BFE $(F, \sigma)$ such that $c(F, \sigma) = c$, then there exists a subnetwork $\tilde{c}$ of $c$ such that (i) $\tilde{c}$ contains all weak and asymmetric ties in $c$, and (ii) $\tilde{c}$ satisfies triadic closure and the RPrivIC.*

Part (i) shows when network outcomes can emerge under local monitoring in a BFE. 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 RPrivIC can be satisfied as in the perfect monitoring case. However, unlike under perfect monitoring, 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. These ties are essential to make network formation incentive compatible (condition 13). In addition, the network $\tilde{c}$ must satisfy the RPrivIC: it must contain enough strong or asymmetric ties to ensure that each player's cooperation constraint (condition 14) holds. Strong and asymmetric ties required for cooperation to be incentive compatible must therefore be clustered with other ties in $\tilde{c}$. As such, the network consists 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]). Our primary contribution is in showing that the combination of strategic interaction and robust local monitoring generates such patterns.

# 4 Examples

Under Assumption 1, the ties in a network cannot be homogeneous. In a network where all ties are weak ($c_{ij} \leq \delta$ for all $i$ and $j$), players would not have an incentive to cooperate in the continuation phase. In a network where all ties are strong ($c_{ij} \geq \delta$ for all $i$ and $j$), players would receive strictly negative payoffs in the formation phase. However, when ties are heterogeneous, players can use a network to pool incentive constraints, and Propositions 1–3 establish the necessary and sufficient restrictions on network structure for cooperative behavior to emerge. There are two potential sources of heterogeneity in our setting: (i) primitive asymmetries in the formation technology $\Gamma$, and (ii) endogenous asymmetries from differing investments during the forming phase. In this section, we consider some examples of the formation technology $\Gamma$, and illustrate the type of network outcomes that can emerge.

## 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 asymmetries in the primitives, and any heterogeneity in network 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 $\gamma$, networks cannot pool incentive constraints, and the empty network is the only equilibrium outcome.[11]

[11] To illustrate why, consider the setting with perfect monitoring, where PubIC characterizes the network outcomes that can emerge in a SPNE. Let $c$ be a non-empty network with $i \in \mathcal{N}(c)$, and let $\bar{f} \equiv \frac{\sum_{j \in \mathcal{N}_i(c)} F^c_{ij}}{|\mathcal{N}_i(c)|}$. When $\gamma$ is convex, $\frac{\sum_{j \in \mathcal{N}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \geq \gamma\left(\bar{f}\right)$. Since the PubIC requires $\frac{\sum_{j \in \mathcal{N}_i(c)} c_{ij}}{|\mathcal{N}_i(c)|} \leq \min\{\delta, 1-(1-\delta)\bar{f}\}$, it follows that 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 is a

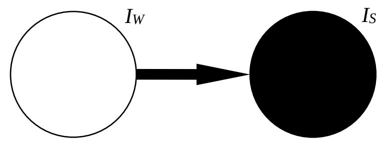



Figure 5: Endogenous asymmetries network

*The network in figure 5 is completely-connected. 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 satisfies triadic closure. By adjusting the absolute size of the population $I$, 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 RPrivIC is satisfied .*

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 strictly positive 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

$$\frac{1}{\delta}\frac{\bar{c}-\delta}{\underline{c}-\delta} \leq \min_{i\in\mathcal{I}} \mathcal{Q}_i(c) \leq \max_{i\in\mathcal{I}} \mathcal{Q}_i(c) \leq \frac{\bar{C}}{\bar{C}-\underline{C}}, \tag{15}$$

i.e., every player sponsors enough ties to ensure that the private cooperation constraint $\frac{1}{\delta}\frac{\bar{c}-\delta}{\underline{c}-\delta} \leq \min_{i\in\mathcal{I}} \mathcal{Q}_i(c)$ 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 perfect monitoring case (Example 4), as long as $\frac{1}{\delta}\frac{\bar{c}-\delta}{\underline{c}-\delta} < \frac{\bar{C}}{\bar{C}-\underline{C}}$, a complete network can 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 sufficiently large population $I$, $\min \mathcal{Q}_i(c) \approx \max \mathcal{Q}_i(c)$, and these ratios can be chosen so that condition (6) is satisfied by adjusting the relative size of the group $\mathcal{I}_S$. As a result a network outcome can emerge in a BFE where all players are connected and receive a strictly positive payoff.

violation of Assumption 1 because $\bar{f} \in [\underline{F}, \bar{F}]$. Hence, the non-empty network $c$ is an equilibrium outcome in a homogeneous formation model with convex technology if and only if players could form and cooperate on the ties in network $c$ even without any strategic linkages between ties.

## 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, $\gamma_W$ and $\gamma_B$, which represent the technology for forming ties within versus between groups. A stylized example are informal risk-sharing arrangements between 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 function $\gamma_W$ describes the technology for forming ties within a village and has both a lower intercept and smaller slope than the function $\gamma_B$, 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 leads to a strictly negative bilateral 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 (15) from the binary formation model again characterizes the RPrivIC,

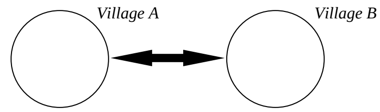



Figure 6: Risk-sharing network

*The network in figure 6 is completely-connected, and satisfies triadic closure. Within villages, farmers form weak ties. Between villages, farmers form strong ties.*

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 RPrivIC 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 $\delta \to 1$ .

The within-between group model can be used to illustrate a 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 access to formal insurance may prefer these to informal risk-sharing and, ceteris paribus, the formal 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 when the between-village ties with farmers in $A$ are removed. However, there is also 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 a 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

Another asymmetric formation technology is the target partner model, where the population is divided into groups and the technology depends on the group membership of the partner. 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."


((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 predict where large firms must be positioned in a core-periphery network.*

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, $\gamma_S$ and $\gamma_L$, which represent the technology for forming ties with small versus large firms. For simplicity, let $\gamma_L$ and $\gamma_S$ 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 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 7(a)).

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 (15) from the binary formation model again characterizes the RPrivIC. In this case, a large firm sponsors $\ell - 1$ ties and receives $ns$ ties, while the small firms sponsor $\ell$ 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 RPrivIC 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 incentive constraints 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 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.

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.

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.

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, 1770â1795.

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.

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

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.

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

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.

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

We fix parameters $\left(\bar{F}, \underline{F}, \delta, \Gamma\right)$ satisfying Assumption 1. Let $\bar{\mathbb{N}} \equiv \mathbb{N} \cup \{\infty\}$; for $\delta \in (0,1)$, $\delta^\infty \equiv 0$. For a history $h^t$, $c(h^t)$ denotes the unique network induced in period $t$ of the game; $c_i(h^t)$ denotes the set of ties of player $i$; and $c_{ij}(h^t)$ denotes player $i$'s tie with player $j$. Note that for an $it$-history $h_i^t$, player $i$'s network neighborhood is the same in every $t$-history $h^t \in h_i^t$ . Hence, we also denote by $c_i(h_i^t)$ and $c_{ij}(h_i^t)$ the network neighborhood, and tie between players $i$ and $j$, uniquely induced by the $it$-history $h_i^t$. 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 $h^t$.

## A.1 Continuation strategies

We start by defining some strategy-profiles for the continuation phase of the game, which we use in the proof of Propositions 1--3. We fix in the following a "target" network $\tilde{c}$.

We start by defining a grim-strategy for a game with public monitoring. Define the strategy profile $\sigma^{0\tilde{c}}$ as follows: for all $i, j \in \mathcal{I}$ and any history $h^t$,

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

In words, players cooperate with partners in any history $h^t$ where there is a tie between these players and the current network is the target network $\tilde{c}$, otherwise players do not cooperate with anyone.

The strategy profile $\sigma^{0\tilde{c}}$ requires public monitoring since player $i$ must be able to verify if $c(h^t) = \tilde{c}$, i.e., if the network implied by the current history is the target network. The strategy profile $\sigma^{1\tilde{c}}$ adapts $\sigma^{0\tilde{c}}$ to a setting with local monitoring: for all $i, j \in \mathcal{I}$ and histories $h^t$,

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

In words, player $i$ cooperates with partners in any history $h^t$ where there is a tie between these players and player $i$'s current network neighborhood is her target network neighborhood, otherwise player $i$ does not cooperate with anyone. This strategy profile only requires local monitoring, since player $i$ must be able to verify only if her current network neighborhood coincides with her target network neighborhood. The strategy $\sigma_i^{1\tilde{c}}$ is the grim-strategy described in Section 3.3.1. 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-response 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 $\left(F_{-i}^{\tilde{c}}, \sigma_{-i}^{1\tilde{c}}\right)$. In that case, player $i$ defects on all of her ties in period $t$ (i.e., $\sigma_{ij}^{2\tilde{c}}(h^t) = D$ for all $j \neq i$). In the second case, the actions observed by player $i$ in $h_i^t$ are consistent with other players following the strategy profile $\left(F_{-i}^{\tilde{c}}, \sigma_{-i}^{1\tilde{c}}\right)$, given the actions chosen by player $i$. In that case, we describe player $i$'s actions by means of a defection plan $\tilde{\alpha}^i = (\tilde{\alpha}_j^i)_{j \neq i} \in \bar{\mathbb{N}}^{I-1}$. As in Section 3.3.1, the interpretation is that $\tilde{\alpha}_j^i$

is the (first) period in which player $i$ defects against player $j$ in the continuation phase. Let $\mathcal{A}_i^{\tilde{c}}(h^t) \subseteq \bar{\mathbb{N}}^{I-1}$ be the set of defection plans $\alpha^i$ 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 $\mathcal{A}_i^{\tilde{c}}(h^t)$ to solve the following problem:

$$\max_{\alpha^i \in \mathcal{A}_i^{\tilde{c}}(h^t)} \sum_{j \neq \mathcal{N}_i(c(h^t))} \left( \frac{(1 - \delta^{\alpha^i_j})(1 - c_{ij})}{1 - \delta} + \delta^{\alpha^i_j} \right). \tag{16}$$

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 (16) 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 $\alpha^i_j$ 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 $\alpha^i_j$. 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 be chosen to define the strategy for player $i$ (using the axiom of choice).

The solution to the optimization problem also has the following property. Suppose $\tilde{\alpha}^i$ solves (16) 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

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

We can now complete the description of $\sigma_i^{2\tilde{c}}$. In a history $h^t$ where the actions observed by player $i$ in $h_i^t$ are consistent with other players following the strategy profile $\left(F_{-i}^{\tilde{c}}, \sigma_{-i}^1\right)$, given the actions chosen by player $i$, player $i$ cooperates with player $j$ in period $t$ if and only if $\tilde{\alpha}_j^i > t$ (i.e., for all $j \neq i$,$\sigma_{ij}^{2\tilde{c}}(h^t) = C$ if and only if $\tilde{\alpha}_j^i > t$ 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 $\sigma^{0c}$, $\sigma^{1c}$, and $\sigma^{2c}$.

**Lemma 1.** *Consider a network $c$. (i) In a game with perfect 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.2 Proof of propositions in Section 3

### Proof of Proposition 1

*Part [i]: If $c$ satisfies the PubIC, then $(F^c, \sigma^{0c})$is a SPNE.*

Consider the continuation phase. First, let $h^t$ be any history such that $c(h^t) \neq c$. The strategy profile $(F^c, \sigma^{0c})$ prescribes that all players in $\mathcal{I} - \{i\}$ defect on all ties. Hence, it is clearly optimal for player $i$ to also defect on all ties. Second, let $h^t$ be a history such that $c(h^t) = c$. The strategy profile $(F^c, \sigma^{0c})$ 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 $\left(F_{-i}^c, \sigma_{-i}^{0c}\right)$ 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, all other ties will be discontinued in the next period; anticipating this reaction, player $i$ should defect on her other ties immediately as well. 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 because $c$ satisfies the PubIC.

Now consider the formation phase. Under that strategy profile $(F^c, \sigma^{0c})$, 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 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. Hence, given the strategy profile $\left(F_{-i}^c, \sigma_{-i}^{0c}\right)$, 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 $c\left(F_i', F_{-i}^c\right) \neq c$, 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 PubIC, the continuation payoff for player $i$ under the strategy profile $(F^c, \sigma^{0c})$ 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 $h^t$, $(F^c, \sigma^{0c})$ is a SPNE. It follows from Lemma 1 that if $c$ satisfies the PubIC, there is a SPNE $(F, \sigma)$ such that $c(F, \sigma) = c$. □

*Part [ii]: If there exists a SPNE $(F, \sigma)$ such that $c(F, \sigma) = c$, then $c$ satisfies the PubIC.*

By way of contradiction, suppose $(F, \sigma)$ is a SPNE where $c(F, \sigma) = c$, and yet $c$ does not satisfy the PubIC. 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 $\sigma$ 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 $(F, \sigma)$ in either the formation or continuation phase of the game. This contradicts that $(F, \sigma)$ is a SPNE. As a result, if there is a SPNE $(F, \sigma)$ where $c(F, \sigma) = c$, then $c$ must satisfy the PubIC. □

**Proof of Proposition 2**

*Part [i]: If $c$ satisfies the PrivIC, then $(F^c, \sigma^{2c})$ is a PBE.*

To establish that $(F^c, \sigma^{2c})$ is a PBE we must find a system of beliefs $\mu$ that is consistent with $(F^c, \sigma^{2c})$ such that each player is playing a best-response for each observable history. Consider the following system of beliefs $\mu$. 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 $\left(F^c_{-i}, \sigma^{2c}_{-i}\right)$, player $i$ believes with probability one that all other players have followed the actions prescribed by the strategy profile $(F^c, \sigma^{2c})$ (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 $\left(F^c_{-i}, \sigma^{2c}_{-i}\right)$, 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 $(F^c, \sigma^{2c})$.

To show that $(F^c, \sigma^{2c}, \mu)$ 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 PrivIC. Hence, $\sigma^{2c}(h_i^t)$ is a $\mu$-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 $\mu$-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 $(F^c, \sigma^{2c})$, 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 $h^t$, $(F^c, \sigma^{2c})$ specifies that all players 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 their remaining ties in period $t$ (and all subsequent periods). As a result, $\sigma_i^{2c}(h^t)$ is a $\mu$-best response to $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$ for the $it$-history $h_i^t$. It follows that $\sigma_i^{2c}$ is a $\mu$-best response in every $it$-history $h_i^t$ of the continuation phase.

Finally, consider the formation phase. Under that strategy profile $(F^c, \sigma^{2c})$, 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. Hence, given the strategy profile $\left(F_{-i}^c, \sigma_{-i}^{2c}\right)$, player $i$'s $\mu$-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 PrivIC, one optimal plan is $\alpha^i$, 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 $(F^c, \sigma^{2c})$ after any $it$-history given the beliefs $\mu$, $(F^c, \sigma^{2c})$ is a PBE. It follows from Lemma 1 that if $c$ satisfies the PrivIC, then there is a PBE $(F, \sigma)$ such that $c(F, \sigma) = c$. □

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

By way of contradiction, suppose $(F, \sigma)$ is a PBE where $c(F, \sigma) = c$, and yet $c$ does not satisfy the PrivIC. Then on the path of play for the strategy profile $(F, \sigma)$, 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

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

and for any system of beliefs $\mu$ that is Bayesian consistent with $(F, \sigma)$ 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 $\alpha^i \in \mathcal{A}^c_i$ such that

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

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

**Proof of Proposition 3**

*Part [i]: If c satisfies triadic closure and the RPrivIC, then* $(F^c, \sigma^{1c})$*is a BFE.*

We consider possible deviations by player $i \in \mathcal{N}(c)$ for a system of beliefs $\mu$ that is Bayesian consistent with $(F^c, \sigma^{1c})$. Start with the continuation phase, and consider an *it*-history $h^t_i$ where $c_i(h^t_i) = c_i$. Then there exists a unique history $h^t \in h^t_i$ such that $\mu(h^t | h^t_i) = 1$, and $c(h^t) = c$. If player $i$ follows $\sigma^{1c}_i(h^t)$, 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 $(F^c, \sigma^{1c})$ 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 $t+1$; $(F^c, \sigma^{1c})$ 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 RPrivIC.

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 $(F^c, \sigma^{1c})$ 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 $(F^c, \sigma^{1c})$) 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 $(F^c, \sigma^{1c})$ for player $i$ in any $it$-history. The argument uses only the assumption that the system of beliefs $\mu$ is Bayesian consistent with $(F^c, \sigma^{1c})$, and so holds for any system of beliefs that is Bayesian consistent with $(F^c, \sigma^{1c})$. The strategy profile $(F^c, \sigma^{1c})$ is therefore a BFE. It follows from Lemma 1 that if $c$ satisfies triadic closure and the RPrivIC, there is a BFE $(F, \sigma)$ such that $c(F, \sigma) = c$. □

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

Let $(F, \sigma)$ be a BFE such that $c(F, \sigma) = c$. We will construct the subnetwork $\tilde{c}$ with the desired properties. Note that Assumption 1 implies that any node in $\mathcal{N}(c)$ must be a sponsor and recipient on at least one tie.

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 $(F_i, \sigma_i)$ 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 $i$ followed $(F_i, \sigma_i)$ 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 $\mu$ that is Bayesian consistent with $(F, \sigma)$ 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 clearly 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 $(F, \sigma)$ 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 $\tau$ (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 RPrivIC. 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 RPrivIC.

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

$$\frac{\frac{1}{\delta}\sum_{j\in\mathcal{R}_i(c')} c_{ij} + \sum_{j\in\mathcal{S}_i(c')} c_{ij}}{|\mathcal{N}_i(c')|} \leq 1 - (1-\delta)\mathcal{Q}_i(c'), \tag{17}$$

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 $(F, \sigma)$ 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 $(F, \sigma)$ 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 $(F_{-i}, \sigma_{-i})$, the expected continuation value of this deviation is therefore

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

The expected continuation value for following the strategy profile $(F, \sigma)$ 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 $(F, \sigma)$ 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 (17).

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

$$\frac{\sum_{i\in\mathcal{N}_i(c')} c_{ij}}{|\mathcal{N}_i(c')|} \leq 1 - (1-\delta)\frac{\sum_{j\in\mathcal{N}_i(c')} F_{ij}^c}{|\mathcal{N}_i(c')|} \tag{18}$$

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 $F_{ij} \geq F_{ij}^c$, 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 $\emptyset$ is strictly negative. But player $i$ can guarantee a non-negative continuation payoff by not investing in any ties. As a result, if $(F, \sigma)$ is a BFE, the inequality (18) must also be satisfied. $\square$

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 RPrivIC. 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 exists 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. We therefore assume, without loss of generality, that we have case (i) and can choose $i$, $j$ and $k$ such that $jk \notin c$. To complete the proof, we will show that this leads to a contradiction.

Consider a history $h^t$ where player $i$ observed actions consistent with the path of play for $(F, \sigma)$ 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 $(F, \sigma)$, there is a Bayesian consistent system of beliefs $\mu$ 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 $\sigma_{-i}$, 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, $\sigma_j$ 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 $\sigma_i$, but player $i$ discontinues $ij$ in $t+2$ instead of $t+1$. Given the system of beliefs $\mu$, 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 $(F, \sigma)$ 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. $\square$