Article

Strategic tie formation for long-term exchange relations

Werner Raub (iD)

Vincent Buskens
Utrecht University, the Netherlands

Vincenz Frey
Bocconi University, Italy

Abstract

Theory and empirical research have established that repeated interactions foster cooperation in social dilemmas. These effects of repeated interactions are meanwhile well known. Given these effects, actors have incentives for strategic tie formation in social dilemmas: they have incentives to establish long-term relations involving repeated interactions. Perhaps surprisingly, models accounting for strategic tie formation are scarce. We introduce and analyze a new game-theoretic model that captures the well-known effects of repeated interactions, while simultaneously endogenizing the formation of long-term relations. We assume strict game-theoretic rationality as well as self-regarding preferences. We highlight the commitment feature of tie formation: through establishing a long-term relation, at cost, actors ensure that they would suffer themselves from future sanctions of own opportunism. This allows for mutually beneficial cooperation in the first place. Some extensions of the model are discussed.

Keywords

Commitments, exchange relations, repeated interactions, social dilemmas, tie formation

Corresponding author:
Werner Raub, Department of Sociology/ICS, Utrecht University, Padualaan 14, 3584 CH Utrecht, the Netherlands.
Email: w.raub@uu.nl

Introduction

Repeated interactions can foster cooperation in situations with two interdependent actors involved in a social dilemma. This is a well-known result of theoretical modeling as well as empirical research (see, for example, Buskens and Raub, 2013; Raub et al., 2014, for overviews). Cooperation in repeated social dilemmas can be based on conditional behavior—reciprocity—of rational and purely self-interested actors. Therefore, when involved in a social dilemma, rational and self-interested actors have incentives for strategic tie formation, that is, for establishing a long-term relation that allows for repeated interactions.

Conditional behavior can be a mechanism of cooperation, too, in the more complex case of interactions embedded in networks of relations. When the behavior of Ego and Alter in a focal interaction may have effects for how future interaction partners of Ego and Alter behave in interactions with, respectively, Ego and Alter, that may increase incentives for cooperation of Ego and Alter in their focal interaction. Also, their behavior in the focal interaction may depend on what Ego knows about Alter's behavior in previous interactions with third parties as well on what Alter knows about previous behavior of Ego vis-à-vis third parties. Consequently, such network effects provide incentives for strategic network formation. Work that simultaneously models network effects as well as strategic network formation meanwhile includes theoretical as well as empirical research (see Jackson and Zenou, 2013, for a collection of papers on models of network effects and network formation; Goyal, 2007; Vega-Redondo, 2007 and Jackson, 2008, for textbooks; and the handbook Bramoullé et al., 2016). Moreover, this work also includes models for the specific case of cooperation in social dilemmas in a network of actors (Frey et al., 2015; Frey, 2017; Raub et al., 2013).

Given the progress with respect to integrated models of network effects and network formation, it is surprising that the more elementary case of simultaneously modeling the effects and the formation of long-term relations for two-actor social dilemmas has been largely neglected. Coleman (1990: 111, 178) did suggest that actors involved in a dilemma-like situation have an interest in creating social structures that mitigate short-term incentives for opportunistic behavior through strengthening actors' long-term concerns, but rigorous analysis supporting this intuition is lacking. For such analysis, a model is needed that comprises two ingredients. On one hand, the model must capture the well-known *effects* of repeated interactions. In this respect, the model is not original. On the other hand, however, the model must likewise capture what these effects imply in the first place for the *formation* of long-term relations—this is the major new feature of our

model. Here, assuming strict game-theoretic rationality as well as self-regarding preferences, we provide a simple model that captures the effects of repeated interactions and at the same time endogenizes the formation of long-term relations. In addition to well-known implications that refer to the effects of long-term relations, the analysis yields new implications on the deeper question concerning conditions such that actors engage in strategic tie formation and establish long-term relations because they anticipate the effects of such relations and with such effects in mind (see Prendergast, 1999: 8, 56 and Batenburg et al., 2003: 140, for similar arguments in analyses of the contractual governance of employment relations and, respectively, of buyer–supplier relations between firms).[1]

The interplay of engaging in long-term relations and coping with social dilemma problems has been studied empirically in economic sociology, broadly conceived, as well as in labor market research (see Frey et al., 2019, for references). For example, Kollock's (1994) experimental study on the emergence of exchange structures has been influential. His experiment mimics economic exchange in two different commodity markets. One of these is the market for rubber in Thailand. Kollock (1994: 314–321) points out that economic exchange between growers and buyers of rubber has social dilemma features due to quality uncertainty and asymmetric information, similar to Akerlof's (1970) classic market for "lemons," namely, used cars. At the time of sale, the buyer cannot determine the rubber's quality. Only quite some time later, after having processed the rubber, the buyer learns whether the grower has invested time and effort needed for a high-quality crop. Under such conditions and without external third-party enforcement of agreements between buyers and growers, buyers will be reluctant to pay a high price, and growers face incentives not to produce high-quality rubber. Hence, exchange partners may end up trading only low-quality rubber, while trading high-quality rubber at an appropriate price would be more beneficial for both grower and buyer. Kollock's second market is the market for rice. This is a commodity such that quality can be ascertained easily and at negligible cost at the time of exchange. Therefore, quality uncertainty and asymmetric information problems as well as incentives for opportunistic behavior are not an issue: exchange between buyers and sellers does not resemble a social dilemma but, rather, resembles exchange on a neoclassical market.

Kollock observes that long-term exchange relations between particular growers and buyers prevail on the rubber market so that the same grower and buyer engage in repeated exchange over a series of seasons. Such long-term relations are not typically found for the market for rice, where immediate and impersonal transactions prevail. Kollock's argument is that growers

and buyers on the rubber market establish long-term exchange relations because in this way incentives for opportunistic behavior are mitigated since buyers can condition their future behavior vis-à-vis the same grower on the grower's past behavior and growers can establish reputations for cooperative behavior. In his experiment with subjects in the role of buyers and, respectively, sellers, Kollock (1994: 321–326) used two conditions. One condition included quality uncertainty for buyers. In the other condition, quality was public knowledge. Kollock (1994: 328–334) indeed found that long-term relations between buyers and sellers were more prevalent under quality uncertainty.

Related research in economic sociology, using experimental as well as survey designs, has shown that social dilemma features of economic exchange induce actors to prefer partners with whom they share other and noncommercial relations (e.g. DiMaggio and Louch, 1998). Economists, stimulated by Kollock's study and likewise employing experimental designs, focused on labor market applications and on how social dilemma features of employer–employee interactions due to principal–agent problems are mitigated through the endogenous formation of long-term employment relations (e.g. Brown et al., 2004). In Kollock's study as well as subsequent work, a typical assumption is that actors can choose repeatedly between different kinds of partners, such as partners with whom they interacted previously or partners with whom they had no prior interactions. A new feature of our model is that actors can choose only once, preceding further interactions, between a sequence of interactions with the same partner or with changing partners.

A key feature of establishing long-term relations with an eye on mitigating incentives for opportunistic behavior and on fostering cooperation in social dilemmas is that strategic tie formation involves voluntary commitments. Here, other than in much of the literature in sociology and social psychology (see Kollock, 1994: 315–316, for references), "commitment" does not refer to moral obligations, feelings of attachment, or affective bonds. In our context, "commitment" refers, first, to a pattern of behavior in exchange, namely, to voluntary repeated exchanges with the same partner (Kollock, 1994: 315–316, 324–325). Second, "commitment" refers to a counterintuitive feature that merits attention: through strategic tie formation, an actor voluntarily ensures that he[2] himself would suffer from future costs of own opportunism. Through strategic tie formation, actors "bind their own hands," and in this way induce mutually beneficial cooperation. More specifically, through strategic tie formation, Ego ensures that trading partner Alter has information on Ego's past behavior. Hence, Alter can condition his behavior in future exchanges with Ego on Ego's present behavior. In particular, Alter could punish Ego

in future exchanges, should Ego behave opportunistically in their present exchange. Therefore, through strategic tie formation, Ego voluntarily reduces his own incentives for opportunistic behavior by making himself vulnerable to Alter's sanctions (see Raub, 1993: 248–249, for a similar scenario with voluntarily providing information as a means of mitigating own incentives for opportunistic behavior and thus inducing cooperation; see Gambetta and Przepiorka, 2019, for a similar approach in a related context). In this respect, strategic tie formation is a means of "hostage posting" in the sense of Schelling (1960) and Williamson (1985; see Raub, 2004, for a game-theoretic model of hostage posting as a mechanism of cooperation).

In a programmatic contribution, Coleman (1988) advocated as a strategy for theory formation in sociology "to import the economist's principle of rational action for use in the analysis of social systems proper, including but not limited to economic systems, and to do so without discarding social organization in the process" (p. S97; see also Coleman, 1986: 1323). Our article implements this strategy. We analyze how social organization, namely, long-term relations rather than spot exchange on neoclassical markets, affects exchange behavior, including economic exchange, of rational actors. The new feature of our contribution is that we model such social organization itself as a result of rational action. In this way, our article builds and improves on earlier game-theoretic and related models. Much earlier work employing rational choice principles, including but not exclusively work on cooperation in social dilemmas, likewise addresses the effects of social organization. However, this work mostly neglects to also tackle the question of how social organization can result endogenously from incentive-guided behavior.

This article is organized as follows. We first introduce our game-theoretic model. We then derive implications for effects of long-term relations and for strategic tie formation. Various extensions of the analysis are sketched. We conclude with a summary of our main results, testable predictions for experimental research, and a discussion of macro-implications of the model.

A game-theoretic model of repeated interactions and strategic tie formation

Embedded Prisoner's Dilemmas

We employ the Prisoner's Dilemma (PD) as a canonical model for a social dilemma (Figure 1). In the PD, defection (D_i) is the dominant strategy for each actor i ($i = 1, 2$) and, hence, mutual defection $D = (D_1, D_2)$ is the unique

		Actor 2	
		Cooperation (C_2)	Defection (D_2)
Actor 1	Cooperation (C_1)	R, R	S, T
	Defection (D_1)	T, S	**P, P**

Figure 1. Prisoner's Dilemma $(S < P < R < T)$. The bold-faced cell indicates the unique equilibrium.

equilibrium with payoff P for each actor.[3] Mutual defection is associated with a Pareto-suboptimal outcome since both actors are better off when they cooperate (C_i), while mutual cooperation $C=(C_1, C_2)$, although a Pareto improvement compared to D and a Pareto-optimal strategy combination, is not an equilibrium. The PD is a model for two-sided incentive problems in exchange (Hardin, 1982). For example, a seller has an incentive to sell a bad product for the price of a good one, while the buyer has an incentive to delay payment.

We now embed the PD in a repeated game Γ that is played in rounds $t=0$, 1, 2, ... The repeated game has new as well as well-known features. The new feature is that, in round 0, actors can engage in strategic tie formation. We explain the rules of the game for round 0 below, after having specified the structure of Γ for rounds 1, 2, ...

Depending on what happens in round 0, Γ continues either as a sub-game denoted by Γ^{str} or as a subgame denoted by Γ^{tie}. In round 1 of sub-game Γ^{str}, actor 1 and actor 2 play the PD. Subsequently, in rounds 2, 3, . . ., each actor i plays the PD from Figure 1 with different partners $i(2)$, $i(3)$, ... Hence, subgame Γ^{str} comprises two series of one-shot PDs. In one of these series, the long-lived actor 1 plays the PD with short-lived actors $1(2)$, $1(3)$, . . . in rounds 2, 3, . . . after the first round with actor 2. In the other series, the long-lived actor 2 plays the PD with short-lived actors $2(2)$, $2(3)$ . . . in rounds 2, 3, . . . after the first round with actor 1. Each short-lived actor plays only one PD. There is no information exchange between the actors involved in Γ^{str}.[4] This excludes third-party effects on behavior in the PDs. Due to these properties, Γ^{str} comprises PDs played by strangers.

While Γ^{str} comprises one-shot PDs, the subgame Γ^{tie} is a well-known indefinitely often repeated PD between actor 1 and actor 2 in rounds 1, 2,. . .. We employ the standard assumption that each actor i, in Γ^{tie} and in rounds $t=2, 3, . . .$, knows his own moves as well as those of the other actor j $(j=1, 2; i \neq j)$ in all previous rounds $1, . . ., t-1$.

Remark. Note that the important feature of Γ^{str} is not that each actor i plays the PD only once with j and with each actor $i(2)$, $i(3)$, ... Also, we need not assume that information exchange between the actors involved in Γ^{str} is completely excluded. The important feature is that cooperation is not supported by game-theoretic equilibrium behavior in Γ^{str}. Rather, equilibrium behavior must imply that actors defect throughout all rounds 1, 2, . . . in Γ^{str}. Our assumptions on the structure of Γ^{str} do ensure this property for the subgame and keep the model simple.

Strategic tie formation

Rounds 1, 2, . . . of Γ are preceded by round 0. This round is closely related to the core new feature of our model. In round 0, actors 1 and 2 can choose between playing the subsequent series of PDs with different partners, that is, to play subgame Γ^{str} of one-shot PDs, or they can choose to play the subsequent series of PDs with each other and to enter subgame Γ^{tie}. Thus, in round 0, actors 1 and 2 can invest in strategic tie formation and can establish a long-term relation in which the actors play repeated PDs with each other. With an eye on examples from economic sociology and labor market research, we could interpret establishing such a long-term relation as setting up an agreement for repeated transactions, a joint venture, or as entering a long-term employment relation.

We assume that strategic tie formation is associated with total costs $\tau > 0$. Assuming costless tie formation would trivialize the problem of accounting for social organization in the sense that no trade-off would be involved. Different scenarios for allocating these costs can be conceived as different kinds of institutions: they are rules of the game in which the actors are involved (North, 1990: Chapter 1). Seen from this perspective, we model effects of institutions for establishing long-term relations as well as how institutions and long-term relations interact in affecting cooperation in social dilemmas. Our results for strategic tie formation will reveal that it depends on the parameters of the PD and of Γ whether actors are willing to invest in strategic tie formation under a given institution. Hence, our model allows, too, for inferences on institutional design for fostering cooperation in social dilemmas.[5]

For our model, we employ a simple sharing institution for the costs of strategic tie formation. More precisely, we assume that actor 1 and actor 2 decide simultaneously and independently in round 0 of Γ about their individual investment. Each actor can either invest $\tau/2$ or decide not to invest. If each actor invests $\tau/2$, the strategic tie is formed and the actors enter subgame Γ^{tie} of repeated PDs. Conversely, if at least one actor decides not to

invest, the strategic tie is not formed and the actors enter subgame Γ^{str} of one-shot PDs. In the latter case, an actor who had been willing to invest does not lose his investment. After round 0 and before playing round 1, both actors are informed about each actor's decision in round 0.[6] Our institution for sharing the costs of strategic tie formation corresponds to common assumptions in models of strategic network formation, namely, two-sided link formation (a link is only formed if both actors wish to be linked) with shared costs of links (Jackson, 2008: Chapter 6).

Further assumptions on Γ

Round 1 is always played after round 0. After each round $t = 1, 2, \ldots$ of Γ, the next round $t + 1$ is played with a constant probability w $(0 < w < 1)$, while Γ stops after each round with probability $1 - w$. Actor i's (expected) payoff for Γ therefore equals the sum of tie formation costs, if any, in round 0 and the exponentially discounted payoffs in rounds $1, 2, \ldots$. This implies, for example, that actor i's payoff is

$$U_i = -\frac{\tau}{2} + R + wR + w^2 R + \ldots + w^{t-1} R + \ldots = -\frac{\tau}{2} + \frac{R}{1-w}$$

if i and j both invest in strategic tie formation and if they subsequently cooperate in all PDs in rounds $1, 2, \ldots$. Similarly

$$U_i = P + wP + w^2 P + \ldots + w^{t-1} P + \ldots = \frac{P}{1-w}$$

is actor i's payoff if round 0 of Γ does not result in strategic tie formation and if i as well as all partners of i defect in all PDs in rounds $1, 2, \ldots$.[7]

We further assume that all actors are informed on the structure of Γ (and thus also on the structure of the PD), that they know from each other that they have this information, and so on. The structure of Γ is thus "common knowledge." Finally, we assume that Γ is played as a noncooperative game, that is, actors cannot incur binding agreements or binding unilateral commitments that are not explicitly modeled as moves in the structure of the game.[8]

Analysis of the model

We are interested in conditions for cooperation in social dilemmas as well as conditions for strategic tie formation. Therefore, since we assume rational

behavior of actors, we derive conditions for subgame perfect equilibria[9] of Γ such that both actors invest in strategic tie formation in round 0 of Γ and subsequently cooperate throughout all rounds 1, 2, . . . in all PDs. We refer to such an equilibrium as a "cooperation equilibrium." Mutual defection throughout the repeated game is likewise an outcome of equilibrium behavior. The so-called "folk theorem" implies that the repeated game may have many other equilibria, too. We employ the commonly used assumption for the analysis of an indefinitely often repeated PD that an equilibrium such that both actors cooperate throughout all rounds 1, 2, . . . can be considered as the "solution" of the game because each actor maximizes his own payoff, given the equilibrium strategies of all actors *and* because such an equilibrium is associated with higher payoffs for each actor than the situation where actors defect throughout all PDs in all rounds 1, 2, ... To specify conditions for cooperation equilibria, using backward induction, we first consider equilibrium behavior in subgame Γ^{str} of one-shot PDs and in subgame Γ^{tie} of repeated PDs. Subsequently, we derive conditions such that strategic tie formation in round 0 is implied by equilibrium behavior.

Equilibrium behavior in Γ^{str} and Γ^{tie}

Results on equilibrium behavior in subgames Γ^{str} and Γ^{tie} are well known from earlier literature. Since Γ^{str} comprises only one-shot PDs without information exchange between the actors involved, we have:

Proposition 1—defection without tie formation. Equilibrium behavior in subgame Γ^{str} implies that all actors defect in all PDs in all rounds.

We now turn to subgame Γ^{tie}. Given our assumptions on Γ, we can apply standard theory on indefinitely often-repeated games (e.g. Friedman, 1986) for the analysis of Γ^{tie}. In Γ^{tie}, cooperation can be the result of equilibrium behavior based on reciprocity and conditional strategies. When playing the PD in round $t = 2, 3, . . .$, actor i can condition his behavior on information about the other actor's behavior in rounds $1, . . ., t-1$. If j cooperated in earlier rounds, i can reward this by cooperating himself in round t. Conversely, if j defected in earlier rounds, i can punish this by defecting himself in round t. Thus, with $T > R$, actor j has a short-term incentive to defect rather than cooperate in each PD against a cooperating partner. However, in Γ^{tie} he also has to take into account that defecting in the current round may imply long-term costs in future rounds since actor i may then defect in at least some of those future rounds. Actor j could obtain at best $P < R$ in those future rounds, given i's defection. Anticipating conditional behavior of i, a rational actor j thus has to balance short-term incentives

$(T-R)$ and long-term incentives $(R-P)$. Mutual cooperation in the repeated PD can then be a result of individually rational behavior of actors who take long-term effects of their present behavior into account.

These properties of repeatedly played PDs highlight the commitment feature of strategic tie formation. Through tie formation, each actor j voluntarily ensures that the other actor i receives information on j's previous behavior. This implies that i can condition his future behavior vis-à-vis j on j's present behavior so that j voluntarily makes himself vulnerable to future punishment by i. In particular, j makes himself vulnerable to future punishment by i in case j behaves opportunistically and defects, thus voluntarily reducing j's own benefits from defection. Binding one's own hands in this way can be rational behavior, though, because it induces cooperation of the other actor in the first place. While j's commitment prohibits that j can benefit unilaterally from exploiting i's cooperation through own defection, the commitment does allow for mutual cooperation, which is more attractive for j than mutual defection.

The conditional strategy that is associated with the most attractive reward for the other actor if he cooperates and the most severe punishment for defection is commonly known as a "trigger strategy" (e.g., Friedman, 1986). This is the strategy such that the actor cooperates in round 1 and cooperates in future rounds 2, 3, . . . if he has no information on a defection in a previous round. However, after the first defection has occurred in some round t, the trigger strategy requires that the actor defects in all future rounds $t + 1$, $t + 2$, ... Obviously, if both actors use trigger strategies, they cooperate in all rounds of Γ^{tie}. If long-term incentives can at all be large enough to induce a rational actor to cooperate, these incentives are certainly large enough if the other actor uses a trigger strategy. It follows that an indefinitely often repeated PD has an equilibrium such that both actors cooperate throughout all rounds if and only if triggers strategies used by both actors constitute an equilibrium (see Friedman, 1986: Chapter 3, for details).

For empirical applications, the assumption need not be that actors do indeed use trigger strategies. Cooperation of rational actors in Γ^{tie} does require the use of conditional strategies but these conditional strategies may comprise less severe punishments than implied by a trigger strategy. For example, an actor may be willing to return to cooperation if the "punishment period" for the other actor has been "long enough." Nevertheless, the existence of a trigger strategy equilibrium is a necessary condition for the existence of equilibria based on conditional strategies that imply cooperation throughout the repeated game but involve less extreme sanctioning and thus smaller long-term incentives for honoring trust. Hence, we do assume that cooperation becomes more likely when the conditions for a trigger strategy equilibrium become less restrictive. This is a common assumption,

although often left implicit, in empirical applications (see Buskens and Raub, 2013, for a more detailed discussion of this approach). Indeed, results of a meta-analysis of experiments with indefinitely often-repeated PDs support that cooperation increases in the probability w of future rounds and when the repeated game has an equilibrium such that actors cooperate when implementing the equilibrium strategies (Dal Bó and Fréchette, 2018: 68–71, Results 1 and 2).

The following proposition is a direct implication of the fundamental theorem on trigger strategy equilibria in indefinitely often-repeated games (see, for example, Friedman, 1986: 88–89, for details).[10]

Proposition 2—cooperation after tie formation. Γ^{tie} has an equilibrium such that both actors cooperate throughout all rounds if and only if $w \geq TEMP = (T - R) / (T - P)$.

Note that *TEMP* is a measure for the actors' incentives ("temptation") to defect. It is easily seen that $0 < TEMP < 1$. Proposition 2 shows that cooperation in an indefinitely often-repeated PD constitutes equilibrium behavior if and only if the incentives to defect are compensated by a sufficiently large probability w that Γ continues. The proposition implies that cooperation is facilitated—in the sense that the condition $w \geq TEMP$ becomes less restrictive—if the short-term incentive $T-R$ to defect decreases, if there are increasing costs $T-P$ of a conflict in the sense that both actors simultaneously try to exploit the partner's cooperation by own defection, if the costs $R-P$ of mutual defection compared to mutual cooperation increase, and if the continuation probability w increases. In Axelrod's (1984) formulation, w represents the "shadow of the future": with increasing w, actors' long-term incentives become more important relative to short-term incentives. Note that there always exists sufficiently large w as well as sufficiently small $T-R$ or sufficiently large $T-P$ so that the condition in Proposition 2 is fulfilled.

It should be noted that an indefinitely often-repeated PD, other than a finitely repeated PD, allows for equilibria that imply cooperative behavior without assuming that actors, with some positive probability, have other-regarding preferences (e.g. Fehr and Schmidt, 2006). Given parameters of the game such that the condition in Proposition 2 is fulfilled, cooperation in an indefinitely often repeated PD can be based exclusively on conditional behavior and enlightened self-interest of rational actors, with "enlightened" indicating that actors take long-term effects of own present behavior into account.[11] A meta-analysis of experiments with indefinitely often-repeated PDs (Dal Bó and Fréchette, 2018: 87–89, Result 7) suggests that other-regarding preferences are indeed less important drivers of behavior in such games, with cooperation being mainly driven by actors' strategic behavior.

Strategic tie formation

Having analyzed the effects of strategic tie formation, we now turn to the new feature of our model, namely, investments in tie formation. We do so by specifying conditions for equilibria of Γ that imply tie formation in round 0 and cooperation in subsequent rounds 1, 2, ...

Without tie formation, rational actors defect in subgame Γ^{str} and hence $U_i = P/(1-w)$ is the payoff for actor 1 as well as for actor 2 in Γ^{str}. On the other hand, there are parameter configurations such that cooperation in Γ^{tie} after tie formation in round 0 can be the result of equilibrium behavior. If actor 1 and actor 2 cooperate in Γ^{tie}, each actor's payoff in Γ^{tie} is $U_i = R/(1-w)$. This yields the following new proposition.

Proposition 3—value of strategic tie and net gain from tie formation. Assume that $w \geq TEMP$ so that Γ^{tie} has an equilibrium such that both actors cooperate throughout all rounds. The value of strategic tie formation is then $(R-P)/(1-w)$ for each actor i, while $(R-P)/(1-w)-(\tau/2)$ is the net gain from tie formation for each actor i.

Under our assumptions on equilibrium behavior, the value of strategic tie formation is equal to the difference between the equilibrium payoffs from rounds 1, 2, ... The value of tie formation for actor i thus increases if the costs $R-P$ of mutual defection compared to mutual cooperation increase and if the continuation probability w increases. Note that $(R-P)/(1-w)$ is also the upper bound for the costs of tie formation that a rational actor would be willing to incur.

Proposition 4—investments in strategic tie. Γ has a cooperation equilibrium such that actor 1 and actor 2 invest in tie formation in round 0 and subsequently cooperate in all rounds 1, 2, . . . if and only if

$$w \geq TEMP \text{ and} \tag{1}$$

$$\tau \leq 2\frac{R-P}{1-w}. \tag{2}$$

This is the core new result due to our model. The proposition follows from Propositions 1 to 3, the institution for sharing the costs of strategic tie formation, and the definition of subgame perfect equilibrium. Each actor's equilibrium strategy for Γ is to invest in strategic tie formation in round 0 of Γ, to defect unconditionally in each subgame Γ^{str} and to use a trigger strategy in subgame Γ^{tie}.

Proposition 4 shows that rational actors can invest in strategic tie formation if their cooperation problems are not too large and if the costs of the investment are sufficiently small. Small enough cooperation problems in the sense of $w \geq TEMP$ can be solved with investments in tie formation. Too large cooperation problems in the sense of $w < TEMP$ cannot be solved even with investments in tie formation. In addition, for small enough cooperation problems, each actor's individual costs of strategic tie formation must not exceed the actor's gains from mutual cooperation in Γ^{tie} rather than mutual defection in Γ^{str}.

Some extensions

For simplicity, we used the PD for modeling social dilemmas. Moreover, we have assumed homogeneity in the sense that both actors have the same payoff function for the PD. It would be no problem to relax the homogeneity assumption and adapt our propositions, with core results remaining robust. In the following, we briefly consider (1) extensions of the analysis to other games than the PD, (2) other institutions for the costs of strategic tie formation as well as collective good problems associated with strategic tie formation, and (3) the case of games with incomplete information.

First, one can show robustness of results for a large class of games other than the PD. More precisely—to be shown in more detail in a future article—results analogous to those for the PD hold for repeated games with trigger strategy equilibria in the sense of Friedman (1986: Chapter 3; see Raub et al., 2013: 722–724, for results in the related context of network effects on cooperation in social dilemmas and strategic network formation), including games modeling social dilemmas with 2 as well as $n > 2$ actors.

As an example, consider the Trust Game (Dasgupta, 1988; Kreps, 1990). In this game, actor 1 is the trustor and actor 2 is the trustee. The trustor moves first, choosing between placing or not placing trust. If trust is not placed, the interaction ends, with trustor and trustee each receiving payoff P. If trust is placed, the trustee chooses between honoring and abusing trust. If he honors trust, trustor and and trustee each receive payoff $R > P$. If trust is abused, the payoff for the trustor is $S < P$, while the trustee receives $T > R$. Hence, not placing trust, while placed trust would be abused, is the unique subgame perfect equilibrium of the game. Both actors would be better off, though, if trust would be placed and honored, since $R > P$. The Trust Game thus models a social dilemma. In terms of economic exchange, Kollock's rubber example resembles a Trust Game, with the buyer as trustor and the grower as trustee. In the Trust Game, only the trustee has an incentive for opportunistic behavior. After all, while the trustee can secure $T > R$ by abusing trust, the trustor's best reply against a trustee who honors trust would be

to place trust, since $R > P$. The trustor's equilibrium strategy not to place trust implies protection against the trustee's opportunism rather than an attempt to increase the trustor's payoff by exploiting behavior of the trustee. While incentives for opportunism are one-sided in the Trust Game, they are two-sided in the PD. In the PD, defection is a means to exploit the other actor's cooperation, thus securing payoff $T > R$. However, defection like-wise avoids being unilaterally exploited, thus securing payoff $P > S$.

Now assume that Trust Games rather than the PD are played in rounds 1, 2, . . . of Γ. Strategic tie formation in round 0 implies that trustor and trustee play the Trust Game indefinitely often (and without changing roles) in Γ^{tie}, while the subgame Γ^{str} would include a Trust Game between actor 1 as trustor and actor 2 as trustee in round 1, actor 1 would play one-shot Trust Games with different trustees 1(2), 1(3), . . . in rounds 2, 3, . . . and actor 2 would play one-shot Trust Games with different trustors 2(2), 2(3), . . . in rounds 2, 3, ... Clearly, equilibrium behavior in Γ^{str} then implies that no trust would be placed, while Γ^{tie} has an equilibrium such that trust is placed and honored in all rounds if and only if $w \geq TEMP$. In this equi-librium, the trustor places trust as long as trust has been always honored, threatening abuse of trust with not placing trust in future Trust Games. Then, under the cost sharing rule for investments in strategic tie formation we have used in section "Strategic tie formation" and assuming $w \geq TEMP$, an equilibrium exists such that trustor and trustee invest in strategic tie formation and subsequently place and honor trust throughout all rounds 1, 2, . . . if $\tau \leq 2(R - P) / (1 - w)$.

Next, consider extensions of the analysis that relate to new features of institutions for the costs of strategic tie formation and to collective good problems associated with strategic tie formation. A potential problem with respect to rational actors' investments in strategic tie formation is that the long-term relation with repeated interactions due to tie formation has col-lective good properties. Consider both the PD and the Trust Game. In prin-ciple, actor i could benefit from the long-term relation even if i did not contribute himself to the costs of establishing the long-term relation. Therefore, in principle, there are incentives for free riding and risks of sub-optimal investments in strategic tie formation. Our institutional rule for allocating costs in round 0 implies that opportunities and incentives for free riding are mitigated. Specifically, the long-term relation is established if and only if both actors invest and an actor who is willing to invest does not lose his own investment if the other actor refuses to invest, too. Thus, an actor cannot unilaterally exploit the other actor's investment in strategic tie for-mation. This ensures that investing in round 0, under the conditions in Proposition 4, is consistent with equilibrium behavior.

To see how institutional rules for allocating costs in round 0 can induce incentive problems, consider another simple example: actor 1 and actor 2 decide simultaneously and independently about their individual investment. Each actor can either invest τ or decide not to invest. The strategic tie is formed if and only if at least one actor decides to invest. Each actor who decides to invest, has to pay τ, irrespective of the other actor's choice in round 0. Under such an institution, for $\tau \leq (R - P)/(1 - w)$ and $w \geq TEMP$, each actor would prefer the strategic tie to be formed but would likewise prefer the other actor to invest in strategic tie formation, while not investing himself. Hence, a bargaining problem arises. Also, under the same conditions, when both actors invest, they could subsequently benefit from conditional cooperation but they would waste resources in round 0 in the sense that the investment by only one actor would suffice to enter subgame Γ^{tie} and to be able to benefit from conditional cooperation.

We highlighted the commitment feature of strategic tie formation. In the PD, incentives to defect are two-sided. For this reason, an actor's investments in strategic tie formation do contribute to mitigating that actor's incentives for own opportunistic behavior but likewise to mitigating opportunistic behavior of the other actor. The Trust Game is useful for distinguishing between both motives for strategic tie formation, since incentives for opportunistic behavior are one-sided in this game.

If tie formation would require that only the trustor invests in round 0, an equilibrium exists such that the trustor invests and subsequently trust is placed and honored throughout all rounds if $w \geq TEMP$ and $\tau \leq (R - P)/(1 - w)$ so that the costs of strategic tie formation are low enough for the trustor. On the other hand, under an institution such that only the trustee could invest in strategic formation, there is also an equilibrium such that the trustee alone bears the full costs in round 0 if $w \geq TEMP$ and $\tau \leq (R - P)/(1 - w)$, with trust placed and honored in all subsequent rounds. This shows that it can be rational behavior for an actor who has one-sided incentives for opportunism to commit himself by incurring the full costs of strategic tie formation. Tie formation makes it less attractive for the trustee to behave opportunistically, due to future punishment. This can induce the trustor to place trust in the first place, with gains for both trustors and trustees through trust being placed and honored, compared to the payoffs they obtain when no trust is placed.

One can easily verify that under each of the institutions for allocating the costs of strategic tie formation considered so far, the repeated game Γ always has an equilibrium such that actors do not invest in tie formation and subsequently defect throughout all rounds 1, 2, ... Hence, an equilibrium selection problem emerges not only for the subgame Γ^{tie} but also for Γ itself. Consequently, payoff dominance arguments are needed not only with respect

to conditional cooperation as a solution of Γ^{tie}. Such arguments are also needed with respect to strategic tie formation in round 0 and conditional cooperation in all subsequent rounds as a solution of Γ itself: in both cases, one needs to assume that actors tacitly coordinate on the equilibrium that makes them better off. This is so not only for the PD as a model for a social dilemma. Obviously, the situation is analogous for the other games mentioned above, such as the Trust Game, that are models for social dilemmas.

Finally, we have assumed that actors are *always* rational and would *always* defect in one-shot PDs and also have complete information on each other's preferences. Albeit defensible (see the remark above on other-regarding preferences vs enlightened self-interest as drivers of cooperation in indefinitely often-repeated PDs), this is a simplifying assumption. An implication of this assumption is that strategic tie formation serves exclusively as a means that allows for sanctioning actors' present behavior in future rounds of Γ. Our model, therefore, neglects that strategic tie formation may serve another purpose, too. Namely, via information on how actor *j* has behaved in previous rounds, actor *i* may learn about unobservable characteristics of *j*, such as whether *j* is indeed rational and would indeed always defect in one-shot PDs.

Finitely repeated games with incomplete information (e.g. Kreps et al., 1982) can be used to address these issues. One then assumes that actors do not know for sure about characteristics of their partner such as his preferences or whether he behaves rationally. Rather, actors only know the probability of having a partner with certain characteristics. The partner's behavior can then be used for inferring his unobservable characteristics. Kreps et al. (1982) have shown how to derive conditions for cooperation in finitely repeated games and for learning about unobservable characteristics of the partner in such games. Using an approach analogous to Frey et al. (2015) and Frey (2017), one could derive conditions for strategic tie formation and subsequent cooperation for a finitely repeated PD as well as a finitely repeated Trust Game (see Przepiorka and Diekmann, 2013, for a somewhat similar model, including experimental work, for Trust Games with incomplete information). Note that this includes analyzing learning about unobservable characteristics of the partner not only in the finitely repeated PD or in the finitely repeated Trust Game. Rather, it includes analyzing, too, if actors can learn about characteristics of the partner also from the partner's behavior in round 0. Assuming a finitely repeated game with incomplete information, investments in strategic tie formation still serve commitment purposes in the sense of making an actor voluntarily vulnerable to future sanctions of the partner (the "binding effect" of commitments, see Raub, 2004). In addition, the question arises (see Frey, 2017, for how to approach the issue in the context of strategic network formation) whether

investments in strategic tie formation likewise allow for inferences about unobservable characteristics of the actor who is willing to invest in strategic tie formation (the "signaling effect" of commitment, see Raub, 2004).

Discussion

Employing the theoretical strategy to simultaneously analyze how social organization—in our case, long-term relations with repeated interactions—affects the behavior of rational actors and how social organization results from rational action, we have developed a new game-theoretic model of strategic tie formation. The model captures well-known results on conditions such that rational actors involved in social dilemmas cooperate. Simultaneously—and this is the major new feature—the model endogenizes strategic tie formation by specifying conditions such that rational actors incur costs to establish a long-term relation involving repeated interactions in the first place. The model likewise allows to explicitly specify the value of strategic tie formation. We have emphasized another noteworthy feature, namely, that strategic tie formation involves voluntary commitment. For one, actors commit to repeated interactions with the same partner rather than one-shot interactions with different partners. Likewise, actors incur a commitment in the sense of making it unattractive for themselves to profit from opportunistically exploiting the partner's cooperation.

Our model uses the PD as a canonical example of a social dilemma. The model yields empirically testable predictions on effects of changes in the parameters *T, R*, and *P* of the PD, the costs τ of strategic tie formation, and the continuation probability *w*. Three kinds of predictions are topical. These refer, first, to effects on tie formation, second, to effects on behavior in social dilemmas, and, third, to effects on the relation between tie formation and behavior in social dilemmas. Predictions on effects of changes in *T, R, P*, and *w*, on behavior in social dilemmas are similar to predictions that follow from earlier models (see Buskens and Raub, 2013). Other predictions are new and highlight the added value of our model.

For experimental tests of the model, we can distinguish between two scenarios. The first covers $w < TEMP$. Here, cooperation problems as represented by *TEMP* are large relative to *w*. We predict a small likelihood of tie formation, while the likelihood of tie formation is only weakly associated with the costs τ. We also predict a small likelihood of cooperation as well as small effects of tie formation on subsequent behavior. The second scenario covers $w \geq TEMP$. Cooperation problems are now smaller relative to *w* but rational cooperation presupposes repeated interactions. We predict sizable effects of τ on tie formation as well as sizable effects of tie formation on subsequent behavior in PDs. More specifically, we predict that the

likelihood of tie formation decreases with increasing costs τ. Likewise, we predict that tie formation has a positive effect on the likelihood of subsequent cooperation. These predictions correspond nicely to predictions for network effects on cooperation in social dilemmas and strategic network formation (Raub et al., 2013: 728).

Note that it would be additionally interesting to experimentally compare scenarios such that subjects can mitigate cooperation problems by investing in strategic tie formation with scenarios such that subjects, at costs, can secure external enforcement for agreements on behavior in one-shot interactions. We could easily derive predictions for conditions such that rational actors would prefer investing in strategic tie formation or would prefer investing in external enforcement or would be indifferent between both modes of mitigating cooperation problems.

By way of conclusion, it is useful to highlight macro-implications of the analysis that transcend cooperation in social dilemmas as such. From an economic sociology perspective on market transactions, these implications include that market structure differs depending on whether or not exchange comes with incentives for opportunistic behavior. Without such incentives, a market with one-shot transactions may prevail. With incentives for opportunism, however, one would expect a market structure characterized by long-term relations between market participants (see Kollock, 1994: 313, 320–321, as well as later work in economics, such as Brown et al., 2004). Another likely implication refers to effects of strategic tie formation on inequality. If actors i and j engage in strategic tie formation and exchange repeatedly in subgame Γ^{tie}, they benefit themselves from mutual cooperation and payoffs R rather than P in each transaction. At the same time, they exclude other actors k from exchanges with i and j, thus possibly inducing inequality since these other actors miss opportunities for exchange or at least forego the benefits of mutual cooperation in exchange (see Frey and Van de Rijt, 2016, for an analysis of this issue in a related context, including empirical evidence from an experiment).

ORCID iD

Werner Raub [iD] https://orcid.org/0000-0001-5425-336X

Notes

 1. Strictly speaking, since rational choice models are about actors' behavior, rather than about cognitive processes related to behavior, a more precise formulation might be that the analysis yields implications for the conditions such that actors behave *as if* they anticipate the effects of long-term relations and *as if* they have such effects in mind.
 2. "He," "his," and so on, are used to facilitate readability and without intending gender bias.
 3. See a textbook, such as the one by Rasmusen (2007), for game-theoretic concepts and assumptions that are employed in our analysis.
 4. Given our assumptions, two PD's are played in each round 2, 3, . . . of Γ^{str}, one involving actor 1, the other involving actor 2. We do not need to further specify in which sequence these two PDs are played per round.
 5. Cooperation in social dilemmas is beneficial for the actors directly involved but need not be beneficial for third parties. Members of the Mafia (e.g. Gambetta, 1993) or cartel members (e.g. Stigler, 1964) are involved in social dilemma-like interactions. In these contexts, cooperation is beneficial for the members, while undermining cooperation is desirable from a societal perspective. Therefore, from a societal perspective, institutional design need not aim at fostering but, depending on the specific context, may aim at mitigating cooperation.
 6. Note that, given our cost sharing institution for strategic tie formation, Γ has one subgame Γ^{tie}, while Γ has three subgames Γ^{str}.
 7. Note that we interpret payoffs as cardinal utilities. Note, too, that the model includes discounting of future payoffs due to the probability that the game might end and that we neglect negative time preferences. It would be no problem to include negative time preferences and results would remain robust.
 8. We assume a noncooperative game precisely because we wish to specify conditions such that rational actors will cooperate "endogenously" and without external enforcement in social dilemmas, based exclusively on the embeddedness of the dilemma in a long-term relation.
 9. Subgame perfection is the basic refinement of the Nash equilibrium concept and a common conceptualization of individually rational behavior in situations with strategic interdependence. Henceforth, we exclusively consider subgame perfect equilibria and refer to them for brevity as "equilibria."
10. Basically, it suffices to observe that if actor i uses a trigger strategy, it is a best reply for actor j to likewise use a trigger strategy if it does not pay off for j to defect immediately in round 1. The equilibrium condition $w \geq TEMP$ follows.
11. Coleman (1964) has characterized "enlightened self-interest" when conceiving of "socialization" in a way that contrasts with the common view of "internalization of values," namely, socialization as allowing an actor "to see the long-term consequences to oneself of particular strategies of action, thus becoming more completely a rational, calculating man" (p. 180).

References

Akerlof GA (1970) The market for "lemons": quality uncertainty and the market mechanism. *The Quarterly Journal of Economics* 84: 488–500.

Axelrod R (1984) *The Evolution of Cooperation*. New York: Basic Books.

Batenburg RS, Raub W and Snijders C (2003) Contacts and contracts: temporal embeddedness and the contractual behavior of firms. *Research in the Sociology of Organizations* 20: 135–188.

Bramoullé Y, Galeotti A and Rogers B (2016) *The Oxford Handbook of the Economics of Networks*. Oxford: Oxford University Press.

Brown M, Falk A and Fehr E (2004) Relational contracts and the nature of market interactions. *Econometrica* 72: 747–780.

Buskens V and Raub W (2013) Rational choice research on social dilemmas. In: Wittek R, Snijders TAB and Nee V (eds) *Handbook of Rational Choice Social Research*. Stanford, CA: Stanford University Press, pp. 113–150.

Coleman JS (1964) Collective decisions. *Sociological Inquiry* 34(2): 166–181.

Coleman JS (1986) Social theory, social research, and a theory of action. *American Journal of Sociology* 91: 1309–1335.

Coleman JS (1988) Social capital in the creation of human capital. *American Journal of Sociology* 94: S95–S120.

Coleman JS (1990) *Foundations of Social Theory*. Cambridge, MA: Belknap Press of Harvard University Press.

Dal Bó P and Fréchette G (2018) On the determinants of cooperation in infinitely repeated games: a survey. *Journal of Economic Literature* 56: 60–114.

Dasgupta P (1988) Trust as a commodity. In: Gambetta D (ed.) *Trust: Making and Breaking Cooperative Relations*. Oxford: Blackwell, pp. 49–72.

DiMaggio P and Louch H (1998) Socially embedded consumer transactions: for what kinds of purchases do people most often use networks? *American Sociological Review* 63: 619–637.

Fehr E and Schmidt KM (2006) The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In: Kolm S-C and Ythier JM (eds) *Handbook of the Economics of Giving, Altruism and Reciprocity*. Amsterdam: Elsevier, pp. 615–691.

Frey V (2017) Boosting trust by facilitating communication: a model of trustee investments in information sharing. *Rationality and Society* 29: 471–503.

Frey V, Buskens V and Corten R (2019) Investments in and returns on network embeddedness: an experiment with trust games. *Social Networks* 56: 81–92.

Frey V, Buskens V and Raub W (2015) Embedding trust: a game-theoretic model for investments in and returns on social capital. *Journal of Mathematical Sociology* 39: 39–72.

Frey V and Van de Rijt A (2016) Arbitrary inequality in reputation systems. *Nature Scientific Reports* 6: 38304.

Friedman JW (1986) *Game Theory with Applications to Economics*. New York: Oxford University Press.

Gambetta D (1993) *The Sicilian Mafia: The Business of Private Protection*. Cambridge, MA: Harvard University Press.

Gambetta D and Przepiorka W (2019) Sharing compromising information as a cooperative strategy. *Sociological Science* 6: 352–379.

Goyal S (2007) *Connections: An Introduction to the Economics of Networks*. Princeton, NJ: Princeton University Press.

Hardin R (1982) Exchange theory on strategic bases. *Social Science Information* 21: 251–272.

Jackson MO (2008) *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Jackson MO and Zenou Y (eds) (2013) *Economic Analyses of Social Networks*. London: Edward Elgar.

Kollock P (1994) The emergence of exchange structures: an experimental study of uncertainty, commitment, and trust. *American Journal of Sociology* 100: 313–345.

Kreps DM (1990) Corporate culture and economic theory. In: Alt J and Shepsle KA (eds) *Perspectives on Positive Political Economy*. Cambridge: Cambridge University Press, pp. 90–143.

Kreps DM, Milgrom P, Roberts J, et al. (1982) Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory* 27: 245–252.

North DC (1990) *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Prendergast C (1999) The provision of incentives in firms. *Journal of Economic Literature* 37: 7–63.

Przepiorka W and Diekmann A (2013) Temporal embeddedness and signals of trustworthiness: experimental tests of a game theoretic model in the United Kingdom, Russia, and Switzerland. *European Sociological Review* 29: 1010–1023.

Rasmusen E (2007) *Games and Information: An Introduction to Game Theory*. 4th ed. Oxford: Blackwell.

Raub W (1993) The role of rights in a theory of social action. *Journal of Institutional and Theoretical Economics* 149: 240–251.

Raub W (2004) Hostage posting as a mechanism of trust: binding, compensation, and signaling. *Rationality and Society* 16: 319–365.

Raub W, Buskens V and Corten R (2014) Social dilemmas and cooperation. In: Braun N and Saam NJ (eds) *Handbuch Modellbildung und Simulation in den Sozialwissenschaften*. Wiesbaden: Springer VS, pp. 597–626.

Raub W, Buskens V and Frey V (2013) The rationality of social structure. *Social Networks* 35: 720–732.

Schelling TC (1960) *The Strategy of Conflict*. London: Oxford University Press.

Stigler GJ (1964) A theory of oligopoly. *Journal of Political Economy* 72: 44–61.

Vega-Redondo F (2007) *Complex Social Networks*. Cambridge: Cambridge University Press.

Williamson OE (1985) *The Economic Institutions of Capitalism*. New York: Free Press.