Economics Letters xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/econbase

Sticks and carrots: Two incentive mechanisms supporting intra-group cooperation ☆

Thorsten Janus, Jamus Jerome Lim *

University of Wyoming and the World Bank, Mailstop H10-1004, 1818 H St NW, Washington, DC 20433, United States

ARTICLE INFO

Article history:
Received 21 April 2008
Received in revised form 16 December 2008
Accepted 19 December 2008
Available online xxxx

Keywords:
Intra-group cooperation
Prisoner's dilemma

JEL classification:
D71
D74

ABSTRACT

In this note, we introduce two distinct incentive mechanisms that support dynamic intra-group cooperation in the context of prisoner's dilemma payoffs: rewards for cooperating, and punishments for defection, where the rewarding or punishing party may be outside the relationship.

1. Introduction

Olson (1965) was among the first to formally pose the puzzle of group formation and cooperation, and this has provoked a large literature seeking to understand group behavior. In this note, we introduce two incentive mechanisms to sustain intra-group cooperation with prisoner's dilemma payoffs. For simplicity, we examine three-agent groups where relations may either be triadic–one person interacting with two others–or tripartite, where all agents interact. Due to shirking incentives, sustained group cooperation requires a system of endogenous enforcement. We posit a reward structure based on a transfer system and a punishment structure based on collective punishment. Both can ensure cooperation.

There is an existing literature that seeks to model institutions and social networks in terms of endogenous enforcement. The use of incentive slackness in triadic relations to tie strategies across two party games or "domains," has been studied by Aoki (2001) and Bernheim and Whinston (1990) while exogenous superior information or enforcement capability among group members compared to non-group members is used in Fearon and Laitin (1996) and Ghatak and Guinnane (1999). Moreover, such an institutional arrangement may itself be endogenous (Okada, 1993). The model here takes the former approach, but we add to the literature in two ways. First, we use transfers to endogenize the amount of incentive slack. Second, our punishment mechanism extends the use of incentive slack to three-party settings.

Tripartite cooperation has also been examined in Hart and Kurz (1983) and Ray and Vohra (1999). While both papers seek to endogenize the process of coalition formation, the latter goes a step further in endogenizing the coalition structure. Our note differs from these papers in that we do not assume that binding contracts can be written. An application in this vein, where contracts are assumed to bind, is the trade bloc model of Burbidge et al. (1997).

2. Analytical framework

The environment is comprised of a group with three agents, which are represented by the set $A=\{1, 2, 3\}$. Each individual i possesses a strategy set S_i, with $S'=S_1\times S_2\times S_3$. Let $q_i\in S_i$ be a feasible action for player i in the stage game, and denote $q=\{q_i, q_j, q_k\}$. Choices are perfectly observable, and players have perfect recall. A pure strategy for a player i is thus a sequence $\{s_{i,t}(\cdot)\}_{t=1}^{\infty}$ mapping the history H_{t-1} of previous action choices to the action choice in period t, $s_{i,t}(H_{t-1})\in S_i$; with the set of all such pure strategies given by Σ_i. These pure strategy profiles induce an outcome path $Q(s)=\{q_t\}_{t=1}^{\infty}=\{q_{1,t}, q_{2,t}, q_{3,t}\}_{t=1}^{\infty}$.

These agents interact in bilateral relationships over an infinite horizon (or, alternatively, over a finite horizon with no known termination time). These relationships are summarized in Fig. 1.

Agents have individual welfare given by

$$V_{i,0}=\sum_{t=0}^{\infty}(\delta_i)^t v_{ij,t}(q_{i,t}, q_{j,t}, q_{k,t}; \beta_{i,jk,t}), \tag{1}$$

☆ The authors thank Donald Wittman for helpful comments and suggestions. Financial support from the Graduate Division, University of California, Santa Cruz (Janus) and the Institute on Global Conflict and Cooperation (Lim) are gratefully acknowledged. The findings, interpretations and conclusions expressed herein are those of the authors and do not necessarily reflect the view of the World Bank Group, its Board of Directors, or the governments they represent.

* Corresponding author. Tel.: +1 202 458 9120 (office); fax: +1 202 614 9120.
E-mail addresses: tjanus@uwyo.edu (T. Janus), jlim@worldbank.org (J.J. Lim).

0165-1765/$ – see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.econlet.2008.12.012

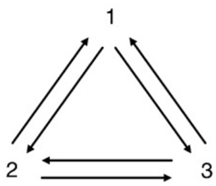

Fig. 1. Agent relationships.

where $V_{i,0}$ is the date 0 payoff to agent i given the agent's own action $q_{i,t}$ and those of the other players $q_{j,t}$ and $q_{k,t}$, which result from the strategies $s_{i,t}$, $s_{j,t}$, and $s_{k,t}$ employed at time t, respectively; $0<\delta_i<1$ is i's subjective discount rate; and $\beta_{i,jk,t}$ is a measure of the external benefit accruing to i for a relationship between individuals j and k at time t.

In the prisoner's dilemma, let $s_{i,t}=\{c,d;c,d;\tau_{ij};\tau_{ik}\}$, where $\tau_{ij}(\tau_{ik})\geq 0$ is a transfer from i to j (i to k). $v_{ij}(d,c)>v_{ij}(c,c)>v_{ij}left(d,d>v_{ij}(c,d)$ and similarly for the state game between i and k. Therefore, strategy d (c) is defection (cooperation) and the payoff structure does not depend on time.

Assumption 1. *Nash reversion strategy* Agents employ the *Nash reversion strategy* in the group formation game for the infinite horizon. In other words, a player i plays a strategy s_i along the equilibrium path $Q(s)$ until one of the three players defects, and the Nash equilibrium of the stage game $q_{i,t}=\{d;d;\tau_{ij}=0;\ \tau_{ik}=0\}$ is played thereafter.

This trigger strategy is fairly standard in the literature, but here we assume that if even if just one player deviates, all players–even those who have not experienced deviation directly in the prisoner's dilemma–revert to the static Nash equilibrium. In addition, we have included the cessation of transfers as well as the corresponding flows of external benefits.

There is experimental evidence that Nash reversion strategies are employed in repeated game settings, including the prisoner's dilemma, for two person games at least (Engel-Warnick and Slonim, 2004; Selten and Stoecker, 1986). Nash reversion also allows us to show equilibrium existence for the strongest threat possible; if a cooperative equilibrium does exist, one can always look for lighter punishments to support it. We will say that group formation is sustainable if all agents cooperate pairwise in their prisoner's dilemmas.

Definition 1. A subgame perfect pure strategy Nash equilibrium in the three-player game is a triple $\{q_1^*, q_2^*, q_3^*\}$ induced by a profile $s=(s_1^*, s_2^*, s_3^*)\in\Sigma_1\times\Sigma_2\times\Sigma_3$ such that for every h_{t-1} the restriction $s|h_{t-1}$ to the subgame starting at t satisfies: $\forall i,\ j\in A,\ i\neq j$: $\{s_i'\neq s_i^*$ such that $V_i(s_i|h_{t-1},\ s_j^*|h_{t-1},\ s_k^*|h_{t-1})>V_i^*(s_i^*|h_{t-1},\ s_j^*|h_{t-1},\ s_k^*|h_{t-1})\}$.

We now introduce the central proposition of this note.

Proposition 1 (Group formation). *For $0<\delta_1$, δ_j, $\delta_k<1$ and $A=\{1, 2, 3\}$, there exist Nash reversion strategies involving rewards and punishments such that group formation is sustainable even when one or more incentive compatibility constraints in the bilateral prisoner's dilemma are violated.*

Proof. See Appendix A. □

Remark. The result for group formation above is relatively comprehensive. That is, if the conditions in Proposition 1 are satisfied such that the equilibrium in Definition 1 exists, then incentive constraints can even be violated for *both* players in the same prisoner's dilemma.

The proposition thus allows us to assert that, with appropriate Nash reversion strategies, group formation in the context of a game with bilateral prisoner's dilemma payoffs can be sustained even if any two agents meeting alone would not cooperate. This extra cooperation is sustained either due to a transfer rewards system–carrots–or collective punishment threats—sticks.

For example, if agent 3 enjoys an external benefit from cooperation between 1 and 2, he or she can make 1 internalize this by paying 1 a transfer in every period as long as 1 cooperates with 2. If cooperation ceases, so does the transfer. This is essentially a repeated contingent use of the Coase theorem. Alternatively, though 2 and 3 may both be tempted to defect on each other, the threat that–if that were to occur–1 will defect on both of them in a collective punishment (and hence remove their surplus from their interactions with 1), can remove the temptation.

3. Applications

We offer two applications: Rewards in criminal gangs, and discipline within military platoons.

Given the high risks involved in gang participation, the finding that street-level members of gangs often earn more or less the minimum wage presents a puzzle (Levitt and Venkatesh, 2000). Even after taking into account the discounted value of potential future earnings–given the tournament structure of gang participation–membership in a gang appears to be suboptimal, especially given the likelihood that criminal discount rates are relatively high (Paternoster and Brame, 1998).

However, the enforcement of the cooperative outcome in gangs appears to be bolstered by rewards offered for cooperative behavior. These rewards may be monetary (a larger share of crime-related profits), nonpecuniary (enhanced recognition from promotion up the gang's hierarchy, or a greater sense of belonging), or both. In addition, these rewards usually accompany circumstances where group cohesion may be threatened. For example, Klein (1995) discusses how members of street gangs enjoy enhanced cohesion after engaging in sporadic criminal or violent activity, while Decker and van Winkle (1996) document how proceeds from the sale of drugs are considered rewards that are a part of street culture. These can be viewed as *ex post* transfers to members, offered by the gang leader, in exchange for the external benefits–usually viewed as leaders' prestige–that he or she obtains as a result of the gang's continued existence.

Punishment in the military is often meted out to the entire group. Ostensibly, this is to build a sense of group "intermindedness," or more precisely, a sense of unity through social solidarity (Spindler, 1948). This is especially cogent in boot camp–where recruits are transitioning from civilian to military life–and in paramilitary organizations (Archer, 1999), where looser institutional structures make it harder to prevent desertion.

Such punishment allows military organizations to support group cooperation, since punishments for deviant behavior in the platoon target the collective for the transgressions of any one soldier. For example, mass punishments remain common in boot camps even though the practice is expressly prohibited by military law (Heckathorn, 1988). In terms of our model, collective punishments of this form ensure that compliance rapidly emerges in the platoon.

Appendix A

Proof of Proposition 1. Before proceeding with the proof, it is useful to establish the following lemma.

Lemma 1 (Folk theorem). *$\exists\delta_i<1$ such that for any $\delta_i\in[\underline{\delta}, 1]$ $\forall i$, a cooperative outcome path $Q=\{q_{i,t}, q_{k,t}\}_{t=1}^{\infty}=\{c,c\}_{t=1}^{\infty}$ can be sustained with the Nash reversion strategy, where $\underline{\delta}$ is the lowest discount factor that can sustain cooperation, given the payoffs V_{ij}.*

Proof. The lemma is a direct result of the Folk Theorem for repeated games applied to the context here; see Fudenberg and Maskin (1986) for a formal treatment. □

Please cite this article as: Janus, T., Lim, J.J., Sticks and carrots: Two incentive mechanisms supporting intra-group cooperation, Economics Letters (2009), doi:10.1016/j.econlet.2008.12.012

Lemma 2 shows the proposition when we allow for both transfers and spillovers from relationships. Lemma 3 shows the special case when there are no spillovers and no transfers, and only the stick mechanism–Nash reversion by all three players following a single defection–sustains cooperation.

Lemma 2. *In the group formation game, there exist Nash reversion strategies with transfers $\tau_{ij}>0$, $i,j=\{1, 2, 3\}$, $i\neq j$ such that group formation is sustainable, even if one or more incentive compatibility constraints are violated inpairwise games in the absence of transfers, provided $\delta_i>0$ and $\beta_{i,jk}>0$ are sufficiently large or $\delta_i>0$ and the incentive slack of one of the players in a different relationship are sufficiently large.*

Proof. Even with the Nash reversion strategies, not all agents may be willing to play the prisoner's dilemma with both the other agents. However, we can develop the proof for the full cooperation case and then specialize it. Using $\tau_{ij}=-\tau_{ji}$, a nonempty set $\{\tau_{12},\tau_{13},\tau_{23}\}$ that satisfies the following conditions are sufficient for the Nash equilibrium:

$$v_{12}(d,c)-v_{12}(c,c)+v_{13}(d,c)-v_{13}(c,c)<\frac{\delta_1}{1-\delta_1}\times\{[v_{12}(c,c)-v_{12}(d,d)]+[v_{13}(c,c)-v_{13}(d,d)]+(\beta_{1,23}-\tau_{12}-\tau_{13})\},\tag{A.1a}$$

$$v_{21}(d,c)-v_{21}(c,c)+v_{23}(d,c)-v_{23}(c,c)<\frac{\delta_2}{1-\delta_2}\times\{[v_{21}(c,c)-v_{21}(d,d)]+[v_{23}(c,c)-v_{23}(d,d)]+(\beta_{2,13}+\tau_{12}-\tau_{23})\},\tag{A.1b}$$

$$v_{31}(d,c)-v_{31}(c,c)+v_{32}(d,c)-v_{32}(c,c)<\frac{\delta_3}{1-\delta_3}\times\{[v_{31}(c,c)-v_{31}(d,d)]+[v_{32}(c,c)-v_{32}(d,d)]+(\beta_{3,12}+\tau_{13}+\tau_{23})\}.\tag{A.1c}$$

Without loss of generality, assume that $\beta_{1,23}$ is large and that Eq. (A.1a) is satisfied, but Eqs. (A.1b) and (A.1c) are not. Then there potentially exist values of the transfers τ_{12} and τ_{13} such that the direction of the inequalities in Eqs. (A.1b) and (A.1c) reverse, while the inequality in Eq. (A.1a) remains unchanged. This requires $\hat{\tau}_{12}+\hat{\tau}_{13}<\beta_{1,23}$, where $\hat{\tau}_{ij}$ denotes the minimum required transfer to sustain cooperation between *i* and *j*. The bound on 1's willingness to subsidize the game between 2 and 3 depends on the external benefit, $\beta_{1,23}$, he or she enjoys, given cooperation between 2 and 3. Notice that if Eq. (A.1b) (if Eq. (A.1c)) is also satisfied, then the lower bound of $\beta_{1,23}$ can be lower, since $\hat{\tau}_{23}$ may be positive (negative).

However, not all players may be willing to play the prisoner's dilemma with each other. For example, if 1 and 3 do not play the prisoner's dilemma, then Definition 1 can still be satisfied with $\beta_{2,13}=0$ so that Eq. (A.1a) simplifies to

$$v_{12}(d,c)-v_{12}(c,c)<\frac{\delta_1}{1-\delta_1}\{[v_{12}(c,c)-v_{12}(d,d)]+\beta_{1,23}-\tau_{12}-\tau_{13}\},\tag{A.2a}$$

$$v_{21}(d,c)-v_{21}(c,c)+v_{23}(d,c)-v_{23}(c,c)<\frac{\delta_2}{1-\delta_2}\{[v_{21}(c,c)-v_{21}(d,d)]+[v_{23}(c,c)-v_{23}(d,d)]+\tau_{12}-\tau_{23}\},\tag{A.2b}$$

$$v_{32}(d,c)-v_{32}(c,c)<\frac{\delta_3}{1-\delta_3}\{[v_{32}(c,c)-v_{32}(d,d)]+\beta_{3,21}+\tau_{13}+\tau_{23}\}.\tag{A.2c}$$

Again without loss of generality, for $\beta_{1,23}$ large, Eq. (A.2a) is satisfied and transfers can be used to satisfy the other equations, again up the value of the external benefit 1 enjoys. If, furthermore, 1 and 2 do not play the prisoner's dilemma so that $v_{12}(\cdot)=v_{21}(\cdot)=0$, Definition 1 simply requires

$$\beta_{1,23}>\tau_{12}+\tau_{13},\tag{A.3a}$$

$$v_{23}(d,c)-v_{23}(c,c)<\frac{\delta_2}{1-\delta_2}\{[v_{23}(c,c)-v_{23}(d,d)]+(\tau_{12}-\tau_{23})\},\tag{A.3b}$$

$$v_{32}(d,c)-v_{32}(c,c)<\frac{\delta_3}{1-\delta_3}\{[v_{32}(c,c)-v_{32}(d,d)]+(\tau_{13}+\tau_{23})\},\tag{A.3c}$$

which again is satisfied for sufficiently large $\beta_{1,23}$. The cooperation conditions for all other configurations of agents playing the prisoner's dilemma are analogous. □

Lemma 3. *For $0<\delta_i<1$, $i=\{1, 2, 3\}$, there exist Nash reversion strategies with punishments such that group formation is sustainable, even if one or more incentive compatibility constraints are violated, provided every player has incentive slack in at least one relationship and that $\delta_i>0$ is sufficiently large.*

Proof. To remove the possibility of the transfer mechanism sustaining cooperation, assume that $\beta_{i,jk}=\tau_{ij}=0\,\forall i,j,k$. If both of an agent's incentive compatibility constraints are violated, by Lemma 1, the agent will not cooperate. Therefore, we consider only cases with at most one of each agent's incentive constraints violated. In this case, group formation could still be possible as the Nash equilibrium. Without loss of generality, let $v_{12}(d,c)-v_{12}(c,c)>\frac{\delta_1}{1-\delta_1}[v_{12}(c,c)-v_{12}(d,d)]$ and $v_{13}(d,c)-v_{13}(c,c)<\frac{\delta_1}{1-\delta_1}[v_{13}(c,c)-v_{13}(d,d)]$. Given Nash reversion group formation requires

$$v_{12}(d,c)-v_{12}(c,c)+v_{13}(d,c)-v_{13}(c,c)<\frac{\delta_1}{1-\delta_1}\{[v_{12}(c,c)-v_{12}(d,d)]+[v_{13}(c,c)-v_{13}(d,d)]\},\tag{A.4a}$$

$$v_{21}(d,c)-v_{21}(c,c)+v_{23}(d,c)-v_{23}(c,c)<\frac{\delta_2}{1-\delta_2}\{[v_{21}(c,c)-v_{21}(d,d)]+[v_{23}(c,c)-v_{23}(d,d)]\},\tag{A.4b}$$

$$v_{31}(d,c)-v_{31}(c,c)+v_{32}(d,c)-v_{32}(c,c)<\frac{\delta_3}{1-\delta_3}\{[v_{31}(c,c)-v_{31}(d,d)]+[v_{32}(c,c)-v_{32}(d,d)]\},\tag{A.4c}$$

and Eq. (A.4a) can still be satisfied for $v_{13}(c,c)$ sufficiently large. If both Eqs. (A.4b) and (A.4c) are satisfied in an analogous manner, then group formation is possible. □

Taken together, Lemma 2 and 3 exhaust all possible cases, which concludes the proof of the proposition. □

References

Aoki, Masahiko, 2001. Toward a Comparative Institutional Analysis. MIT Press, Cambridge, MA.

Archer, David, 1999. Exploring "bullying" culture in the para-military organisation. International Journal of Manpower 20 (1/2), 94–105 (February).

Bernheim, B. Douglas, Whinston, Michael D., 1990. Multimarket contact and collusive behavior. RAND Journal of Economics 21 (1), 1–26 (Spring).

Burbidge, John B., DePater, James A., Myers, Gordon M., Sengupta, Abhijit, 1997. A coalition-formation approach to equilibrium federations and trading blocs. American Economic Review 87 (5), 940–956 (December).

Decker, Scott H., van Winkle, Barrick, 1996. Life in the gang: family, friends, and violence. Cambridge Studies in Criminology. Cambridge University Press, New York, NY.

Engle-Warnick, Jim, Slonim, Robert L., 2004. The evolution of strategies in a repeated trust game. Journal of Economic Behavior and Organization 55 (4), 553–573 (December).

Fearon, James D., Laitin, David D., 1996. Explaining interethnic cooperation. American Political Science Review 90 (4), 715–735 (December).

Fudenberg, Drew, Maskin, Eric S., 1986. The Folk Theorem in repeated games with discounting or with incomplete information. Econometrica 54 (3), 533–554 (May).

Ghatak, Maitreesh, Guinnane, Timothy W., 1999. The economics of lending with joint liability: theory and practice. Journal of Development Economics 60 (1), 195–228 (October).

Hart, Sergiu, Kurz, Mordecai, 1983. Endogenous formation of coalitions. Econometrica 51 (4), 1047–1064 (July).

Heckathorn, Douglas D., 1988. Collective sanctions and the creation of prisoner's dilemma norms. American Journal of Sociology 94 (3), 535–562 (November).

Klein, Malcolm W., 1995. The American Street Gang: Its Nature, Prevalence, and Control. Oxford University Press, New York, NY.

Levitt, Steven D., Venkatesh, Sudhir Alladi, 2000. An economic analysis of a drug-selling gang's finances. Quarterly Journal of Economics 115 (3), 755–789 (August).

Okada, Akira, 1993. The possibility of cooperation in an n-person prisoners' dilemma with institutional arrangements. Public Choice 77 (3), 629–656 (November).

Olson, Mancur, 1965. The logic of collective action: public goods and the theory of groups. Harvard Economic Studies. Harvard University Press, Cambridge, MA. revised edition.

Paternoster, Raymond, Brame, Robert, 1998. The structural similarity of processes generating criminal and analogous behaviors. Criminology 36 (3), 633–670 (August).

Ray, Debraj, Vohra, Rajiv, 1999. A theory of endogenous coalition structures. Games and Economic Behavior 26 (2), 286–336 (January).

Selten, Reinhard, Stoecker, Rolf, 1986. End behavior in sequences of finite prisoner's dilemma supergames. Journal of Economic Behavior and Organization 7 (1), 47–70 (March).

Spindler, G. Dearborn, 1948. The military—a systematic analysis. Social Forces 27 (1), 83–88 (October 1948–May 1949).

Please cite this article as: Janus, T., Lim, J.J., Sticks and carrots: Two incentive mechanisms supporting intra-group cooperation, Economics Letters (2009), doi:10.1016/j.econlet.2008.12.012