Journal of Economic Behavior and Organization 158 (2019) 288–310

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

Partners or rivals? An experimental study of a two-stage tournament

Hong Chao[a,*], Chun-Yu Ho[b], Shaoqing Huang[c], Xiangdong Qin[c], Jiajia Cong[d]

[a] *Department of Finance, School of Economics, Faculty of Economics and Management, East China Normal University. 3663 North Zhongshan Rd, Shanghai 200062, China*
[b] *Department of Economics, University at Albany, State University of New York, New York, NY 12222, USA*
[c] *Antai College of Economics & Management, Shanghai Jiao Tong University. 1954 Huashan Rd, Shanghai 200030, China*
[d] *Department of Industrial Economics, School of Management, Fudan University. 670 Guoshun Rd, Shanghai 200433, China*

ARTICLE INFO

Article history:
Received 11 January 2018
Revised 9 October 2018
Accepted 27 November 2018
Available online 6 December 2018

JEL classification:
C92
D74
D82
D83

Keywords:
Two-stage tournament
Alliance
Hold-up problem
Appropriation
Informal agreement

ABSTRACT

We consider a two-stage tournament in which two alliances compete in stage one. Members of the winning alliance then compete against each other in stage two. Members' investment in stage one could increase their alliance's winning probability, but could also be appropriated and used against them by their partners-turned-rivals in stage two. This hold-up problem creates a negative incentive for within-alliance cooperation. We test this theoretical benchmark in an experiment. In a second experiment, we investigate whether *ex-ante* informal agreements deter appropriation. We find that such agreements are honored, and thus encourage investments, only under the fixed matching condition.

1. Introduction

Competition between teams is a widespread phenomenon. Team members must cooperate to win between-team competitions, but they may also free ride on their partners' contributions (Olson and Zeckhauser, 1966). Apart from the free-riding problem, members may come into a direct conflict outside the scope of their team, making them cautious in cooperating inside the team. Potential competition between team members is particularly detrimental to cooperation if members are able to use the resources that partners have invested in the team to compete against them outside the team. Although studies have reported that between-team competition and peer rewards or sanctions can mitigate the within-team free-riding problem (Erev et al., 1993; Fehr and Gächter, 2000; Eriksson and Villeval, 2012; Pan and Houser, 2017), cooperation between team members is still difficult to establish when such a hold-up problem is present.

This hold-up problem is often observed in political alliances. Members of an alliance may use resources invested by partners and exploit partners' exposed weaknesses to compete against them outside the scope of the alliance. For exam-

* Corresponding author.
E-mail addresses: hchao@fem.ecnu.edu.cn, chaohong@sjtu.edu.cn (H. Chao).

https://doi.org/10.1016/j.jebo.2018.11.023

ple, to strengthen the Alliance against Nazis in World War II, the United States aided the Soviet Union with about $12.5 billion in war materials and other supplies from 1941 to 1945 via the Lend-Lease program. In particular, Soviets received 15,000 airplanes, 9000 tanks and self-propelled guns, 362,000 trucks, 47,000 jeeps, and other non-military supplies including foodstuffs, telegraph wire and cable, metals, chemical substances, petroleum products, and production-line equipment (Weeks, 2010). The aid received by the Soviets was often misused or even transferred or resold to its counterparts in order to strengthen its political influence in Eastern Europe and the Middle East. For example, sugar and tires were sold to Iranians; trucks were sold or donated to Poland. In response to the mounting concerns about the misuse of aided resources in the postwar competition, the United States limited the amount and restricted the conditions of the aid to the Soviets, especially near the end of World War II (Herring, 1969; Minkova, 2018). These worries subsequently manifested themselves. During the Korean War (1950–1953), American troops captured much American equipment. Evidently, the Soviets provided military aid to Korea using the very same supplies they had received from the United States several years earlier (Weeks, 2010).

Another example is the Peloponnesian war (431–404 BCE) between the Delian League led by Athens and the Peloponnesian League led by Sparta. In the final phase of the war, Athens fought the coalition of Sparta and Persia. However, being afraid that Sparta would become a powerful rival in the future, Persia was slow to furnish promised funds and ships. For example, Persia intentionally skimped a portion of Spartan soldiers' pay and delayed the aid of Phoenician fleet (Meiggs, 1979; Kagan, 1991). Persia's worry became true after the victory of the Peloponnesian League. Sparta invaded Persia with its powerful fleet and with the support of tax revenue collected from Athens' city-states, both of which were the legacy of Persian's aid during the Peloponnesian war (Buckler, 2003).[1]

We build a two-stage tournament model for the aforementioned situation. In the first stage, two alliances with two members in each alliance compete. The output of each alliance depends on its members' investments and a random shock. The alliance with higher output wins the tournament, and the first-stage prize is equally shared between its members. In the second stage, the members of the winning alliance can take advantage of the resource their former partner invested in the first stage to increase their own output. Their individual output depends on the appropriated resource and a random shock. The member with the higher output wins the second-stage prize. Importantly, alliance members anticipate the second-stage appropriation behavior of their partners-turned-rivals. This makes them reluctant to invest in their alliance in the first stage, as investing more resources could jeopardize their success in the second stage. This reluctance to invest results in a diminution in the probability of winning the first-stage tournament; hence, members must consider this tradeoff while making investment decisions. This feature distinguishes our model from the multi-stage tournaments studied in the literature (e.g., Altmann et al., 2012; Ludwig and Lünser, 2012).

Engel and Kleine (2015) study a two-stage game in which two players make innovation investments in the first stage, with one's likelihood of success depending only on one's own investment. If one player's innovation succeeds, the counterpart can appropriate the innovation to increase profit in the second stage. Surprisingly, they find that the second-stage appropriation of the counterpart does not suppress one's first-stage innovation investment, even if such appropriation reduces the innovator's profit. The setting they consider involves neither cooperation nor competition. In contrast, we study how the first-stage within-alliance cooperation is affected if there is appropriation between ex-partners in the second-stage competition. Ke et al. (2013) use a two-stage contest model to study how the experience of fighting a common rival in the first stage influences alliance members' willingness to turn against each other in the second stage. They find that the anticipation of future conflict decreases partners' contributions to the fight against common enemies. However, they do not consider the role of appropriation behavior, which is the focus of our study.

Our model predicts that when a member anticipates its partner to choose a high (low) appropriation level in the second stage, the member will choose a low (high) investment level in the first stage. To test this, we design a laboratory experiment that allows us to isolate the effect of the appropriation level in a controlled environment. In three treatments of the experiment, the second-stage appropriation levels are exogenously set to zero, low, and high, respectively. We find that the subjects' investments under the zero and low appropriation conditions are significantly higher than those under the high appropriation condition, in line with our prediction.

In a second step, we study the setting in which subjects can endogenously choose the appropriation level. Theoretically, subjects will always choose the highest possible level of appropriation. A third-party enforceable punishment, while certainly effective in deterring appropriation and thus encouraging investment, is often absent in practice. Hence, we are interested in whether an *ex-ante* informal agreement to deter future appropriation can effectively induce investment.

Studies have found that if non-binding communication is possible, subjects may have an incentive to agree on coordinating strategies that maximize their joint payoffs (Crawford, 1998). Even if the strategies specified in the agreement do not maximize their individual payoff (e.g., in social dilemma games), subjects may still honor the agreement if they are afraid of triggering future noncooperation from other members. Some studies have reported that informal agreements effectively mitigate the free-riding problem (Sutter and Strassmair, 2009; Cason et al., 2012), while other studies have found that an informal agreement between individuals with conflicting interests is more likely to be broken (Seidmann, 1990; Charness, 2000; Ke et al., 2015). In our setting, aside from the free-riding problem, a direct conflict of interest arises between the two

[1] Although we study the hold-up problem within the framework of political alliances, this problem also exists in many other cooperative situations. For example, at the individual level, researchers may cooperate in a joint project by sharing resources but use each other's resources in their projects and compete outside the joint project (e.g., in promotion competition). At the firm level, members of a research and development (R&D) alliance may copy their partners' invested knowledge and skills and use them to produce competing products outside their alliance (Katz and Ordover, 1990).

former partners in the second stage, as a decision to refrain from appropriating increases one's likelihood to lose the second stage. Thus, the effect of an informal agreement in this setting is unclear *ex ante*.

To investigate the effect of an informal agreement, we conduct a second experiment with a 2×2 factorial design for four treatments; we vary the communication possibility (possible vs. impossible) and the matching method (random vs. fixed). In contrast to the first experiment, the subjects in these four treatments are free to choose between a low and high appropriation level. In the two treatments with communication opportunities, the subjects can communicate about their intended appropriation choices before the game. We find that when the communication opportunity is absent, the subjects almost always choose the high appropriation level under both random and fixed matching methods. When communication is possible, the subjects under random matching rarely adhere to their agreements to choose the low appropriation level, making it difficult to establish within-alliance cooperation. Under fixed matching, however, the agreements are frequently honored due to reputation concerns, and this encourages the subjects to develop long-term within-alliance cooperative strategies; thus, largely increasing their investments.[2]

Our study is different from the previous literature on the hold-up problem in which one trading party is unwilling to make the *ex-ante* relationship-specific investments since the *ex-post* returns on investment are to be shared with the other party who does not share in the investment cost (Klein et al., 1978; Williamson, 1985).[3] The novelty of our study lies in its focus on the hold-up problem generated by the appropriation behavior of a member's partners-turned-rivals in a two-stage tournament. We are also the first to examine how informal agreements mitigate this hold-up problem. Studies have found that communication about investments improves within-team cooperation (Sutter and Strassmair, 2009; Cason et al., 2012). In contrast, our subjects can only make known their intended appropriation choices in the pre-play structured communication. We are thus able to examine how the informal agreements against future appropriation affect the current investment choices.

The remainder of this paper is organized as follows. Section 2 outlines a simple model to deliver testable hypotheses. Section 3 examines the role of appropriation on investment. Section 4 examines the role of communication and matching protocol on increasing investment and decreasing appropriation level. Section 5 concludes.

2. A simple two-stage tournament model

2.1. The first-stage tournament

We follow the framework of a rank-order tournament (Lazear and Rosen, 1981) to build our model. We consider two alliances with two risk-neutral members (i.e., political parties in our context) in each alliance who act simultaneously. In the first stage, the two alliances compete against each other. The output of each alliance *i*, y_i, depends on the investments of its two members, *ik* and *im*, according to the following function:

$$y_i = f(a_{ik} + a_{im}) + \varepsilon_i, \quad i, k, m = 1, 2, \text{ and } k \neq m.$$

The variables a_{ik} $(0 \le a_{ik} \le \bar{a})$ and a_{im} $(0 \le a_{im} \le \bar{a})$ represent the investments of members *ik* and *im*, respectively, where $\bar{a}$ is the common upper bound of each member's endowed political resource. We assume that the investments are not verifiable. Although the invested resources of the two members can be heterogeneous, their marginal products on the output, *f*, are assumed to be the same. The random shock, ε_i, is uniformly distributed over the interval $[-\bar{\varepsilon}, +\bar{\varepsilon}]$ and is independent and identically distributed (*i.i.d.*) for both alliances. Investments are costly for the members in an alliance and assumed to take a linear function: $C(a_{ik}) = \theta a_{ik}$, $i,k = 1,2$.

If y_i is higher than the output of the other alliance, say y_j, then alliance *i* wins the first-stage tournament. The members in alliance *i* equally share the fixed prize *V*, as their investments are unverifiable, but the members in alliance *j* receive no reward. The probability of alliance *i* winning the first-stage tournament is $P_i = P(y_i > y_j) = F_{\varepsilon_j - \varepsilon_i}(f(a_{ik} + a_{im} - a_{jk} - a_{jm}))$, which is the cumulative distribution function (CDF) of ε_j-ε_i.

2.2. The second-stage tournament

In the second stage, the two partners in the winning alliance become rivals and compete against each other. To improve their individual output, members may appropriate and use the first-stage resource invested by their former partner. The

2 The finding that communication can promote cooperation in fixed matching but not random matching is consistent with that of Buckley et al. (2018) who study the effects of different communication channels on countervailing the free-riding incentives to solve the over-harvesting problem in common pool resources.

3 Previous studies have examined different mechanisms to mitigate the classic hold-up problem, including but not limited to reciprocity and fairness concern (Sonnemans et al., 2001), group identity (Morita and Servátka, 2013), communication (Ellingsen and Johannesson, 2004) and joint ownership (Fehr et al., 2008), etc.

individual output depends on their appropriated resource according to the following function:[4]

$$x_{ik} = g\beta_{ik}a_{im} + \mu_{ik}, \quad i, k, m = 1, 2, \text{ and } k \neq m,$$

where $\beta_{ik} \in [0, \bar{\beta}]$ is the appropriation level selected by member *ik*. We assume, although the appropriation is costless that there is an exogenous upper bound of appropriation level $\bar{\beta} \in [0, 1]$. $\beta_{ik}a_{im}$ is the amount of resource appropriated by member *ik* from its former partner, member *im*, and *g* is the marginal product of the appropriated resource. The random shock μ_{ik} is uniformly distributed over the interval $[-\bar{\mu}, +\bar{\mu}]$ and *i.i.d.* for both members.

If the individual output of member *ik* is higher than that of member *im*, member *ik* wins the second-stage prize *W*, and member *im* receives no second-stage reward. The probability of member *ik* winning the second-stage tournament is given by $P_{ik} = P(x_{ik} > x_{im}) = F_{\mu_{im}-\mu_{ik}}(g\beta_{ik}a_{im} - g\beta_{im}a_{ik})$, which is the CDF of μ_{im}-μ_{ik}.

2.3. Symmetric equilibrium of the two-stage model

Given the output and cost functions, the expected total payoff of each member *ik* is

$$E\pi_{ik} = P_i\left(\frac{V}{2} + P_{ik}W\right) - \theta a_{ik}, \; i, k = 1, 2. \tag{1}$$

As a benchmark, we solve this two-stage game by backward induction and derive the symmetric subgame perfect Nash equilibrium predictions.

In the second stage, after observing its former partner's first-stage investment, a_{im}, member *ik* chooses an appropriation level β_{ik} to maximize its second-stage expected payoff: $E\pi_{ik}^{2nd-stage} = P_{ik}W, \; i, k = 1, 2$. The first-order condition (FOC) of $E\pi_{ik}^{2nd-stage}$with respect to β_{ik} is

$$\frac{\partial E\pi_{ik}^{2nd-stage}}{\partial \beta_{ik}} = \frac{\partial P_{ik}}{\partial \beta_{ik}}W = \frac{ga^*W}{2\bar{\mu}},$$

where a^* is each member's first-stage investment level in the symmetric equilibrium. As there is no cost of appropriation, a payoff-maximizing member chooses the highest possible level of appropriation for any positive first-stage investment. The symmetric equilibrium thus entails $\beta^* = \bar{\beta}$.

In the first stage, member *ik* chooses a_{ik} to maximize its expected total payoff $E\pi_{ik}$ given in Eq. (1). We thus obtain the following FOC:

$$\frac{\partial E\pi_{ik}}{\partial a_{ik}} = \frac{\partial P_i}{\partial a_{ik}} \cdot \frac{V+W}{2} + \frac{1}{2} \cdot \frac{\partial P_{ik}}{\partial a_{ik}} \cdot W - \theta. \tag{2}$$

The first term in Eq. (2) represents the marginal revenue of investment in which $\frac{\partial P_i}{\partial a_{ik}} = \frac{f}{2\bar{\varepsilon}}$ is the marginal probability of winning the first-stage tournament for the investment. The second term represents the negative cross-stage effect of investment in which $\frac{\partial P_{ik}}{\partial a_{ik}} = -\frac{g\beta^*}{2\bar{\mu}}$ is the marginal probability of winning the second-stage tournament for the investment. It is negative, as investing more in the first stage increases the amount of resource being appropriated by a former partner and hence the probability of being defeated in the second stage. The last term θ measures the marginal cost of investment.

3. Appropriation and investment

3.1. Hypotheses development

Given our parametric specification, Eq. (2) can be rewritten as follow:

$$\frac{\partial E\pi_{ik}}{\partial a_{ik}} = \frac{f}{2\bar{\varepsilon}} \cdot \frac{V+W}{2} - \frac{1}{2} \cdot \frac{g\beta^*}{2\bar{\mu}} \cdot W - \theta. \tag{3}$$

Eq. (3) shows that a larger β^*, ceteris paribus, implies a stronger negative cross-stage effect of investment. If β^* is sufficiently large, so that $\beta^* > T$, i.e., $\bar{\beta} > T$ as $\beta^* = \bar{\beta}$, where $T = \bar{\mu}(f(V+W) - 4\theta\bar{\varepsilon})/gW\bar{\varepsilon}$, then Eq. (3) is negative, implying that members should choose the lowest investment level, $a^* = 0$. If $\beta^* < T$, *i.e.*, $\bar{\beta} < T$, Eq. (3) is positive, implying that $a^* = \bar{a}$.[5] This theoretical prediction generates the following hypothesis.

[4] In an alternative model, we consider the case where members can invest in the second stage, and their individual output is an additive function of appropriated resources, second-stage investments, and a random shock. In this case, the equilibrium second-stage investment is independent of both the appropriation level and the first-stage investment. Thus, incorporating the second-stage investment does not change our prediction about the relationship between the second-stage appropriation and the first-stage investment. To focus on our prediction and to make it easier to implement the game in the lab, we assume away the second-stage investments.

[5] Assuming a convex cost function, e.g., $\theta a^2/2$, yields an inner solution $a^* = (\frac{f}{2\bar{\varepsilon}} \cdot \frac{V+W}{2} - \frac{1}{2} \cdot \frac{g\beta^*}{2\bar{\mu}} \cdot W)/\theta$, a negative relationship between β^* and a^* is also predicted. Since the subjects in our experiment have to consider various strategic and production uncertainties simultaneously while making investment choices, a convex cost function would further exaggerate the complexity of the experimental setting. We thus use a linear cost function.

Hypothesis 1. *A subject's first-stage investment level is negatively correlated with his/her partner's second-stage appropriation level.*

Intuitively, if a member anticipates that its partner will choose a high appropriation level in the second stage, it will refrain from investing in the first stage. Conversely, a member should be willing to invest if it anticipates a low appropriation level from its partner.

3.2. Experimental design and procedure

To test Hypothesis 1, we conduct an experiment with three treatments, *Zero, Low,* and *High*, in which we exogenously set the appropriation level β to 0, 0.1, and 0.9, respectively. The baseline treatment *Zero* allows us to examine the subjects' behavior in the absence of an appropriation.[6] The subjects are informed of the value of the appropriation level at the beginning of the experiment. We use a random matching protocol (in particular, an absolute stranger design) in which the subjects are paired with a different alliance partner in each round. All alliances within a session are also randomly paired in each round to play the game.

The two subjects in each alliance independently choose their investment levels a from a set of integers {0, 1, …, 20} to compete against the paired alliance in stage one. The marginal products f and g are both set to 20. The first-stage prize V and the second-stage prize W are 2000 and 18,000. The random terms ε and μ are independently and uniformly distributed over the integer intervals [−40, 40] and [−20, 20].[7] The marginal cost of investment θ is 100. The subjects receive an endowment of 2000 in each round; thus, even if a subject invests 20 and loses the first-stage tournament, he/she receives zero at worst.[8] Given these parameters, the threshold for appropriation level T is equal to 0.53. Therefore, in the *Zero* and *Low* treatments, where $\beta = 0 < T$ and $\beta = 0.1 < T$, the equilibrium investments a are 20. In the *High* treatment, where $\beta = 0.9 > T$, the equilibrium investment a is 0.

The experiment, conducted in the Smith Experimental Economics Research Center at Shanghai Jiao Tong University in China, uses the z-Tree program (Fischbacher, 2007). We run four between-subject sessions for each treatment. The sample consists of 144 students from multiple disciplines divided into 12 sessions, each with 12 participants. The subjects play ten formal rounds in each session under a single treatment. Three trial rounds are played before the formal rounds, where there are no payments or interactions between the subjects, and the subjects are informed that they are playing a computer counterpart. In each formal round, after the subjects make investment choices, the computer randomly determines ε for each alliance and compares the first-stage outputs of the paired alliances. The computer then randomly determines μ for each subject in the winning alliances and compares the second-stage outputs of the paired subjects.

At the end of each round, the subjects are informed of their partner's first-stage choice, the random term of their alliance, the first-stage outputs of their alliance and the paired alliance, and whether they receive a first-stage prize and enter into the second stage. They are also informed of their random term, their second-stage outputs and those of their partner, whether they receive a second-stage prize, their current-round earnings, and their cumulative earnings. A history table with the information from previous rounds is provided on the computer screen during the session. An example of instructions can be found in Appendix A. We only use neutral words in the instructions to avoid any possible framing effect.

To ensure that the experiments are completed in a reasonable amount of time, the subjects are given 60 s to make each choice. If they do not make an investment decision within the given time, the computer randomly chooses an integral number for them from the set of {0, 1, …, 20}. In our dataset, 2 out of 1440 investment decisions are made by the computer; we drop these auto-selected observations from our analysis. The average earnings are RMB58.45, including a show-up fee of RMB5 (USD1 = RMB6.21 at the time of the experiments). Each session lasts 55 minutes on average.

Finally, the subjects complete a post-experimental survey. In addition to providing demographic data, the subjects report their math ability, risk tolerance, individualism, and belief of fairness. The variable *Math* is characterized by a number from 1 to 7, with a larger number indicating a higher math ability. The variable *Risk Tolerance* is characterized by a number from 1 to 10, with a larger number indicating a higher degree of risk tolerance. The *Individualism* of subjects is characterized by a number from 1 to 5, where 1 indicates they tried to earn the most for their group and 5 indicates they tried to earn the most for themselves. The subjects report their belief about *Fairness* by choosing a number from 1 to 10, where 1 indicates they believe that most people are selfish and 10 indicates they believe that most people are fair. In our empirical analysis, we standardized the mixed scales into one common scale. Table 1 reports the descriptive statistics of the subjects' characteristics.

---

[6] We thank one referee for suggesting running this baseline treatment.

[7] We set the value of W larger than that of V to mimic the situation in which the main proportion of prize is rewarded at the second stage. For example, in the elections held in Nepal and Catalonia in 2017, the parties in the first-stage winning alliances (the left alliance in Nepal and the pro-independence alliance in Catalonia) are then competing for leadership positions in government, including the position of president (Bhattarai, 2017; Burgen and Jones, 2017; Hatton and Parra, 2017). In addition, we set the interval of μ to be a half of that of ε since the former is the random shock of each subject's individual output while the latter is the random shock of each alliance's output which is the sum of its two members' outputs.

[8] Note that the expectation of negative earnings may generate undesirable side-effects (Friedman and Sunder, 1994). The endowment can be regarded as the prize for a party in the losing alliance, and the first-stage prize can be regarded as the prize gap between the winning and losing alliances. The theoretical predictions in the text remain unchanged with this adjustment (Lazear and Rosen, 1981).

Table 1
Descriptive statistics of subjects.

Variable	Unit	Definition	Zero, low & high	Communication × matching
Low	0/1	1 indicates the *Low* treatment, and 0 indicates the other treatments	0.33 (0.47)	
High	0/1	1 indicates the *High* treatment, and 0 indicates the other treatments	0.33 (0.47)	
Communication	0/1	1 indicates communication is possible and 0 otherwise		0.50 (0.50)
Fix	0/1	1 indicates the fixed matching protocol, and 0 indicates the random matching protocol.		0.50 (0.50)
Female	0/1	1 indicates female, and 0 indicates male	0.26 (0.44)	0.35 (0.48)
Math	[0,1]	A larger number indicates a greater math ability.	0.64 (0.22)	0.63 (0.23)
Risk tolerance	[0,1]	A larger number indicates a higher degree of risk tolerance.	0.53 (0.25)	0.52 (0.26)
Individualism	[0,1]	A larger number indicate the subject's greater incentive to earn the most only for himself/herself instead of for his/her alliance.	0.45 (0.35)	0.46 (0.34)
Fairness	[0,1]	A larger number indicates the subject believes that most people are fair while dividing interests instead of selfish.	0.34 (0.29)	0.29 (0.25)
Observation			1,438	1,904

Note: We report the mean of each variable followed by its standard deviation in brackets. We exclude any observations in which the investment levels were randomly chosen by the computer because the subjects did not make decisions within their time limit.
Column Zero, Low & High: The sample of *Zero, Low* and *High* treatments. *Zero* (*Low; High*) indicates the treatment of zero (low; high) appropriation level. There are 480, 480, and 478 observations in the *Zero, Low*, and *High* treatments, respectively.
Column Communication × Matching: The sample of *NC-Random, C-Random, NC-Fix*, and *C-Fix* treatments. *NC-Random* indicates the treatment with no communication and the random matching protocol. *C-Random* indicates the treatment with communication and the random matching protocol. *NC-Fix* indicates the treatment with no communication and the fixed matching protocol. *C-Fix* indicates the treatment with communication and the fixed matching protocol. There are 475, 476, 475, and 478 observations in the treatment of *NC-Random, C-Random, NC-Fix*, and *C-Fix*, respectively.

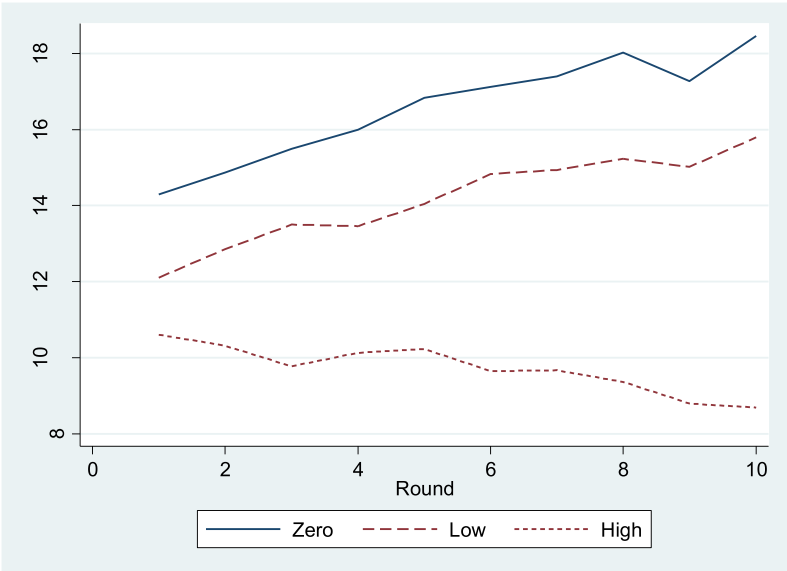

Fig. 1. Average Investments over Time for the Zero, Low & High Treatments.
Note: Zero, Low & *High* indicates the treatments of zero, low and high appropriation levels, respectively.

3.3. Results

Fig. 1 depicts the average individual investment in the *Zero, Low,* and *High* treatments over time and presents two interesting results. First, the average individual investment in the *Zero* treatment is higher than that in the *Low* treatment, which, in turn, is higher than that in the *High* treatment. Second, the average individual investment in the *Zero* and *Low* treatments increases over time, whereas that in the *High* treatment decreases over time. The second result suggests that there is a learning effect over time.

Panel A of Table 2 reports that the mean of individual investment in the *Zero* treatment is 16.58, that in the *Low* treatment is 14.18, and that in the *High* treatment is 9.72. The differences in means for individual investment between the *Zero* and *Low* treatments, between the *Zero* and *High* treatments, between the *Low* and *High* treatments are significantly different

Table 2
Equality tests of investment and low appropriation choice across treatments.

Treatment	Investment	N	Low appropriation choice	N
Panel A: zero, low & high				
Zero	16.58	4		
Low	14.18	4		
High	9.72	4		
Zero – Low	2.40** [0.95]			
Wilcoxon rank-sum test	< 0.05			
Zero – High	6.86*** [1.08]			
Wilcoxon rank-sum test	< 0.05			
Low – High	4.46*** [0.82]			
Wilcoxon rank-sum test	< 0.05			
Panel B: Communication x Matching				
NC-Random	11.03	4	0.06	4
C-Random	8.93	4	0.07	4
NC-Fix	12.99	4	0.12	4
C-Fix	17.42	4	0.36	4
NC-Random – C-Random	2.10 [1.16]		0.01 [0.04]	
Wilcoxon rank-sum test	> 0.10		> 0.10	
NC-Fix – NC-Random	1.96 [1.23]		0.06 [0.03]	
Wilcoxon rank-sum test	< 0.10		> 0.10	
C-Fix – NC-Fix	4.43*** [1.14]		0.24*** [0.03]	
Wilcoxon rank-sum test	< 0.05		< 0.05	

Note: Each observation for the non-parametric tests is the average of the interested variable of all subjects in a session over ten periods. Standard errors of two-sided equal-mean t-tests are in brackets. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. We report the p-values of two-sided Wilcoxon rank-sum tests.
Panel A: The sample of *Zero, Low*, and *High* treatments. Panel B: The sample of *NC-Random, C-Random, NC-Fix*, and *C-Fix* treatments. The definition of *Zero, Low, High, NC-Random, C-Random, NC-Fix*, and *C-Fix* are stated in the note of Table 1.

from each other. As all of the parameters are the same except for the exogenous appropriation levels in the three treatments, the difference in individual investments between these treatments suggests a negative effect of appropriation on investments. These results also suggest that the anticipation of future competition *per se* does not significantly weaken the current cooperation if the appropriation is impossible or mild; however, it strongly undermines cooperation if the within-alliance appropriation is sufficiently severe.

To control for the effects of the subjects' characteristics on individual investment, we estimate a panel Tobit model with random effects of individual investment to examine the treatment effect of the appropriation level as follows:

$$a_{it} = \delta_L Low_i + \delta_H High_i + x_{it}\gamma + v_i + v_{it}. \tag{4}$$

The dependent variable is the investment of subject i in period t. The variable Low_i is an indicator equal to one if subject i belongs to the *Low* treatment and zero otherwise. The variable $High_i$ is an indicator equal to one if subject i belongs to the *High* treatment, and zero otherwise. The vector of variables x_{it} includes the subjects' characteristics and period-specific fixed effects (FEs). The latter set of variables captures the learning effect over time. The error terms have two components, namely an unobserved individual heterogeneity v_i following a normal distribution N(0,σ_v) and a random error v_{it} following a normal distribution.

The empirical results of Eq. (4) are reported in Column 1 of Table 3. Encouragingly, the coefficients δ_L and δ_H are negative and significant at 1% level, and δ_L is less negative than δ_H. In sum, we provide evidence to support Hypothesis 1 and summarize the results as follows:

Result 1. *The investment level in the Zero treatment is significantly higher than that in the Low treatment, which in turn is significantly higher than that in the High treatment.*

3.4. Deviation from the Nash equilibrium

Despite the anticipated effect of appropriation on individual investment, the individual investments are still different from their Nash equilibrium levels in the three treatments. Specifically, the equilibrium levels of investment are 20 in the *Zero* and *Low* treatments and 0 in the *High* treatment.[9] We also observe a larger deviation of investments from the Nash

[9] The average investments of *Zero* and *Low* treatments are significantly different from 20 at the 5% and 1% levels, respectively. The average investment of *High* treatment is significantly different from 0 at the 1% level.

Table 3
Treatment effects on investments and low appropriation choice.

Dependent variable	Investment			Low appropriation choice
Data set	Zero, low & high	Communication × matching	C-Fix	Communication × matching
Model	(1)	(2)	(3)	(4)
Low (δ_L)	−5.10***			
	(0.90)			
High (δ_H)	−9.72***			
	(0.91)			
Communication (δ_C)		−1.39*		0.20
		(0.82)		(0.29)
Fix (δ_F)		1.50*		0.26
		(0.83)		(0.28)
Communication × fix (δ_{CF})		9.01***		0.69*
		(1.16)		(0.37)
Rotating low (δ_{RL})			11.97***	
			(2.77)	
Both low (δ_{BL})			14.63**	
			(6.26)	
Female	−0.38	−0.21	1.26	0.20
	(0.88)	(0.64)	(2.61)	(0.21)
Math	2.77	2.14	3.99	−0.24
	(1.84)	(1.37)	(4.37)	(0.42)
Risk tolerance	0.27	−0.26	0.59	−0.16
	(1.55)	(1.11)	(4.30)	(0.37)
Individualism	−1.31	−4.42***	0.95	−1.18***
	(1.09)	(0.95)	(3.77)	(0.31)
Fairness	−0.48	3.41***	4.13	0.45
	(1.31)	(1.19)	(3.62)	(0.36)
Period FE	Yes	Yes	Yes	Yes
Observations	1,438	1,904	478	954
Number of subjects	144	192	48	190
Log likelihood	−3471.00	−4671.37	−724.62	−319.61
Diagnostic tests (p-value)				
LR test for $\mathrm{Var}(v_i)=0$	<0.01	<0.01	<0.01	<0.01
H_0: $\delta_L=\delta_H$	<0.01			
H_0: $\delta_C+\delta_{CF}=0$		<0.01		<0.01
H_0: $\delta_F+\delta_{CF}=0$		<0.01		<0.01
H_0: $\delta_{RL}=\delta_{BL}$			>0.10	

Note: All models in Column 1–3 (4) are estimated with panel Tobit (Probit) models including random effects at subject level. LR Test for $\mathrm{Var}(v_i)=0$ tests the null hypothesis that the percent contribution to the total variance of the panel-level variance component is zero and rejects the null of no random effect in all columns. The definitions of the explanatory variables are given in Table 1. Standard errors are given in brackets. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

equilibrium in the *High* treatment than in the *Zero* and *Low* treatments. In this sub-section, we explain the asymmetric deviation in the three treatments with the presence of bounded rationality and social preference.[10]

To incorporate the behavioral factor of social preference induced by group identification, following Charness and Rabin (2002) and Chen and Li (2009), we assume that the expected utility of each member *ik* is a weighted average of her expected payoff and the expected payoff of her partner *im*:

$$EU_{ik} = (1-s)E\pi_{ik} + sE\pi_{im}, \quad i,k,m = 1,2 \text{ and } k \neq m,$$

where the function of expected payoffs $E\pi_{ik}$ and $E\pi_{im}$ is given in Eq. (1), and $s \in [0,1]$ measures the degree of prosociality. A larger s represents that the member is more concerned about her partner's payoff and is less concerned about her own payoff.[11]

Given that the appropriation levels β are exogenously set, members in alliances only choose the first-stage investment level to maximize their expected utility. However, members with bounded rationality might be erroneous in their decision makings. We adopt the logit Quantal Response Equilibrium (QRE) method (McKelvey and Palfrey, 1995) to model members' noisy decision-making behavior. Let $EU_{ik}(a_{ik},\ p_{-ik})$ denotes the expected utility of member *ik* to invest $a_{ik} \in A = \{L, M, H\}$ when the other three members (the partner and the two members in the competing alliance) choose investment levels

[10] These two factors are common explanations for the deviation of observed choices from theoretical benchmarks in the experimental studies of public goods, multistage tournaments, and contests (Andreoni, 1995; Houser and Kurzban, 2002; Sheremeta, 2013). Other explanations are reviewed in Sheremeta (2013).

[11] A salient group identity evoked by the between-alliance competition can cause a blurring of the boundaries between individual and alliance welfare and induce the subjects to pursue the collective interests of their alliance, according to the social identity theory (Tajfel and Turner, 1979; Chen and Li, 2009).

Table 4
Estimated s and λ at the QRE.

Parameter	*Estimate*	Choice	*Observed choice probability*			*Predicted choice probability*		
			Zero	*Low*	*High*	*Zero*	*Low*	*High*
s – Prosociality	0.20*** (0.00)	$a_{ik} = L$	0.08 (0.26)	0.04 (0.19)	0.14 (0.34)	0.04	0.06	0.26
λ - Precision parameter	0.59*** (0.04)	$a_{ik} = M$	0.13 (0.34)	0.33 (0.47)	0.72 (0.45)	0.19	0.20	0.51
Likelihood	1134.75	$a_{ik} = H$	0.79 (0.41)	0.63 (0.48)	0.14 (0.35)	0.77	0.74	0.23

Note: Observations = 1,438. Choice $a_{ik} = L$ for $0 \le \text{investment} < 7$, Choice $a_{ik} = M$ for $7 \le \text{investment} < 14$, and Choice $a_{ik} = H$ for $14 \le \text{investment} \le 20$. The equations embed with those two parameters are illustrated as follows: take member *ik* as an example; the total expected utility EU_{ik} for member *ik* to invest $a_{ik} \in A = \{L,M,H\}$ is: $EU_{ik}(a_{ik},p_{-ik}) = (1-s)E\pi_{ik} + sE\pi_{im}$, where the parameter *s* represents each member's degree of prosociality. In the logit QRE, the other three members choose investment levels according to the mixed strategy profile p_{-ik}. The probability that member *ik* chooses investment level a_{ik} is proportional to $\exp(\lambda EU_{ik}(a_{ik},p_{-ik}))$ with precision parameter λ. We report the parameter estimate and then followed with their standard errors given in brackets. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

according to the mixed strategy profile p_{-ik}. The probability that member *ik* chooses an investment level a_{ik} is

$$p(a_{ik}) = \frac{\exp(\lambda EU_{ik}(a_{ik},\ p_{-ik}))}{\sum_{a'_{ik} \in A} \exp\left(\lambda EU_{ik}\left(a'_{ik},\ p_{-ik}\right)\right)}, \quad i, k = 1, 2$$

Thus, the likelihood of an investment level being chosen is positively dependent on the expected utility it yields. The precision parameter $\lambda \ge 0$ determines the sensitivity of the choice probabilities with respect to expected utilities. A larger λ represents that the probability of choosing each investment level is closer to that under utility maximization.

The parameters *s* and λ are estimated by maximum likelihood.[12] We reduce the state space of investment to make the computation tractable because we need to keep track of the probability distribution and payoff of each possible outcome for computing expected utility. More specifically, we categorize choice $a_{ik} = L$ for $0 \le a_{ik} < 7$, $a_{ik} = M$ for $7 \le a_{ik} < 14$, and $a_{ik} = H$ for $14 \le a_{ik} \le 20$. The middle panel of Table 4 reports that the respective proportions of $a_{ik} = L$, $a_{ik} = M$, and $a_{ik} = H$ are 0.08, 0.13, and 0.79 in the *Zero* treatment, 0.04, 0.33, and 0.63 in the *Low* treatment, and 0.14, 0.72 and 0.14 in the *High* treatment. While the respective Nash equilibria are $a_{ik} = H$ in the *Zero* and *Low* treatments and $a_{ik} = L$ in the *High* treatment, it indicates that the larger deviation from the Nash equilibrium in the *High* treatment is maintained after we aggregate the data. We expect the data aggregation may affect the quantitative results of our structural estimates but not the qualitative results.

The left panel of Table 4 reports that the parameters s and λ are estimated significantly at 0.20 and 0.59, respectively. It suggests that our subjects are not completely individualistic and have bounded rationality. Further, many experimental studies of public goods and contests estimate the value of λ to lie between 0.4 and 1 (Goeree et al., 2002; Gronberg et al., 2012; Lim et al., 2014), and our results are consistent with those studies.

Consistent with the observed empirical pattern, the right panel of Table 4 reports that our model generates a larger deviation from the Nash equilibrium in the *High* treatment than the *Zero* and *Low* treatments. Intuitively, in the *Zero* and *Low* treatments, bounded rationality pushes the investments below the Nash equilibrium of $a_{ik} = H$, whereas social preference encourages the subjects to raise investments; therefore, only the bounded rationality pushes the outcome away from the Nash equilibrium. In the *High* treatment, both bounded rationality and social preference raise the investments above the Nash equilibrium of $a_{ik} = L$.

4. Informal agreement to deter appropriation

The above results suggest a negative relationship between investment and appropriation level. Our model also predicts that the subjects who are free to make an appropriation decision always choose the highest possible appropriation level. Thus, when the upper bound of appropriation level is sufficiently high (i.e., $\bar{\beta} > T$), subjects will be discouraged from investing. In such a setting, we are interested in whether an informal agreement can induce subjects to voluntarily reduce their appropriation level and hence boost their partner's investment. We explore the effects of informal agreements under both a single interaction and finitely repeated interactions between alliance members.

4.1. Hypotheses development

We derive the hypotheses more formally in Appendix B and briefly summarize them here. We first consider the case of a single interaction. When pre-play communication is possible, alliance members may try to reach an agreement to deter appropriation, seeking to ensure their investments are protected and not used against them in the second stage. However, under single interaction, subjects should have no incentive to honor such an agreement, as they derive a benefit but bear no cost from breaking the agreement. The communication, hence, only serves as cheap talk, and subjects behave the same as in a one-shot game with no communication.

[12] We thank one referee for suggesting the structural analysis.

Hypothesis 2. *When pre-play communication about appropriation behavior is possible in a one-shot game, the subjects will not honor an informal agreement to deter appropriation. Thus, the investment and appropriation levels are the same as those in a one-shot game without communication.*

Under the condition of finitely repeated interactions, whether the subjects honor the agreement or not is affected by their future interaction with the fixed partner. Following Kreps et al. (1982), we assume there are two types of subjects: reciprocal and rational. A reciprocal subject will honor the agreement to deter appropriation in any given round if his/her partner complied with the agreement in the previous round; otherwise, he/she will retaliate by choosing the high appropriation level in the current and all subsequent rounds. A rational subject, however, is only interested in his/her own payoff maximization. Given uncertainty about the partner's type, even a rational subject will honor the agreement in all but the last few rounds if he/she believes it is sufficiently likely the partner is reciprocal, as his/her expected payoff from inducing the partner's future cooperation by honoring the agreement outweighs that from breaking the agreement to take advantage of the partner now. Consequently, the investment levels are higher, and the appropriation levels are lower than those in a repeated game without communication in which an agreement against appropriation is impossible.

Hypothesis 3. *When pre-play communication about appropriation behavior is possible in a repeated game, the subjects will honor an informal agreement to deter appropriation in all but the last few rounds. Thus, the investments are higher and the appropriation levels are lower than those in a repeated game without communication.*

4.2. Experimental design and procedure

To test Hypotheses 2 and 3, we conduct a second experiment with a 2×2 factorial design; we vary the communication possibility (possible vs. impossible) and the matching method (random vs. fixed). The four treatments are named *C-Random, NC-Random, C-Fix*, and *NC-Fix*. The basic setting and parameter choices are the same as in the first experiment except that the subjects are free to choose between low (0.1) and high (0.9) levels of appropriation in the four treatments, conditional on reaching stage two.

We use a random matching protocol for *C-Random* and *NC-Random* and a fixed matching protocol for *C-Fix* and *NC-Fix*. In the fixed matching protocol, the subjects are paired with the same partner for all ten rounds, but the matching between competing alliances is still random in each round to prevent potential collusion between alliances. An example of the experimental instructions can be found in Appendix A.

In addition, in the treatments of *C-Random* and *C-Fix*, the subjects can discuss their intended appropriation choices in a 60-s pre-play communication stage at the beginning of each round. They can communicate by sending any of the following four sentences:

"I will choose 0.1 in the second stage";
"I suggest that you choose 0.1 in the second stage";
"I agree"; and
"I disagree."

With these sentences, subjects are able to reach an informal agreement to choose low appropriation levels.[13]

We run four between-subject sessions for each of the four treatments. The sample consists of 192 students from multiple disciplines divided into 16 sessions, each with 12 participants. The last column of Table 1 reports the descriptive statistics of our subjects' characteristics in this sample. As in the first experiment, the subjects are given 60 s to make each choice; otherwise, the computer randomly makes a decision for them. We drop from our analysis 16 of the 1920 investment decisions and 6 of the 960 appropriation decisions that are auto-selected observations. The average earnings are RMB60.31, including a show-up fee of RMB5. Each session lasts 70 minutes on average.

4.3. Results

Sections 4.3.1 and 4.3.2 examine the effects of communication and matching protocol on subjects' investment and appropriation choices respectively. Section 4.3.3 explores whether the informal agreements to deter appropriation are reached and honored when communication is possible, and how they affect subjects' investments in the *C-Random* and *C-Fix* treatments.

4.3.1. Effects of communication and matching protocol on investment

Fig. 2 depicts the average individual investment across the treatments of *NC-Random, C-Random, NC-Fix*, and *C-Fix*, and finds two interesting results. First, the effect of communication is stronger under the fixed matching protocol than under the random matching protocol. There is a positive difference in average individual investment between *C-Fix* and *NC-Fix*

[13] This structured communication design is deliberately chosen to prevent direct communication about investments, as our interest is how communication about appropriation affects investments. We also exclude the possibility that a subject could claim or suggest that their partner choose the high appropriation level of 0.9, as we aim to study the informal agreements on deterring appropriation rather than those on promoting appropriation. In practice, it is also unlikely that parties will reach an agreement to appropriate more from their partners.

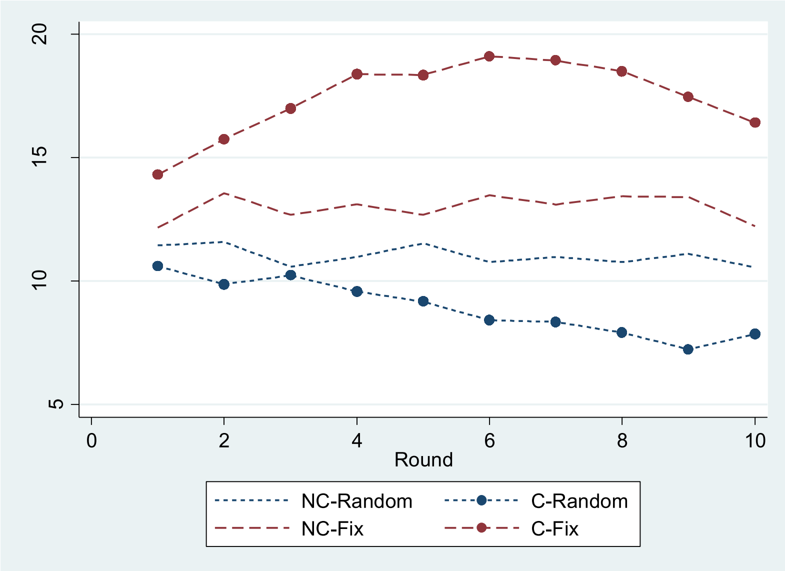

Fig. 2. Average Investments over Time for the four Communication x Matching Treatments.
Note on Communication x Matching treatments: For the sample with *NC-Random, C-Random, NC-Fix*, and *C-Fix* treatments, *NC-Random* indicates the treatment with no communication and the random matching protocol. *C-Random* indicates the treatment with communication and the random matching protocol. *NC-Fix* indicates the treatment with no communication and the fixed matching protocol. *C-Fix* indicates the treatment with communication and the fixed matching protocol.

treatments, whereas there is a negative difference in average individual investment between *C-Random* and *NC-Random* treatments. Second, the average individual investment in the *NC-Fix* treatment is higher than that in the *NC-Random* treatment.

Consistently, Panel B of Table 2 reports that the average individual investment in the *C-Fix* treatment is 17.42, in the *NC-Fix* treatment is 12.99, and their difference is significant (at the 1% level for the *t*-test and 5% level for the Wilcoxon rank-sum test). It suggests a positive effect of communication possibility on investments in the fixed matching protocol. Further, the average individual investment in the *NC-Random* treatment is 11.03, and that in the *C-Random* treatment is 8.93, but their difference is insignificant. Finally, the difference of investment between the *NC-Fix* and *NC-Random* treatments is marginally significant (at the 10% level for the Wilcoxon rank-sum test but insignificant for the *t*-test), suggesting a positive effect of the fixed matching protocol on investments in the absence of communication possibility.

To control for the effects of subjects' characteristics and learning effect on investment, we estimate a panel Tobit model with random effects of individual investment to examine the treatment effect of communication and matching protocol as follows:

$$a_{it} = \delta_C Communication_i + \delta_F Fix_i + \delta_{CF} Communication_i \times Fix_i + x_{it}\gamma + v_i + v_{it}. \quad (5)$$

The empirical results of Eq. (5) are reported in Column 2 of Table 3. First, the coefficient of δ_C is negative and significant at the 10% level, which indicates that the average individual investment in the *NC-Random* treatment is significantly higher than that in the *C-Random* treatment. One possible cause is the discouraging effect of communication result. In the *NC-Random* treatment, when making investments in stage one, subjects cannot entirely preclude the possibility that their partner may be sufficiently altruistic (or simply irrational) to choose the low appropriation level in stage two. In the *C-Random* treatment, however, if no agreement is reached during communication, subjects can affirm that their partner will not cooperate by choosing the low appropriation level. Even if an agreement against appropriation is reached, it is frequently violated in *C-Random*, as reported later. Thus, the subjects are discouraged from investing as if the appropriation levels are exogenously set to be high.[14] To explore the evidence, we find that the investment levels in the *High* treatment are significantly lower than those in *NC-Random*, but the investment levels in the treatments of *High* and *C-Random* are not significantly different.[15]

[14] Cooper et al. (1989) also found that the failure to coordinate at the communication stage made the game stage more non-cooperative. See Ke et al. (2015) for similar findings.

[15] In both regressions, the dependent variable is the individual investment and the explanatory variable is the dummy for the *High* treatment. We use the data of the *NC-Random* and *High* treatments in the first regression to compare these two treatments. The estimated coefficients are as follows: Constant = 11.03, $p < 0.01$; *High* treatment dummy = −1.32, $p < 0.05$. We then use the data of the *C-Random* and *High* treatments in the second regression. The estimated coefficients are as follows: Constant = 8.89, $p < 0.01$; *High* treatment dummy = 0.82, $p > 0.1$. These results are robust to the inclusion of individual characteristics as explanatory variables.

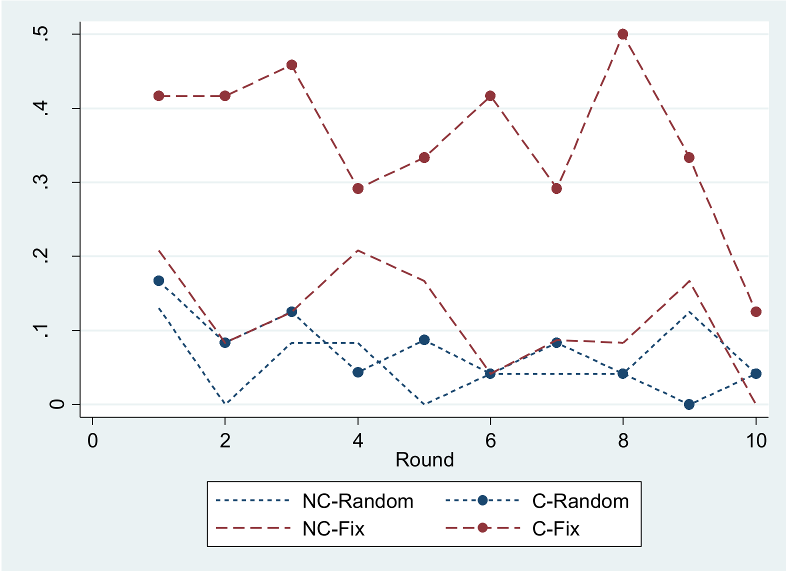

Fig. 3. Proportion of Low Appropriation Choice over Time for the Four Communication x Matching Treatments.
Note on Communication x Matching treatments: For the sample with *NC-Random, C-Random, NC-Fix*, and *C-Fix* treatments. *NC-Random* indicates the treatment with no communication and the random matching protocol. *C-Random* indicates the treatment with communication and the random matching protocol. *NC-Fix* indicates the treatment with no communication and the fixed matching protocol. *C-Fix* indicates the treatment with communication and the fixed matching protocol.

Second, the coefficient δ_F is positive and significant at the 10% level, which indicates that the investment in the *NC-Fix* treatment is significantly higher than that in the *NC-Random* treatment. These results may be caused by the subjects' reputation concerns under repeated interactions, where the subjects increase their present investment to induce future cooperation from their partner (Kreps et al., 1982; Palfrey and Rosenthal, 1994).[16] Third, focusing on the treatments with fixed matching protocol, we find that the investment in the *C-Fix* treatment is significantly higher than that in the *NC-Fix* treatment as the linear restriction $\delta_C + \delta_{CF} = 0$ is rejected at 1% level.

Among the subjects' characteristics, the effect of *Individualism* is significantly negative, suggesting that those who are only concerned about their own earnings tend to refrain from investing to reduce the chance of being defeated in stage two. In contrast, those who try to earn the most for their alliance are more likely to increase their investments and, consequently, the probability for their alliance to win stage one. In addition, the significantly positive effect of *Fairness* suggests that subjects are more willing to invest if they believe their partner is a fair person who will reciprocate their investment by being cooperative.

In sum, we provide evidence to support Hypothesis 3 but not Hypothesis 2 in terms of investment level and summarize the results as follows:

Result 2. *The investment levels in the four communication-matching treatments can be ranked as follows: C-Fix > NC-Fix > NC-Random ≥ C-Random.*[17]

4.3.2. Effects of communication and matching protocol on appropriation choice

This sub-section analyzes the effects of communication and matching protocol on subjects' appropriation choices. Fig. 3 depicts the proportion of low appropriation choices in each round for the four communication-matching treatments. Apparently, there are more low appropriation choices made in the *C-Fix* treatment than in the other three treatments.

Panel B of Table 2 reports that the probability of choosing a low appropriation level in the *NC-Random* and *C-Random* treatments are 0.06 and 0.07, respectively, and their difference is insignificant. The probability to choose a low appropriation level in the *NC-Fix* treatment is 0.12, not significantly different from that in the *NC-Random* treatment. Further, the probability to choose low appropriation level in the *C-Fix* treatment is 0.36, which is significantly higher than that in the *NC-Fix* treatment (at the 1% level for the *t*-test and 5% level for the Wilcoxon rank-sum test).

[16] To explore the evidence, we estimate two panel Tobit regressions and find that the correlation between the subjects' present investment and their partner's investment in the previous round is significantly positive in *NC-Fix* but not significant in *NC-Random*. These results are available upon request from the authors.

[17] The "≥" between *NC-Random* and *C-Random* indicates that the difference between these two treatments is significant according to the random-effects Tobit regression but insignificant according to the non-parametric tests.

Table 5
Winning probability with/without agreements during pre-play communication (alliance level).

First-stage status	No agreement	Agreement	Total
C-Random			
# of cases for all alliances	153	87 (One Low: 73; Both Low: 14)	240
# of cases for winning alliances	81	39 (One Low: 32; Both Low: 7)	120
Winning probability	0.53	0.45	
C-Fix			
# of cases for all alliances	126	114 (One Low: 100; Both Low: 14)	240
# of cases for winning alliances	45	75 (One Low: 65; Both Low: 10)	120
Winning probability	0.36	0.66	
C-Fix − C-Random			
Winning probability	-0.17^{***} [0.06]	0.21^{***} [0.07]	
Wilcoxon rank-sum test	< 0.01	< 0.01	

Note: Each observation for the non-parametric tests is the proportion of alliances in the corresponding category winning the first-stage tournament in a session. One Low (Both Low): The two alliance partners agree that one of them (both of them) will choose the low appropriation level. Standard errors of two-sided equal-mean *t*-tests are in brackets. *** $p<0.01$, ** $p<0.05$, and * $p<0.1$. We report the p-values of two-sided Wilcoxon rank-sum test.

To control for the effects of subjects' characteristics and learning effect on low appropriation choice, we estimate a panel Probit model with random effects of low appropriation choice to examine the treatment effect of communication and matching protocol as follows:

$$\beta_{it} = \delta_C Communication_i + \delta_F Fix_i + \delta_{CF} Communication_i \times Fix_i + x_{it}\gamma + \nu_i + \nu_{it}. \quad (6)$$

The dependent variable is the binary variable of subject *i* in round *t*, choosing a low appropriation level. The empirical results of Eq. (6) are reported in Column 4 of Table 3. First, the insignificant coefficient of δ_C indicates that the probability of choosing a low appropriation level does not differ between *NC-Random* and *C-Random*. Second, the insignificant coefficient of δ_F indicates that the probability of choosing a low appropriation level does not differ between *NC-Fix* and *NC-Random*. Third, focusing on the treatment with a fixed matching protocol, we find that the probability of choosing a low appropriation level in *C-Fix* is significantly higher than that in *NC-Fix*, as the linear restriction $\delta_C + \delta_{CF} = 0$ is rejected at the 1% level. Finally, among the subjects' characteristics, the effect of *Individualism* is significantly negative for the same reason as that discussed in Section 4.3.1.

In sum, we provide evidence to support Hypotheses 2 and 3 in terms of appropriation level. We conclude the ordering of the probability of choosing a low appropriation level among the four treatments in Result 3.

Result 3. *The probability of choosing the low appropriation level in the four communication-matching treatments can be ranked as follows: C-Fix > NC-Fix = NC-Random = C-Random.*

4.3.3. Informal agreements to deter appropriation

In the *C-Random* and *C-Fix* treatments, the subjects are offered pre-play communication opportunities that allow them to reach informal agreements against the appropriation. The two possible agreements are "One Low" (i.e., the alliance partners jointly agree that one of them will choose the low appropriation level in stage two) and "Both Low" (i.e., both partners agree to choose the low appropriation level in stage two).[18] Table 5 summarizes the number of cases in which the alliances do or do not reach an agreement and the probability for these alliances to win the first stage. In each treatment, 24 alliances engage in within-alliance communication in each of the ten rounds. We thus have 240 communication cases for all alliances and one half of them for the winning alliances.

Table 5 reports that in the *C-Random* treatment, the alliances reach an agreement in 87 out of 240 cases. Within these cases, the alliances win the first-stage tournament in 39 cases. The winning probability for the alliances with an agreement is only 45%. Within the 153 communication cases in which no agreement is reached, the alliances win the first stage in 81 cases. The winning probability for the alliances with no agreement is 53%. These results suggest that an agreement does not provide the alliances with a relative advantage to win in the *C-Random* treatment.

In the *C-Fix* treatment, however, the alliances reach an agreement in 114 out of 240 cases. Within these, the alliances win the first stage in 75 cases. The winning probability for those with an agreement is thus 66%. Within the 126 cases in which no agreement is reached, the alliances win the first stage in 45 cases. The winning probability for those with no agreement is only 36%. These results suggest that in the *C-Fix* treatment, the alliance with an agreement are motivated to invest, and their chances of winning are thus greatly improved relative to those with no agreement.

[18] An agreement is regarded as being reached only if both alliance members explicitly express their intention to follow this plan. For example, when one subject selects "I will choose 0.1 in the second stage," the agreement is reached only if the other subject replies with "I agree" or "I suggest that you choose 0.1 in the second stage." If the other subject replies with "I disagree" or "I will choose 0.1 in the second stage" or keeps silent, then it is regarded that no agreement has been reached. We have hired two independent coders to perform the content coding and then checked the reliability of their coding. Given our simple coding rule, any disagreement between the two coders was typically a coding error which was then corrected.

Table 6
Mapping of agreements to actual appropriation choices for first-stage winners (individual level).

Agreements on appropriation choices	Actual appropriation choices				
	Both High	*A* Low *B* High	*A* High *B* Low	Both Low	Total
Panel A: C-Random					
No agreement	**143**	9	9	0	161
A Low *B* High	**26**	5	0	1	32
A High *B* Low	**24**	0	5	1	30
Both Low	**12**	1	1	0	14
Total	205	15	15	2	237
Panel B: C-Fix					
No agreement	**82**	4	4	0	90
A Low *B* High	4	**59**	0	2	65
A High *B* Low	4	0	**59**	2	65
Both Low	0	1	1	**18**	20
Total	90	64	64	22	240

Note: *A* represents the subject himself/herself. *B* represents his/her partner.

Table 5 also reports the cross-treatment comparison of the winning probability. The alliances with no agreement are significantly less likely to win in *C-Fix* than in *C-Random*; however, the alliances with an agreement are significantly more likely to win in *C-Fix* than in *C-Random*. The relative success of agreement in *C-Fix* is owing to the high frequency of honoring agreements, which inspires the subjects to cooperate within alliances.

To illustrate the frequency of honoring agreements in these two treatments, we report the mapping from agreements to actual appropriation choices for the first-stage winners at the individual level in Table 6. The rows represent the categories of agreements, and the columns represent the categories of actual appropriation choices. Each entry lists the number of observations in the corresponding category of agreement and that of the actual choice. Note that at the individual level, there are two subtypes of "One Low" agreements: the agreement of "*A* low *B* high" (only the subject chooses the low appropriation level) and the agreement of "*A* high *B* low" (only the subject's partner chooses the low appropriation level).

Panel A of Table 6 reports that the subjects agree to choose the low appropriation level by reaching an agreement of "*A* low *B* high" or "Both Low" in 19.41% of cases (i.e., $(32+14)/237$, as shown in the last column) in the *C-Random* treatment. However, only 10.87% of these agreements are actually honored (i.e., $(5+0)/(32+14)$, as shown on the diagonal).[19]

Panel B of Table 6 reports that the subjects agree to choose the low appropriation level by reaching an agreement of "*A* low *B* high" or "Both low" in 35.42% of cases (i.e., $(65+20)/240$, as shown in the last column) in the *C-Fix* treatment. 90.59% of these agreements are honored (i.e., $(59+18)/(65+20)$, as shown on the diagonal).

These results support Hypotheses 2 and 3 in terms of agreement breach. A two-sided Wilcoxon rank-sum test using session averages as observations shows that the proportions of low appropriation commitments being fulfilled in these two treatments are significantly different ($p\text{-}value = 0.02$). The higher frequency of honoring agreements leads to a significantly greater proportion of low appropriation choices in the *C-Fix* treatment relative to the *C-Random* treatment. Result 4 summarizes these findings.

Result 4. *The agreements reached in C-Fix relate to winning the first stage, but those in C-Random do not. Further, the subjects in C-Fix are more likely than those in C-Random to comply with their agreements to choose the low appropriation level, which leads to a significantly greater proportion of low appropriation choices in C-Fix than in C-Random.*

In the *C-Fix* treatment, a subject is more likely to agree to a low appropriation choice in a given period if his/her partner in the previous period agreed to a low appropriation choice or actually honored the promise (p-value < 0.01 for both correlations).[20] In contrast, in the *C-Random* treatment, a subject's choice on agreeing to a low appropriation choice in a given period does not depend on whether or not his/her partner in the previous period agreed to a low appropriation choice or honored the promise (p-value > 0.1 for both correlations).

In short, the subjects' behavior in agreement negotiation in the *C-Random* treatment is independent of their random-matched partner's previous behavior in the agreement negotiation and execution, whereas the subjects' behavior in agreement negotiation in the *C-Fix* treatment highly depends on their fixed-matched partners' previous behavior in the agreement negotiation and execution. The high frequency of honoring agreements encourages the subjects to reach agreements in the *C-Fix* treatment.

Due to the high frequency of honoring agreements, the subjects in the *C-Fix* treatment are able to develop long-term (cross-period) cooperative strategies with their fixed partner via the agreements. Among the 24 alliances in *C-Fix*, one adopts

[19] Although the behavior of honoring informal agreements even in some single-interaction settings was also reported in the literature and explained with, e.g., guilt aversion (Battigalli and Dufwenberg, 2007), lie aversion (Lundquist et al., 2009), taste for consistency (Cialdini et al., 1995), or exposure to social disapproval (Rege and Telle, 2004), our results suggest that these behavioral factors may not play an important role in our setting.

[20] We use the Spearman's rank correlation test to examine the correlation of interested variables in this part. We thank one referee for suggesting the cross-period analysis of agreement negotiation.

Table 7
Effects of individual characteristics.

Dependent variable	Committing to Low appropriation choice	Following Low appropriation choice commitment	Adopting a cooperating strategy
Data set		*C-Fix*	
Model	(1)	(2)	(3)
Female	−0.28	−0.040	−0.53
	(0.35)	(0.42)	(0.67)
Math	−0.16	0.40	1.18
	(0.57)	(0.67)	(1.06)
Risk tolerance	−0.90	−1.50**	−2.88**
	(0.55)	(0.67)	(1.17)
Individualism	−1.84***	−2.46***	−3.22***
	(0.49)	(0.61)	(1.01)
Fairness	1.13**	0.84	1.85**
	(0.46)	(0.54)	(0.90)
Period FE	Yes	Yes	Yes
Observations	480	480	48
Number of subjects	48	48	48
Log likelihood	−225.40	−163.21	−17.76
Diagnostic tests (*p*-value)			
LR test for Var(v_i) = 0	<0.01	<0.01	N/A

Note: Columns 1–2 are estimated with panel Probit models including random effects at the individual level. LR Test for Var(v_i) = 0 tests the null hypothesis that the percent contribution to the total variance of the panel-level variance component is zero, and rejects the null of no random effect in all columns. Column 3 is estimated with Probit because there is no variation in the dependent variable across the period. The definition of explanatory variables is given in Table 1. Standard errors are given in brackets. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

the "Both Low" strategy in which the two subjects reach the "Both Low" agreement in each round and adhere to it upon reaching stage two. Further, eight alliances adopt the "Rotating Low" strategy in which the partners agree to rotate the "*A* low *B* high" and "*A* high *B* low" agreements in each round and adhere to the agreement upon reaching stage two.[21,22] We find that in an average round, about 77% of the alliances that adopt cooperative strategies win the first-stage tournament, which suggests that these alliances tend to invest more and are thus more likely to win.

We estimate a panel Tobit model with random effects of investment to examine the effects of the "Both Low" strategy and the "Rotating Low" strategy as follows:

$$a_{it} = \delta_{RL} Rotating\ Low_i + \delta_{BL} Both\ Low_i + x_{it}\gamma + v_i + \nu_{it}. \quad (7)$$

The variable *Rotating Low*$_i$ takes the value one if the subject adopts the "Rotating Low" strategy and 0 otherwise. The variable *Both Low*$_i$ is defined analogously. The control variables include subjects' characteristics and period FEs.

Column 3 of Table 3 reports the results of Eq. (7). The coefficients of *Rotating Low*$_i$ and *Both Low*$_i$ are significantly positive, suggesting that the subjects adopting either cooperative strategy invest significantly more than the other subjects in the *C-Fix* treatment. However, we cannot reject the null hypothesis that the effects of the two strategies on the investments are equal at any conventional level. Result 5 summarizes these findings.

Result 5. *The subjects' behavior in agreement negotiation is independent of their random-matched partner's previous behavior in the agreement negotiation and execution in C-Random but highly dependent on their fixed-matched partners' previous behavior in the agreement negotiation and execution in C-Fix. The subjects in C-Fix are thus able to develop either of the two long-term cooperative strategies with their fixed partner: "Both Low" and "Rotating Low." The subjects adopting either cooperative strategy invest significantly more than the other subjects in C-Fix and are thus more likely to win the first-stage tournament.*

Further, the above results suggest that even in the presence of communication possibility and repeated interaction, not all subjects cooperate with their partners. A subject's decisions to reach and/or honor an agreement and form a cooperative strategy may also depend on his/her characteristics. We estimate three panel Probit models to examine the effects of subjects' characteristics on the decisions of 1) committing to the low appropriation choice in the agreement, 2) following the commitment, and 3) adopting a cooperative strategy. The results are reported in Table 7.

First, we find that only the effect of *Individualism* is significantly negative in all three models, suggesting that self-oriented subjects are less likely than group-oriented subjects to commit to the low appropriation choice, to follow these

[21] The subjects need at least one round to learn about their partner's strategy and will not honor their agreement in the last round. Thus, we regard an alliance as adopting a cooperative strategy if it sticks to this strategy for at least eight rounds.

[22] It has been reported that contestants in a tournament with communication may collude by adopting an "effort rotation" scheme in which they take turns expending effort and winning the tournament (Sutter and Strassmair, 2009). In our case, the "Rotating Low" strategy ensures both partners receive the second-stage prize with the equal frequency upon reaching that stage. In contrast, under the "Both Low" strategy which one of the two partners wins the second stage mainly depends on the second-stage luck. Thus, the "Both Low" strategy is inferior to the "Rotating Low" strategy in terms of fairness and thus less likely to be adopted by the subjects.

commitments, and to adopt cooperative strategies.[23] Second, the effect of *Fairness* is significantly positive in Columns 1 and 3, suggesting that those who believe their partner to be fair are more likely to commit to the low appropriation choice and to build a cooperative relationship with their partner than those who believe they are paired with a selfish partner. Third, the effect of *Risk Tolerance* on the inclination to honor a commitment and adopt a cooperative strategy is significantly negative (see Columns 2–3). A possible explanation is that more risk-seeking (i.e., risk-tolerant) subjects tend to engage in opportunistic activities and renege on their promises and are thus less likely to develop long-term cooperative relationships with their partners.

5. Conclusions

We study a two-stage tournament in which the alliance members cooperate in the between-alliance tournament in the first stage. Upon victory, they can appropriate and use their partners' first-stage investments to compete against each other in the second stage. Anticipating the second-stage appropriation behavior of their partners, members are reluctant to invest in the first stage. Using experimental data, we test this hold-up problem and find evidence for the negative relationship between appropriation levels and investments. We further explore whether an informal agreement against appropriation can encourage investments. The results suggest that when pre-play communication is possible, the subjects are more likely to break their agreements under random matching than under fixed matching. Consequently, the investment levels are not improved by the communication possibility under random matching. In contrast, the high frequency of honoring agreements under fixed matching motivates the subjects to develop long-term cooperative strategies and thus encourages investments.

Two possible implications can be derived from our results. First, reputation concerns of the alliance members improve the effectiveness of informal agreements against appropriation; thus, a reputation-evaluation system can help promote the success of alliances. Second, we find that even in the presence of communication possibility and repeated interaction, not all members build cooperative relationships with their partners. In particular, those who are inclined to earn the most for their alliance rather than for themselves and those who believe their partners to be fair when dividing interests are more likely to develop cooperative strategies. Thus, to create a healthy environment for long-term cooperative relationships, alliances should aim to promote a mutually beneficial group norm, to reinforce the group identification of members, and to inculcate their beliefs in collective goals.

Acknowledgment

Shaoqing Huang would like to acknowledge financial support from the National Social Science Foundation of China (Project Number: 11CJY007).

Appendix A. Experimental instructions

General instructions

Welcome to the Smith Experimental Economics Research Center. You are now participating in a session of an economic decision-making experiment.

Your earnings will be calculated in a fictitious currency called experimental points. At the end of the session today, you will be paid 1 RMB for each 1200 experimental points earned.

Your earnings are anonymous; no other participant in the room will know your earnings unless you choose to tell them. You will never learn the identity of your counterpart, nor will they ever learn who you are.

Please do not talk, exclaim, or otherwise communicate with the other participants during the session. Interactions with your counterpart will take place through the computer program. If you have a question, please raise your hand and a monitor will come to you to answer your question privately. Any participants violating these rules will be asked to leave the session and will forfeit their earnings.

Are there any questions before we continue? If so, please raise your hand.

Specific instructions for Zero, Low *and* High *treatments [treatments with selectable second-stage choice] <treatments with communication> {treatments with fixed matching}*

There are 10 experimental rounds in the following setting. In each round, you will be randomly and anonymously matched with a counterpart who is another participant in this room. Your two-person group will be matched with another randomly formed two-person group. After each round, you will be matched with a different counterpart. Your new

[23] The variable *Individualism* captures the degree of group identification. The social psychology literature has noticed that enhanced group identification improves the effectiveness of communication in promoting cooperation (Tajfel, 1982; Kerr and Kaufman-Gilliland, 1994). Here, however, we limit our attention to how the effectiveness of informal agreements against opportunistic behavior (particularly, appropriation behavior) can be improved by group identity.

two-person group will then be matched with another newly randomly formed two-person group. Therefore, you will be matched with 10 different counterparts and your group will be matched with 10 different groups during the whole session.

{There are 10 experimental rounds in the following setting. Before the experiment, you will be randomly and anonymously matched with a counterpart who is another participant in this room. In each of the 10 rounds, your two-person group will be randomly matched with another two-person group in this room. Therefore, you will be matched with the same counterpart during the whole session but your group will be randomly matched with another group in each round.}

You will be making one decision in each of the 10 rounds [There are two stages in each of the 10 rounds. You will be making one decision in each stage]. You will have 60 s to make each of your decisions, which will be noted at the top of each screen. If you do not make a decision within 60 s, the computer will randomly choose a decision for you.

<Before the first stage in each round, you have 60 s to communicate with your counterpart through the chat window on the screen by selecting the sentences from a given list.>

Your earnings in each round will depend on the decision you make, the decision your counterpart makes, the decisions the participants in the matched group (which is called "counter-group" in the following instructions) make, and chance. If you follow the instructions carefully and make good decisions, you could earn a considerable amount of money.

There are two stages in each round. In the first stage all of the participants will simultaneously and independently choose an integral number between 0 and 20. At the end of each round, you will learn the choice of your counterpart but not the choices of any other participants. After all of the participants make their choices, the computer will randomly select an integral number between −40 and 40 for each two-person group. Each integral number between −40 and 40 is equally likely to be selected by the computer. Your payoffs in the first stage will depend on the numbers you and your counterpart choose, the numbers the participants in the counter-group choose, and the random numbers selected by the computer.

We will first use these numbers to calculate the output of your group:

$$\begin{aligned} Your\ group's\ output &= 20 * (your\ first - stage\ choice) + 20 * (counterpart's\ first - stage\ choice) \\ &\quad + your\ group's\ random\ number. \end{aligned}$$

We will use the same method to calculate the output of your counter-group (the two participants in the counter-group are called A and B):

$$\begin{aligned} Counter - group's\ output &= 20 * (A's\ first - stage\ choice) + 20 * (B's\ first - stage\ choice) \\ &\quad + counter - group's\ random\ number. \end{aligned}$$

Choosing higher numbers increases your group's output. However, choosing higher numbers is more costly. Your cost is:

$$Your\ cost = 100 * \ (your\ first - stage\ choice).$$

We will use the same method to calculate the costs of the other participants.

At the beginning of each round, you will receive 2000 experimental points as fixed earnings for each round.

In the first stage, if your group's output is larger than your counter-group's output, your group will receive a prize of 2000 experimental points. You and your counterpart will equally split this prize; thus each of you will receive 1000 experimental points. If your group's output is smaller than your counter-group's output, your group will receive no prize and this round then ends. If your group's output is equal to your counter-group's output, the computer will randomly select one group to receive the prize, whereas the other group will receive no prize.

Your first-stage earnings are:

$$Your\ first - stage\ earnings = 1000\ \ (if\ your\ group's\ output\ \ >\ \ counter - group's\ output).$$

If your group's output is larger than your counter-group's output, you and your counterpart will enter into the second stage.

Are there any questions before we continue? If so, please raise your hand.

Specific instructions for Low *[*High*] {*Zero*} treatment*

In the second stage, you and your counterpart do not need to make any choices. The computer will randomly select two integral numbers between −20 and 20; one for you and one for your counterpart. Each integral number between −20 and 20 is equally likely to be selected by the computer. Your second-stage output will depend on the number your counterpart chooses in the first stage, and your random number selected by the computer in the second stage. {Your second-stage output will depend on your random number selected by the computer in the second stage.}

We will use these numbers to calculate the second-stage outputs of you and your counterpart:

$$\begin{aligned} Your\ second - stage\ output &= 0.1[0.9] * 20 * (counterpart's\ first - stage\ choice) \\ &\quad + your\ second - stage\ random\ number. \end{aligned}$$

$$\{Your\ second - stage\ output = your\ second - stage\ random\ number.\}$$

$$Counterpart's\ second - stage\ output = 0.1[0.9] * 20 * (your\ first - stage\ choice) + counterpart's\ second - stage\ random\ number.$$

$$\{Counterpart's\ second - stage\ output = counterpart's\ second - stage\ random\ number.\}$$

In the second stage, if your output is larger than your counterpart's output, you will receive a prize of 18,000 experimental points. If your output is smaller than your counterpart's output, you will receive no prize in the second stage. If your output is equal to your counterpart's output, the computer will randomly select you or your counterpart to receive the prize, whereas the other person will receive no prize.

Your second-stage earnings are:

$$Your\ second - stage\ earnings = 18000\ (if\ your\ second - stage\ output > counterpart's\ second - stage\ output).$$

Therefore, choosing a higher number in the first stage increases your group's first-stage output and thus increases the probability that your group receives the first-stage prize and enters into the second stage. However, choosing a higher number increases your cost and your counterpart's second-stage output and thus decreases the probability that you receive the second-stage prize. {However, choosing a higher number increases your cost.}

Are there any questions before we continue? If so, please raise your hand.

Reminder:

You have been matched with a counterpart in this round that you have never been matched with before.

In the first stage, you and your counterpart each choose an integral number between 0 and 20. You and your counterpart will learn the choices of each other at the end of this round. The computer will randomly select an integral number between −40 and 40 for your group.

$$Your\ group's\ output = 20 * (your\ first - stage\ choice) + 20 * (counterpart's\ first - stage\ choice) + your\ group's\ random\ number.$$

Your group will be matched with another group newly formed in this round. The methods for calculating the counter-group's output, costs, and earnings are the same as yours.

$$Your\ first - stage\ earnings = 1000\ (if\ your\ group's\ output > counter - group's\ output).$$

If your group's output is larger than your counter-group's output, your group will enter into the second stage.

In the second stage, you and your counterpart do not need to make any choices. The computer will randomly select two integral numbers between −20 and 20; one for you and one for your counterpart.

$$Your\ second - stage\ output = 0.1[0.9] * 20 * (counterpart's\ first - stage\ choice) + your\ second - stage\ random\ number.$$

$$\{Your\ second - stage\ output = your\ second - stage\ random\ number.\}$$

$$Counterpart's\ second - stage\ output = 0.1[0.9] * 20 * (your\ first - stage\ choice) + counterpart'ssecond - stage\ random\ number.$$

$$\{Counterpart's\ second - stage\ output = counterpart's\ second - stage\ random\ number.\}$$

$$Your\ second - stage\ earnings = 18000\ (if\ your\ second - stage\ output > counterpart's\ second - stage\ output).$$

$$Your\ cost = 100 * (your\ first - stage\ choice).$$

$$Your\ earnings\ in\ this\ round = 2000\ (\text{fixed earnings in this round}) + Your\ first - stage\ earnings + Your\ second - stage\ earnings - Your\ cost.$$

In the next round, you will be matched with another counterpart that you have never been matched with before and your group will be matched with another group newly formed in the next round.

Please enter your first-stage choice from (0–20):________. Please enter your choice within 60 s; otherwise, the computer will randomly choose a number for you.

Specific instructions for treatments with selectable second-stage choice <treatments with communication> {treatments with fixed matching}

In the second stage, you and your counterpart each choose a number between two numbers: 0.1 and 0.9. The computer will randomly select two integral numbers between −20 and 20; one for you and one for your counterpart. Each integral number between −20 and 20 is equally likely to be selected by the computer. Your second-stage output will depend on the number your counterpart chooses in the first stage, the number you choose in the second stage, and your random number selected by the computer in the second stage.

We will use these numbers to calculate the second-stage outputs of you and your counterpart:

$$\begin{aligned} &Your\ second - stage\ output = your\ second - stage\ choice * 20 * (counterpart's\ first - stage\ choice) \\ &\quad + your\ second - stagerandom\ number. \end{aligned}$$

$$\begin{aligned} &Counterpart's\ second - stage\ output = counterpart's\ second - stage\ choice * 20 * (your\ first - stage\ choice) \\ &\quad + counterpart's\ second - stage\ random\ number. \end{aligned}$$

In the second stage, if your output is larger than your counterpart's output, you will receive a prize of 18,000 experimental points. If your output is smaller than your counterpart's output, you will receive no prize in the second stage. If your output is equal to your counterpart's output, the computer will randomly select you or your counterpart to receive the prize, whereas the other person will receive no prize.

Your second-stage earnings are:

$$Your\ second - stage\ earnings = 18000\ (if\ your\ second - stage\ output > counterpart's\ second - stage\ output).$$

Therefore, choosing a higher number in the first stage increases your group's first-stage output and thus increases the probability that your group receives the first-stage prize and enters into the second stage. However, choosing a higher number in the first stage increases your cost and your counterpart's second-stage output and thus decreases the probability that you receive the second-stage prize. Furthermore, choosing a higher number in the second stage increases your second-stage output and thus increases the probability that you receive the second-stage prize.

Are there any questions before we continue? If so, please raise your hand.

Reminder:

You have been matched with a counterpart in this round that you have never been matched with before {You have been matched with the same counterpart in this round}.

In the first stage, you and your counterpart each choose an integral number between 0 and 20. You and your counterpart will learn the choices of each other at the end of this round. The computer will randomly select an integral number between −40 and 40 for your group.

$$\begin{aligned} Your\ group's\ output\ =\ &20 * (your\ first - stage\ choice) + 20 * (counterpart's\ first - stage\ choice) \\ &+ your\ group's\ random\ number. \end{aligned}$$

Your group will be matched with another group newly formed in this round {Your group will be randomly matched with another group in this round}. The method of calculating the counter-group's output, costs, and earnings are the same as yours.

$$Your\ first - stage\ earnings = 1000\ (if\ your\ group's\ output > counter - group's\ output).$$

If your group's output is larger than your counter-group's output, your group will enter into the second stage.

In the second stage, you and your counterpart each choose a number between two numbers: 0.1 and 0.9. The computer will randomly select two integral numbers between −20 and 20; one for you and one for your counterpart.

$$\begin{aligned} Your\ second - stage\ output\ =\ &your\ second - stage\ choice * 20 * (counterpart's\ first - stage\ choice) \\ &+ your\ second - stage\ random\ number. \end{aligned}$$

$$\begin{aligned} &Counterpart's\ second - stage\ output = counterpart's\ second - stage\ choice * 20 * (your\ first - stage\ choice) \\ &\quad + counterpart's\ second - stage\ random\ number. \end{aligned}$$

$$Your\ second - stage\ earnings = 18000\ (if\ your\ second - stage\ output > counterpart's\ second - stage\ output).$$

$$Your\ cost = 100 * (your\ first - stage\ choice).$$

$$\begin{aligned} Your\ earnings\ in\ this\ round = &\ 2000(\text{fixed earnings in this round}) + Your\ first - stage\ earnings + Your\ second \\ &- stage\ earnings - Your\ cost. \end{aligned}$$

In the next round, you will be matched with another counterpart that you have never been matched with before and your group will be matched with another group newly formed in the next round {In the next round, you will be matched with the same counterpart and your group will be randomly matched with another group}.

<You have 60 s to communicate with your counterpart through the chat window on the screen by selecting the sentences from the following list:

Table B1
Second-stage payoff matrix of appropriation choice.

	Low	*High*
Low	(9000, 9000)	(0, 18,000)
High	(18,000, 0)	(9000, 9000)

"I will choose 0.1 in the second stage;"
"I suggest you to choose 0.1 in the second stage;"
"I agree;" and
"I disagree."

(The *subjects are asked to make the first-stage choice after 60 s of communication.*)>

Please enter your first-stage choice from (0 - 20):________. Please enter your choice within 60 s; otherwise, the computer will randomly choose a number for you.

(*If the group's output was smaller than the counter-group's output, then this round ends for the subject. If the group's output was larger than the counter-group's output, then the subject enters into the second stage of which the instructions are shown as follows.*)

Reminder:

In the first stage of this round, you chose ____, your counterpart chose ____, your cost is ____, your group's random number was ____, your group's output was ____, your counter-group's output was ____. Your group's output was larger than your counter-group's output. Thus you and your counterpart each receive 1000 experimental points and enter into the second stage.

In the second stage, you and your counterpart each choose a number between two numbers: 0.1 and 0.9. The computer will randomly select two integral numbers between −20 and 20; one for you and one for your counterpart.

$$Your\ second-stage\ output = your\ second-stage\ choice * 20 * (counterpart's\ first-stage\ choice) + your\ second-stage\ random\ number.$$

$$Counterpart's\ second-stage\ output = counterpart's\ second-stage\ choice * 20 * (your\ first-stage\ choice) + counterpart's\ second-stage\ random\ number.$$

$$Your\ second-stage\ earnings = 18000\ (if\ your\ second-stage\ output > counterpart's\ second-stage\ output).$$

$$Your\ earnings\ in\ this\ round = 2000(\text{fixed earnings in this round}) + Your\ first-stage\ earnings + Your\ second-stage\ earnings - Your\ cost.$$

In the next round, you will be matched with another counterpart that you have never been matched with before and your group will be matched with another group newly formed in the next round {In the next round, you will be matched with the same counterpart and your group will be randomly matched with another group}.

Please enter your second-stage choice between the two numbers 0.1 and 0.9:________. Please enter your choice within 60 s; otherwise, the computer will randomly choose a number for you.

Appendix B. The effects of informal agreements in the one-shot game and the finitely repeated game

When pre-play communication is possible, alliance members may try to reach an informal agreement to ensure their investments protected and not used against them in stage two. We assume that the partners reach an agreement that both will choose the low appropriation level in stage two.

B.1. One-shot game

In a one-shot game, if the two partners enter into the second stage, they have no concern for future interaction between them while making the appropriation choice. Their second-stage payoff matrix of appropriation choice is shown in Table B.1. If both members adhere to the agreement by choosing the low appropriation level or if both of them break the agreement by choosing the high appropriation level, the expected payoff for each member is 9000 (i.e., one half chance to win the second-stage prize 18,000). But if one member chooses the high appropriation level while the other chooses the low appropriation level, the former receives 18,000 and the latter receives 0. The unique Nash equilibrium is thus (high, high). Anticipating this, both members will not invest in the first stage as in a one-shot game without communication. We thus derive Hypothesis 2 in the text.

B.2. Finitely repeated game

In our setting of 10-round repeated interaction, the partners within an alliance are fixed but their competing alliances are different in each round. If the two partners enter into the second stage in a round, their reputation concerns for future interaction may affect their decisions about whether or not to comply with their agreement.

Following Kreps et al. (1982), suppose there are two types of members in the game: reciprocal and rational. A reciprocal member will comply with the agreement by choosing the low appropriation level in any given round unless his/her partner breaks the agreement by choosing the high appropriation level in the previous round. If his/her partner breaks the agreement in the previous round, the reciprocal member will retaliate by choosing the high appropriation level in the current round and all the subsequent rounds. In contrast, a rational member only aims at maximizing his/her own payoff. In the following analysis, we show that even a rational member may have an incentive to comply with the agreement in all but the last few rounds.

The member type is private information. We assume that a rational member *ik* holds a belief that with a probability δ_{im} his partner *im* is reciprocal and with a probability 1-δ_{im} member *im* is rational. Suppose that in round *n* the two partners *ik* and *im* enter into the second stage. Both members then need to make the decision whether or not to comply with the agreement. If member *ik* pretends to be reciprocal by complying with the agreement, his expected payoff from the second stage of round *n* to round 10 is at least

$$E\pi_{ik}(n, low) = \delta_{im}[9000 + (8000 - 5000q)(10 - n)] + (1 - \delta_{im})[5000r(10 - n)], \tag{B1}$$

where *q* represents the probability that both members in the competing alliance invest 20 in the first stage and *r* represents the probability that both members in the competing alliance invest 0 in the first stage. Thus 1-*q*-*r* represents the probability that one member in the competing alliance invests 20 but the other invests 0. Note that theoretically a member will invest 0 (20) in the first stage if he/she believes that his/her partner will choose the high (low) appropriation level in the second stage.

The first line in (B1) represents member *ik*'s expected payoff if member *im* is reciprocal. In that case, member *ik* obtains an expected payoff of 9000 (i.e., one half chance to win the prize 18,000) in the second stage of round *n*. Since member *ik* plays the reciprocal strategy by choosing the low appropriation level from round $n+1$ to round 10 unless the partner breaks the agreement, his expected payoff is

$$\begin{aligned}&\frac{1}{2}18,000 + \left[q\left(\frac{1}{2}\left(1000 + \frac{1}{2}18,000\right) - 2000\right) + (1 - q)\left(1\left(1000 + \frac{1}{2}18,000\right) - 2000\right)\right](10 - n)\\ &\quad = 9000 + (8000 - 5000q)(10 - n).\end{aligned}$$

Note that member *ik* invests 20 from round $n+1$ to round 10 given that a reciprocal partner always chooses the low appropriation level unless a retaliation is triggered. Since both members in alliance *i* invest 20 (and bear the investment cost, 2000), the competing alliance will share an equal chance to win the first stage if both of its members invest 20 (with the probability *q*). But if one member or both members in the competing alliance invests 0 (with the probability $1-q$), alliance *i* will win the first stage.

The second line in (B1) represents member *ik*'s expected payoff if member *im* is rational. If member *im* breaks the agreement by choosing the high appropriation level in round *n*, member *ik*'s second-stage expected payoff in round *n* is 0. Both members will then choose the high appropriation level and thus invest 0 from round $n+1$ to round 10. In each future round, alliance *i* will have one half chance to win the first stage if both members in the competing alliance invest 0 (with the probability *r*) but will lose if at least one member in the competing alliance invests 20 (with the probability $1 - r$). Member *ik*'s expected payoff is thus

$$0 + \left[r\left(\frac{1}{2}\left(1000 + \frac{1}{2}18,000\right)\right)\right](10 - n) = 5000r(10 - n).$$

But member *im* may also mimic a reciprocal member and comply with the agreement in round *n* to induce future within-alliance cooperation. In that case, member *ik*'s expected payoff is strictly improved. Thus by playing the reciprocal strategy member *ik*'s lowest expected payoff from the second stage of round *n* to round 10 is given by (B1).

On the other hand, if member *ik* reveals his rationality by choosing the high appropriation level in the second stage of round *n*, his expected payoff from the second stage of round *n* to round 10 is at most

$$E\pi_{ik}(n, high) = 18,000 + 5000r(10 - n). \tag{B2}$$

In the second stage of round *n*, member *ik* obtains 18,000 if member *im* is reciprocal or if she is rational but mimics a reciprocal member. However, member *ik* obtains an expected payoff of 9000 if member *im* is rational and breaks the agreement. In either case, both members will choose the high appropriation level and thus invest 0 from round $n+1$ to round 10. Thus by breaking the agreement member *ik*'s highest expected payoff from the second stage of round *n* to round 10 is given by (B2).

The comparison between (B1) and (B2) yields

$$E\pi_{ik}(n, low) - E\pi_{ik}(n, high) = \delta_{im}[9000 + (8000 - 5000(q + r))(10 - n)] - 18000. \tag{B3}$$

If δ_{im} is not too small and $q+r$ is not too large, there exists an $n^*_{ik}<10$ such that (B3) is positive for all $n \le n^*_{ik}$. The same logic holds for member *im* if she is rational and believes member *ik* is reciprocal with a probability δ_{ik}. There will exist an $n^*_{im}<10$ such that $E\pi_{im}(n, low) > E\pi_{im}(n, high)$ for all $n \le n^*_{im}$. If n^*_{ik} is smaller (larger) than n^*_{im}, member *ik* (member *im*) will be the first one to break the agreement. Both members will choose the high appropriation level henceforth.[24] Thus even if a member is rational, he/she may still have an incentive to comply with the agreement to induce future within-alliance cooperation in all but the last few rounds. As a result, the investment levels are higher than and the appropriation levels are lower than those in a repeated game without communication in which an agreement against appropriation is impossible. We thus derive Hypothesis 3 in the text.

References

Altmann, S., Falk, A., Wibral, M., 2012. Promotions and incentives: the case of multistage elimination tournaments. J. Labor Econ. 30 (1), 149–174.
Andreoni, J., 1995. Cooperation in public-goods experiments: kindness or confusion. Am. Econ. Rev. 85 (4), 891–904.
Battigalli, P., Dufwenberg, M., 2007. Guilt in games. Am. Econ. Rev. Pap. Proc. 97 (2), 170–176.
Bhattarai, K.D., 2017. What caused the left alliance's landslide victory in Nepal? The Diplomat Dec 19.
Buckler, J., 2003. Aegean Greece in the Fourth Century BC. Brill.
Buckley, N.J., Mestelman, S., Muller, R.A., Schott, S., Zhang, J., 2018. The effects of communication on the partnership solution to the commons. Environ. Resour. Econ. 70 (2), 363–380.
Burgen, S., Jones, S., 2017. All you need to know about the Catalonia election. The Guardian Dec 21.
Cason, T.N., Sheremeta, R.M., Zhang, J., 2012. Communication and efficiency in competitive coordination games. Games Econ. Behav. 76 (1), 26–43.
Charness, G., 2000. Self-serving cheap talk: a test of Aumann's conjecture. Games Econ. Behav. 33 (2), 177–194.
Charness, G., Rabin, M., 2002. Understanding social preference with simple tests. Q. J. Econ. 117 (3), 817–869.
Chen, Y., Li, S., 2009. Group identity and social preference. Am. Econ. Rev. 99 (1), 431–457.
Cialdini, R.B., Trost, M.R., Newsom, J.T., 1995. Preference for consistency: the development of a valid measure and the discovery of surprising behavioral implications. J. Pers. Soc. Psychol. 69 (2), 318–328.
Cooper, R., Dejong, D.V., Forsythe, R.E., Ross, T.W., 1989. Communication in the battle of the sexes game: some experimental results. RAND J. Econ. 20 (4), 568–587.
Crawford, V., 1998. A survey of experiments on communication via cheap talk. J. Econ. Theory 78 (2), 286–298.
Ellingsen, T., Johannesson, M., 2004. Promises, threats and fairness. Econ. J. 114 (495), 397–420.
Engel, C., Kleine, M., 2015. Who is afraid of pirates? An experiment on the deterrence of innovation by imitation. Res. Policy 44 (1), 20–33.
Erev, I., Bornstein, G., Galili, R., 1993. Constructive intergroup competition as a solution to the free rider problem: a field experiment. J. Exp. Soc. Psychol. 29 (6), 463–478.
Eriksson, T., Villeval, M.C., 2012. Respect and relational contracts. J. Econ. Behav. Organ. 81 (1), 286–298.
Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. Am. Econ. Rev. 90 (4), 980–994.
Fehr, E., Kremhelmer, S., Schmidt, K., 2008. Fairness and the optimal allocation of ownership rights. Econ. J. 118 (531), 1262–1284.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Friedman, D., Sunder, S., 1994. Experimental Methods: A Primer for Economists. Cambridge University Press, Cambridge.
Goeree, J., Holt, C., Palfrey, T., 2002. Quantal response equilibrium and overbidding in private-value auctions. J. Econom. Theory 104 (1), 247–272.
Gronberg, T., Andrew Luccasen III, R., Turocy, T., Van Huyck, J., 2012. Are tax-financed contributions to a public good completely crowded-out? Experimental evidence. J. Public Econ. 96 (7-8), 596–603.
Hatton, B., Parra, A., 2017. Catalonia election: pro-independence parties keep their majorities but what happens now? Independent, Dec 22.
Herring, G.C., 1969. Lend-lease to Russia and the origins of the Cold War, 1944-1945. J. Am. Hist. 56 (1), 93–114.
Houser, D., Kurzban, R., 2002. Revisiting kindness and confusion in public goods experiments. Am. Econ. Rev. 92 (4), 1062–1069.
Kagan, D., 1991. The Fall of the Athenian Empire. Cornell University Press.
Katz, M.L., Ordover, J.A., 1990. R&D cooperation and competition. In: Baily, M.N., Winston, C.M. (Eds.), Brookings Papers on Economic Activity: Microeconomics. Brookings Institution Press, Washington D.C.
Ke, C., Konrad, K.A., Morath, F., 2013. Brothers in arms – An experiment on the alliance puzzle. Games Econ. Behav. 77 (1), 61–76.
Ke, C., Konrad, K.A., Morath, F., 2015. Alliances in the shadow of conflict. Econ. Inq. 53 (2), 854–871.
Kerr, N.L., Kaufman-Gilliland, C.M., 1994. Communication, commitment and cooperation in social dilemmas. J. Pers. Soc. Psychol. 66 (3), 513–529.
Klein, B., Crawford, R., Alchian, A., 1978. Vertical integration, appropriable rents, and the competitive contracting process. J. Law Econ. 21 (2), 279–326.
Kreps, D.M., Milgrom, P., Roberts, J., Wilson, R., 1982. Rational cooperation in the finitely repeated prisoners' dilemma. J. Econom. Theory 27 (2), 245–252.
Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. J. Polit. Econ. 89 (5), 841–864.
Lim, W., Matros, A., Turocy, T., 2014. Bounded rationality and group size in Tullock contests: experimental evidence. J. Econ. Behav. Organ. 99 (1), 155–167.
Ludwig, S., Lünser, G.K., 2012. Observing your competitor – the role of effort information in two-stage tournaments. J. Econ. Psychol. 33 (1), 166–182.
Lundquist, T., Ellingsen, T., Gribbe, E., Johannesson, M., 2009. The aversion to lying. J. Econ. Behav. Organ. 70 (1-2), 81–92.
Mckelvey, R., Palfrey, T., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10 (1), 6–38.
Meiggs, R., 1979. The Athenian Empire. Oxford University Press.
Minkova, K.V., 2018. Lend-Lease in early post-war Soviet-American relations. History 63 (2), 614–635.
Morita, H., Servátka, M., 2013. Group identity and relation-specific investment: an experimental investigation. Eur. Econ. Rev. 58 (1), 95–109.
Olson, Mancur, Zeckhauser, Richard, 1966. An economic theory of alliances. Rev. Econ. Stat. 48 (3), 266–279.
Palfrey, T.R., Rosenthal, H., 1994. Repeated play, cooperation and coordination: an experimental study. Rev. Econ. Stud. 61 (3), 545–565.
Pan, X., Houser, D., 2017. Social approval, competition and cooperation. Exp. Econ. 20 (2), 309–332.
Rege, M., Telle, K., 2004. The impact of social approval and framing on cooperation in public good situations. J. Public Econ. 88 (7-8), 1625–1644.
Seidmann, D.J., 1990. Effective cheap talk with conflicting interests. J. Econom. Theory 50 (2), 445–458.
Sheremeta, R.M., 2013. Overbidding and heterogeneous behavior in contest experiments. J. Econ. Surv. Spec. Issue 27 (3), 491–514.
Sonnemans, J., Oosterbeek, H., Sloof, R., 2001. On the relation between asset ownership and specific investments. Econ. J. 111 (474), 791–820.
Sutter, M., Strassmair, C., 2009. Communication, cooperation and collusion in team tournaments: an experimental study. Games Econ. Behav. 66 (1), 506–525.

[24] Note that member *ik* cannot update his belief δ_{im} if member *im*, even being rational, mimics a reciprocal member. Member *ik* only updates δ_{im} to be zero when member *im* first breaks the agreement. The same logic holds for member *im*'s belief.

Tajfel, H., 1982. Social psychology of intergroup relations. Annu. Rev. Psychol. 33 (1), 1–39.
Tajfel, H., Turner, J., 1979. An integrative theory of intergroup conflict. In: Worchel, S., Austin, W. (Eds.), The Social Psychology of Intergroup Relations. Brooks/Cole Press, California.
Weeks, A.L., 2010. Russia's Life-Saver: Lend-Lease Aid to the U.S.S.R. in World War II. Lexington Books, United Kingdom.
Williamson, O., 1985. The Economic Institutions of Free Capitalism. Free Press, New York.