Alliances in the Shadow of Conflict[*]

Changxia Ke[†] Kai A. Konrad[‡] Florian Morath[§]

July 1, 2014

Abstract

Victorious alliances often fight about the spoils of war. This paper presents an experiment on the determinants of whether alliances break up and fight internally after having defeated a joint enemy. First, if peaceful sharing yields an asymmetric rent distribution, this increases the likelihood of fighting. In turn, anticipation of the higher likelihood of internal fight reduces the alliance's ability to succeed against the outside enemy. Second, the option to make non-binding non-aggression declarations between alliance members does not make peaceful settlement within the alliance more likely. Third, higher differences in the alliance players' contributions to alliance effort lead to more internal conflict and more intense fighting.

Keywords: Conflict; Contest; Alliance; Endogenous internal conflict; Hold-up problem; Non-aggression pact; Experiment

JEL Codes: D72; D74

[*]We thank Klaus Abbink, Jordi Brandts, Subhasish Chowdhury, Werner Güth, Luisa Herbst, Henrik Orzen, Alex White, participants in the Workshop on "Alliances and Alliance Formation in Conflicts" in Berlin (2011), the SFB meeting in Mannheim (2012), the ESA European Conference (2012) in Cologne, the CESifo Area Conference on Behavioural Economics in Munich (2012), the Workshop on "Contests: Experiments and Theory" in Munich (2013) and seminar participants at Tsinghua University, University of Macao, and University of Augsburg for helpful comments and suggestions. Moreover, we thank MELESSA of the University of Munich for providing laboratory resources and Nina Bonge, Bernhard Enzi, and Verena Hefner for research assistance. Financial support from the German Research Foundation (DFG, grant no. SFB-TR-15) is gratefully acknowledged.

[†]Queensland University of Technology, School of Economics and Finance, Business School. changxia.ke@qut.edu.au.

[‡]Max Planck Institute for Tax Law and Public Finance, Munich, and Social Science Research Center Berlin. kai.konrad@tax.mpg.de.

[§]Corresponding author. Max Planck Institute for Tax Law and Public Finance, Department of Public Economics, Marstallplatz 1, 80539 Munich, Germany. Phone: +49 89 24246-5254. Fax: +49 89 24246-5299. E-mail: florian.morath@tax.mpg.de.

1 Introduction

Members of alliances who have jointly defeated an opponent must decide how to divide the prize of victory. Such distributional conflict may be resolved peacefully or may involve resource-wasteful fighting. This paper studies how players of a successful alliance cope with the distributional conflict within the alliance. We explore the behavior of members of a victorious alliance when they decide whether to split peacefully or to enter into a fight, and we trace the determinants of this choice behavior.

The question of whether and when alliances break up fighting over the spoils of victory has attracted considerable interest in economics and political science.[1] Perhaps the most illustrative examples are in the context of military conflict. Political history provides many examples of alliance members deciding to turn against each other as soon as the goal of the alliance has been accomplished (Beilenson 1969, O'Connor 1969, Bunselmeyer 1975). Consider, for instance, the Hitler-Stalin alliance against Poland that led to the invasion of Poland and its division between Nazi Germany and the Soviet Union on the brink of the Second World War. The two alliance members kept peace between them for quite some time, but later Germany attacked the Soviet Union.[2] Other important examples are coalitions of different groups that ally to defeat the incumbent ruler. Once this goal is achieved, they may share the power in a peaceful democratic regime, or enter into a struggle for supreme power. The problem of sharing a prize that has been won jointly also occurs in politics when several parties or politicians team up in an effort to win an election or come into power by other means, and then have to decide whether to share power peacefully or to enter into a dispute (or costly bargaining process) over the division of power. The two Roman triumvirates are legendary historical examples; in both triumvirates, the members of the alliance turned against each other not long after jointly reaching power. Many more examples can be found in other countries and historical episodes.[3]

[1] See Kimball (2006) and Johnson and Leeds (2011) for recent surveys and further empirical analyses. Formal model analysis on alliances include, for instance, Morrow (1991, 1994, 2000) and Niou and Ordeshook (1994). Mattes and Vonnahme (2010) analyze non-aggression pacts as a special type of alliance.

[2] Stalin's concerns about the stability of the alliance between Russia and Germany in the months prior to 'Operation Barbarossa' in June 1941, despite British warnings, are discussed by Reynolds (2002).

[3] Similar problems also emerge where firms may team up and form a research joint venture in a patent contest, but then have to decide how to position themselves when marketing the product they have jointly innovated.

Violent conflict between former alliance partners may be a frequent outcome, but it is not an automatism. The break-up of an alliance after victory may be avoidable and is, to a large extent, a matter of choice. In the examples above, former alliance members could typically decide whether or not to enter into a fight with their former ally. Violent conflict is what political scientists have tried to explain more generally, given that peaceful sharing is seemingly the more natural and often Pareto superior outcome. Several reasons for why countries may end up in a violent fight rather than find a less wasteful negotiation outcome are outlined in the survey by Jackson and Morelli (2011).[4] A fundamental question concerns the relation between balance of power and the emergence of conflict (Organski 1958, Claude 1962, Blainey 1988, Wagner 1994); some theories suggest that the distribution of power matters for the allocation outcome, but does not have a strong impact on the probability of fighting (Wittman 1979). Alliances are considered to play an important role for the probability of resource-wasteful fighting.[5] Some alliances such as those based on non-aggression pacts seemingly try to avoid violent conflict among the signatory countries.

Empirical work on the interaction between alliances and the resolution of conflict suffers from severe endogeneity problems. The existence of an alliance or the signing of a non-aggression pact between two countries is typically not an exogenous event, but a consequence of the specific conflict, which makes causal inference very difficult.[6] Our experimental analysis can cope with these problems and allows causal inference. We consider distributional conflict among players who have jointly achieved a common goal in an alliance formed under the rules of the game and as an exogenous event. By imposing conflict with an outgroup, we analyze a victorious alliance's ability to avoid resource-wasteful internal fight about rent distribution. In the experiment, if the alliance wins against a joint adversary, the alliance members earn a prize of victory but need to determine how to share this prize, and they are given two options: They can either split the prize according to an exogenous rule or decide to break up and fight about the prize. Using several

[4] These include incomplete information (see Fearon 1995 for a detailed account), commitment problems (e.g., Garfinkel and Skaperdas 2000, Beviá and Corchón 2010), and equilibrium selection in frameworks with multiple equilibria (e.g., Slantchev 2003, and Konrad and Leininger 2011).

[5] Studies such as Levy (1981) highlight the diversity of alliances.

[6] See Levy (1981) for a careful consideration and Mattes and Vonnahme (2010) for an empirical assessment of the relationship between alliances and fighting and a discussion of the endogeneity issues.

treatments we can test for the importance of possible determinants of the choice of whether to fight and of the fighting intensity in case a distributional conflict takes place.

The first key distinction is about whether a member of a victorious alliance is disadvantaged by the share he receives in case of a peaceful settlement or whether peaceful sharing results in an equal distribution of the prize. The relation between potential asymmetries in the distribution of power and the resource allocation can affect the players' willingness to settle peacefully (Wittman 1979). We implement such asymmetries by allocating unequal peaceful shares to former alliance members who are of equal strength in an upcoming distributional conflict, and we study the effect of such imbalance on the emergence of conflict. The observed behavior is in line with the theoretical predictions in Konrad and Morath (2012a, 2012b): Disadvantaged players trigger distributional conflict even if this reduces their expected material payoff compared to the peaceful outcome. Moreover, alliance members correctly anticipate a higher likelihood of internal conflict and reduce their effort when fighting jointly against a common enemy.[7]

Second, we test whether non-binding declarations on peaceful intra-alliance sharing ("non-aggression treaties") can help to solve the distributional conflict inside an alliance. The study of the effectiveness of treaties by which conflict parties mutually declare to abstain from military conflict has some tradition in political science but comes to inconclusive results.[8] We consider two further treatments in which the alliance members must make a declaration about whether they intend to split the alliance's prize of victory peacefully or to fight inside the alliance. This declaration is made prior to the fight between the alliance and the outgroup, but it is non-binding. In one of the treatments, the declaration is made secretly to the operators of the experiment, but is not observed by the co-players. In the other treatment, the declaration becomes public information. In both treatments the players can later freely choose between the option to split peacefully or to fight. There is no monetary cost or disadvantage from not sticking to the initial

[7] This second result also reconfirms results in an earlier paper (Ke et al. 2013) in which victorious alliances were forced into conflict. There, the exogenously imposed internal conflict caused a hold-up problem when alliance members chose their contributions to alliance effort. In the theory of contests, this hold-up problem has been emphasized and studied by Katz and Tokadlidu (1996), Wärneryd (1998), and Esteban and Sákovics (2003).

[8] For recent contributions see, for instance, Mattes (2008), Leeds and Savun (2007), Long et al. (2007), and Mattes and Vonnahme (2010). Mattes and Vonnahme (2010) attribute the potential effectiveness of non-aggression pacts to the increase of reputational cost of aggression.

declaration, but players may feel some mental or reputational cost. In case the declarations are made public, we find that most alliance members intend to make use of the ex ante (non-binding) declarations of non-aggression, but this does not help to reduce the likelihood of internal conflict. The treatment with unobserved declarations reveals information on the players' true prize sharing intentions and on factors that made them deviate from their initial declaration.

Finally, we find that behavior inside the alliance is important for alliance players' choices of conflict and the intensity of such internal conflict: A higher asymmetry in the alliance players' effort contributions when fighting the joint enemy makes it more likely that former alliance members oppose the peaceful split and leads to more effort expended in the subsequent internal distributional conflict. Given that, in our setting, alliance players face a joint history when deciding about peaceful settlement, this reveals additional information on determinants of the emergence of conflict by showing that the (relative) effort contributions to the conflict with the outgroup crucially influence behavior in the subsequent stages of the game.

Apart from the literature already mentioned above, this research is related to several further strands of the literature. We use simple Tullock (1980) lottery contests as a generic description of conflict.[9] This type of strategic interaction has attracted considerable attention among theorists and has been used by experimental economists and psychologists.[10] A small subset of this literature also considers experiments on group contests or collective action problems with contest elements (e.g., Bornstein et al. 2002, Parco et al. 2005, Gunnthorsdottir and Rapoport 2006, Amegashie et al. 2007, Abbink et al. 2010, Kugler et al. 2010, Sheremeta 2010, Ahn et al. 2011, Cason et al. 2012, and Ke et al. 2013).[11] These contributions do not consider an endogenous choice about internal fighting and the interplay of this decision with the performance of the alliance in a conflict with an outside player or outside group.[12] Morgan et al. (2012) were among

[9] See Skaperdas (1996) for an axiomatization and Konrad (2009) for a survey on several microfoundations.

[10] This research program confirmed that players generally expend more effort in the lottery contest than what would be expected from maximization of monetary payoffs; see Millner and Pratt (1989, 1991), Davis and Reilly (1998), Potters et al. (1998), Anderson and Stafford (2003), Sheremeta (2010, 2011), Price and Sheremeta (2011). Explanations for this overdissipation range from non-monetary utility of winning (Sheremeta 2010, 2011; Chen et al. 2011), spite or inequality aversion (Abbink et al. 2012, Balafoutas et al. 2012) and mistakes (Schmidt et al. 2013, Lim et al. 2014) to judgemental biases (Amaldoss and Rapoport 2009, Sheremeta and Zhang 2010) and free-endowment (in-house money) effects (Price and Sheremeta 2014).

[11] For a comprehensive recent review of contest experiments see Dechenaux et al. (2012).

[12] An emerging strand of literature studies the difference between individual and group decision-making in both

the first to consider experiments on conflict with endogenous participation. In their framework more than two players could enter and fight, and whether entry caused an expected monetary payoff higher than the default payment depends on the number of entrants. Abbink and Brandts (2009) consider the emergence of conflict between groups where the resource distribution in case of peace is proposed by one of the groups, but they do not consider distributional conflicts within successful groups and the endogenous break-up of groups. Lacomba et al. (2014) also focus on endogenous resource allocations and allow the defeated player to destroy (part of) the resources to be transferred to the winner, which serves as a means of avoiding costly conflict. The experimental work that is most closely related to our work is by McBride and Skaperdas (2009). They also consider endogenous fighting decisions but analyze a dynamic, possibly infinitely repeated conflict which takes place between two players only. Fighting today that ends with the defeat of one of the two players is a way to eliminate any potential for future conflict, and conflict today therefore becomes more likely when the future becomes more important.[13] Our main new findings are on endogenous break-up of coalitions where alliance players choose endogenously whether to fight among themselves or share peacefully, and on the role of asymmetries in peaceful division rules and of non-aggression agreements in this context. Apart from the evident role of alliances in conflict, our setup offers additional insight on specific factors that influence a player's decision to trigger resource-wasteful conflict.

2 Theoretical framework and experiments

2.1 Theoretical framework

Base treatment. The basic experimental design implements a framework that builds on the alliance paradox analyzed by Esteban and Sákovics (2003). Two alliance players (called A and

non-competitive and competitive settings (e.g., Bornstein 2003, Charness et al. 2007, Chen and Li 2009, Sutter 2009). Our experiment considers alliance members' individual (rather than collective) decisions in the pursuit of a joint cause. This feature applies to a wide array of contests.

[13] This effect is based on Garfinkel and Skaperdas (2000), and there is a larger literature related to the considerations of bargaining in the shadow of possible future conflict. The threat-point in bargaining problems is often a resource-wasteful fight between the negotiating players. For a survey on this issue see Fearon (1995); further important contributions to this question are Skaperdas and Syropoulos (1996) and Anbarci et al. (2002).

B) are in a contest with a stand-alone player (called C) for a monetary prize of value $v = 450$. The game consists of three stages.

In stage 1, the alliance of A and B fights against C in a contest that follows the rules of a standard Tullock (1980) lottery contest. All players choose independently and simultaneously an amount of effort $x_i \geq 0$, $i \in \{A, B, C\}$. A player's cost of effort is normalized to be equal to the effort itself, and it cannot be recovered, regardless of whether or not a player wins the contest. The vector (x_A, x_B, x_C) of chosen efforts is publicly observed, and a random device determines whether the alliance AB or the stand-alone player C wins the prize. The probability for AB to win this lottery contest is equal to

$$p_{AB} = \frac{x_A + x_B}{x_A + x_B + x_C} \tag{1}$$

if at least one of the effort components is strictly positive, and equal to 1/2 if all three players expend zero effort. The probability that C wins is equal to $p_C = 1 - p_{AB}$. If C wins, C obtains the full prize and the game ends, with monetary payoffs $\pi_C = v - x_C$, $\pi_A = -x_A$, and $\pi_B = -x_B$.

If the alliance of players A and B wins against C, the players enter into stage 2, where A and B are asked to independently and simultaneously choose between an equal split of the prize (A and B each obtain $v/2$) and a contest for the entire prize value. If both players choose the equal split ("split"), then the game ends. The payoffs in this case are $\pi_A = v/2 - x_A$, $\pi_B = v/2 - x_B$, and $\pi_C = -x_C$. If both players A and B choose the contest ("fight"), then the game enters into stage 3 for sure. If one player chooses "fight" and the other player chooses "split", then, with probability $1 - \varepsilon$, the game enters into stage 3, and with probability ε the prize is split peacefully and equally between A and B.[14]

[14] Although our experiment aims at explaining distributional conflict in situations where one player alone can trigger this conflict, we leave a small probability ε (in the experiment, $\varepsilon = 0.1$) that the peaceful split is implemented in case of diverging votes. This is done in order to incentivize individual choices and to eliminate the (trivial) equilibrium in case of $\varepsilon = 0$ in which both players choose "fight", simply because they expect the other player to choose "fight" and expect to have no influence on the outcome. In the three-stage game considered, multiplicity of the subgame equilibrium would be particularly problematic for the interpretation of the results (making stage 1 choices depending on beliefs about the selected subgame equilibrium and allowing for finite horizon trigger strategies). We can only speculate if the ε probability affects individual behavior, but with $\varepsilon > 0$, we expect the observed voting choices to be more informative about the individual preferences, and we keep this design choice constant across all treatments.

In stage 3 (if reached), the alliance players' decisions in stage 2 are made common knowledge, and the two players A and B enter into a Tullock lottery contest. Each of them must independently choose an effort $y_i \geq 0$, $i \in \{A, B\}$; cost of effort is again equal to the effort itself and must be paid independently of the contest outcome. The prize of $v = 450$ is allocated to A and B, respectively, with probabilities

$$q_A = \frac{y_A}{y_A + y_B} \text{ and } q_B = \frac{y_B}{y_A + y_B} \tag{2}$$

if at least one of these efforts is strictly positive, and with probabilities equal to $1/2$ if $y_A = y_B = 0$. Hence, if, for instance, player A wins in stage 3, the monetary payoffs are $\pi_A = v - x_A - y_A$, $\pi_B = -x_B - y_B$, and $\pi_C = -x_C$.

The subgame-perfect Nash equilibrium of this game follows by backward induction. Stage 3 of the game is equivalent to a simple two-player lottery contest with prize value v; as it is well known, equilibrium effort is

$$y_A^* = y_B^* = v/4 = 112.5. \tag{3}$$

Since player $i \in \{A, B\}$ wins the full prize with probability $1/2$, his material payoff in this subgame is $v/2 - y_i^* = v/4$ (minus the cost of stage 1 effort x_i). In comparison, if A and B split equally, each player obtains a payoff of $v/2$ (again minus the cost of stage 1 effort). Thus, in stage 2, both players A and B will choose the peaceful arrangement: there will never be a violent breakup of the alliance if alliance players maximize their monetary payoff.[15]

In stage 1, the value of winning the contest between the alliance AB and the stand-alone player C is equal to $v = 450$ for player C, and it is equal to $v/2 = 225$ for each player A or B, given subgame-perfect play that involves a choice of the peaceful split. Hence, player $i \in \{A, B\}$ chooses x_i to maximize $\pi_i = p_{AB}(v/2) - x_i$, and C chooses x_C to maximize $\pi_C = (1 - p_{AB})v - x_C$.

[15] This equilibrium outcome is robust to a change in the timing of the game. If the order of stages 1 and 2 is reversed (alliance players can make binding prize sharing decisions before the contest with player C takes place), then the equilibrium again involves both alliance players choosing the peaceful arrangement, as well as effort choices as in (4). Note, however, that the (off-equilibrium) efforts $(x_A + x_B, x_C)$ in the contest subgame that would follow a commitment to internal conflict would be lower than those in (4). For a comparison of contest efforts in case of exogenously given internal conflict compared to exogenous peaceful sharing see also Ke et al. (2013).

This results in equilibrium effort contributions of

$$(x_A + x_B)^* = \frac{v}{9} = 50 \text{ and } x_C^* = \frac{2v}{9} = 100. \tag{4}$$

Here, a few points are worth mentioning. First, in equilibrium only the sum of the alliance players' effort is uniquely determined.[16] Second, the alliance players jointly expend much less effort than the stand-alone player, due to the lower share of the prize that each alliance player can win. Third, if, for some reason, internal conflict will be chosen with positive probability, this further reduces an alliance player's expected valuation of winning the fight against the stand-alone player; consequently, in this case, alliance effort $x_A + x_B$ should be even lower.[17]

This three-stage game constitutes the "BASE" treatment in our experiments. Figure 1 summarizes the sequence of actions in the BASE treatment (for the exact experimental procedures see below). Three further treatments each vary the base treatment along exactly one dimension.

Unequal peaceful sharing. The first treatment variation addresses the effect of an imbalance between the peaceful resource allocation and the distribution of power. This imbalance is generated by asymmetries in the shares of the prize that members of a victorious alliance would obtain in case of peace. We exogenously impose that one of the players (A or B) would receive a share of 70% of the prize and the other would receive only 30% of the prize. The fact that the peaceful split involves unequal shares is made common knowledge at the beginning of the game, but which of the players (A or B) would get the larger share is randomly decided only at the beginning of stage 2 (in case the alliance won against player C). Each of the players A and B has the same chance of being the player with the larger share. In this way, the alliance players face symmetric incentives in stage 1, and a player's expected prize value in case of the

[16] Since the probability p_{AB} only depends on the sum of x_A and x_B, player $i \in \{A, B\}$'s *marginal* payoff from increasing x_i depends on $x_A + x_B$ but is independent of the individual efforts x_A and x_B. See Nitzan (1991) for a more detailed equilibrium analysis.

[17] The asymmetric structure of a conflict between an alliance and a single player has the important advantage that there is only a 'one-sided' incentive problem: alliance players do not have to form beliefs about the likelihood of internal conflict in the outgroup and the induced hold-up problem, which considerably simplifies the strategic interaction and the identification of treatment effects.

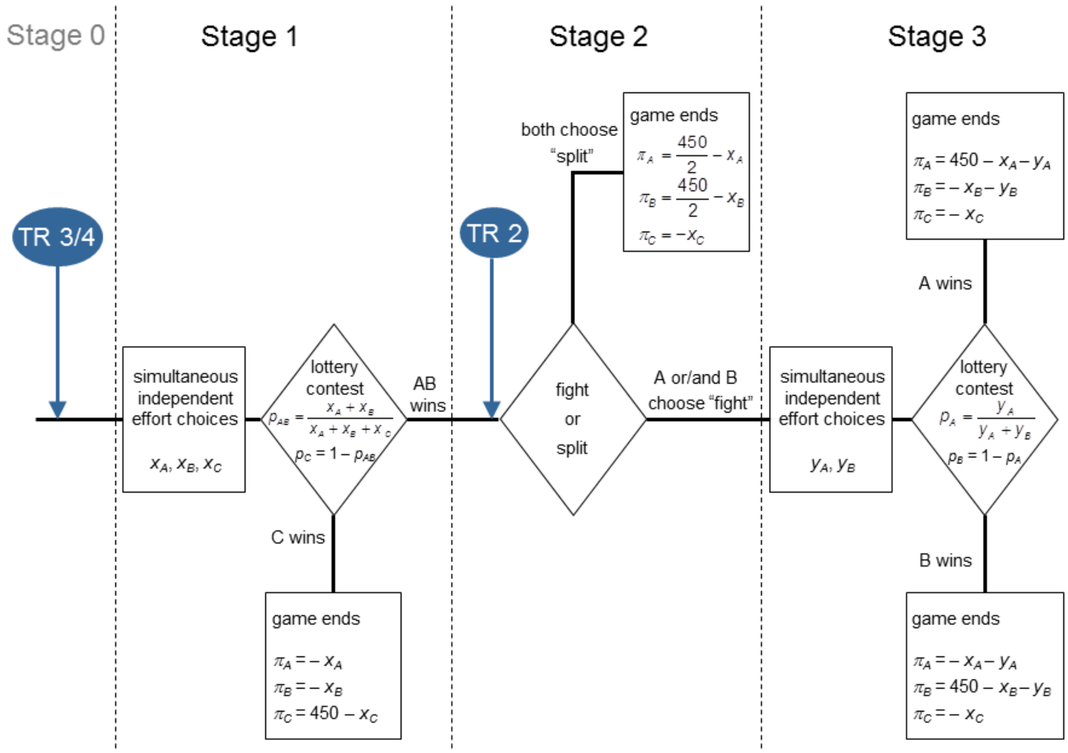

Figure 1: Sequence of actions in the BASE treatment.

peaceful split is exactly the same as in the BASE treatment.[18] Once the game reaches stage 2, players learn who receives the small and who the large share in the prize in case of a peaceful split; then, players A and B simultaneously choose between the two options "split" and "fight", and the game continues as in the BASE treatment above.

A player A or B's expected payoff in the subgame with internal distributional conflict in stage 3 is still equal to $v/4$ (minus stage 1 effort cost) in this treatment. Hence, a player who cares about monetary rewards only strictly prefers the peaceful settlement if and only if his peaceful share of the prize is larger than 25%. Therefore, for players who maximize their monetary payoff, the equilibrium prediction for the choice in stage 2 also remains unchanged: If the alliance defeats the outgroup player, both A and B should choose the peaceful split, independently of whether it is revealed that they get the large or the small share. Moreover, effort choices in stage 1 should be exactly as in the BASE treatment (alliance players still maximize an expected payoff

[18] This aspect is important in order to be able to interpret possible behavioral reactions with respect to stage 1 effort, and ex ante uncertainty about the exact shares in case of a peaceful arrangement is often not implausible to assume.

of $p_{AB}(v/2) - x_i$. If, however, (some) players are motivated by aspects other than pure monetary rewards, an asymmetric split of the prize may affect the players' choices and the likelihood of conflict, as we will discuss later on. The marker TR2 in Figure 1 illustrates this first modification, which results in the second treatment called the "UNEQUAL" treatment.

Ex ante declarations. Forming alliances often involves ex ante agreements about how to divide the spoils of victory. Formal unilateral or mutual declarations of non-aggression are non-binding in an international context and should therefore have no significant causal impact on decisions about war and peace.[19] This view is contested by other theories in political science.

Do non-aggression declarations affect players' choices on how to divide the prize and help alliances to dissolve peacefully? Two further treatments can shed light on this question. There, we add a "stage 0" to the BASE treatment. In this stage, which occurs prior to the choices of efforts (x_A, x_B, x_C) in stage 1, each of the alliance players A and B must make a declaration about whether he intends to fight or to choose the peaceful and equal split of the prize in case their alliance wins against the outgroup player C. This declaration is non-binding (i.e., it can be reversed without direct cost later when the actual decision comes up); moreover, it is restricted to a simple indication of the option that the player intends to choose ("split" or "fight"). Should the game reach the decision stage (stage 2), players A and B can make their actual choice between peaceful sharing or fighting fully independently of their previous declarations. Apart from these declarations in stage 0, all other aspects of the game remain exactly as in the BASE treatment (in particular, the equal shares in case of a peaceful split).[20]

In the first of these two treatments, called "PRIVATE", a player's ex ante declaration is not displayed to the public and hence not observed by other players. It may, however, convey information about the players' true intentions of how to divide the prize, at a point before the conflict with the stand-alone player has taken place. In the second treatment with declarations, called "PUBLIC", the alliance players' ex ante declarations become public information before the

[19] The "realist" school in political science considers the causal relationship between institutions and peace as weak or non-existent; for an outline see Mearsheimer (1994-1995).

[20] In Figure 1, an appropriate adjustment would be to add this declaration stage (stage 0) for players A and B right before their actual choices of stage 1 efforts (indicated by the marker TR3/4).

players enter into stage 1, and this public nature of the declaration is known to all players. As in both the PRIVATE and PUBLIC treatments the declarations are fully non-binding, standard economic theory predicts that they do not affect the subsequent equilibrium play; hence, effort choices and choices of prize sharing should be exactly as in the BASE treatment. In the PUBLIC treatment, however, the declarations may cause behavioral reactions of the co-players: They may affect the effort contributions in stage 1 as well as choices of how to divide the prize, as we will discuss after presenting the details of the experimental procedures. Comparing the declarations in the PUBLIC treatment to the PRIVATE treatment and comparing actual choices of "fight" to initial declarations allows us to analyze when, and because of which factors, alliance members deviate from their initial prize sharing intentions.

2.2 Experimental procedures

The experiment was programmed and conducted with software z-Tree (Fischbacher 2007) and carried out at the University of Munich. The subjects were recruited from the laboratory's subject pool using ORSEE (Greiner 2004) and included students from all fields. The total number of subjects was 282; Table 1 provides information about the sessions that we ran. The experiment is based on a between-subjects design. Each subject participated in 24 rounds of the same treatment (exactly one of the four treatments BASE, UNEQUAL, PRIVATE, and PUBLIC) and kept his individual role as alliance player (A or B) or as stand-alone player (C) throughout all rounds. Anonymity was preserved during the experiment, and payments were made in private. In each session, students were divided into subgroups and randomly rematched within these, in order to eliminate quasi-repeated games effects. The instructions (see appendix) were given to them and read to them by the laboratory staff, and, in addition, an entry quiz guided them through the experiment.

The prize value v was 450 tokens (at a conversion rate of 45 tokens = EUR 1). Individual effort had to be chosen as a number from the set $\{0, 1, 2, ..., 250\}$ (in the contest between the alliance and player C as well as in the potential internal fight between A and B).[21] The subjects

[21] The upper bound of 250, which is clearly above the theory prediction, was chosen to avoid bankruptcy of

Treatment	No. of sessions	No. of subjects	Average earnings	Male subjects	Average age	Economics students
BASE	3	72	EUR 24.8	40%	23	19%
UNEQUAL	3	66*	EUR 24.3	41%	23	17%
PRIVATE	3	72	EUR 24.3	51%	24	25%
PUBLIC	3	72	EUR 24.8	43%	23	11%
Aggregate:	12	282	EUR 24.6	44%	23	18%

*In one session of the UNEQUAL treatment, the number of subjects was only 18 (instead of 24).

Table 1: Summary of the experimental sessions (between-subjects design).

experienced the probabilistic nature of the outcome via a 'fortune wheel' on the computer screen. The fortune wheel is a disc that has two segments in different colors, where the size of the segments is proportional to the relative amounts of efforts $x_A + x_B$, and x_C, respectively. A pointer spins clockwise and then stops in one of the segments, and victory is attributed to AB or to C, depending on the segment in which the pointer comes to a rest.[22]

Apart from a show-up fee of EUR 4, subjects received a fixed payment of EUR 0.6 for each of the 24 independent rounds (which basically served as their total endowment). At the end of the experiment, subjects were paid according to their decisions and outcomes in 6 randomly selected rounds. Positive profit was added to and negative profit was deducted from the fixed payment. Average earnings per subject were EUR 24.6 in total.[23] Before ending the session, subjects were asked to answer an exit questionnaire. The time for a session was very similar across the treatments (roughly 1.5 hours).

subjects. As shown in Section 4, there are some choices at this boundary in the internal fight (compare Table 2). While this upper bound may make the *level* of stage 3 effort less informative, we do not expect it to cause a systematic bias across treatments.

[22]Similarly, in the distributional conflict between A and B, a fortune wheel was used with two segments representing each player's relative effort in stage 3.

[23]The minimum total payment was the show-up fee and the maximum payment was EUR 61.

3 Main hypotheses

We now formulate testable hypotheses on treatment effects with respect to the players' choices. The highlighted hypotheses concentrate on our main question on the emergence of internal conflict (stage 2), but we also discuss the implications for contest effort in stage 1 and stage 3.

In the absence of motivations other than maximizing own monetary payoff, all four treatments have the same subgame-perfect Nash equilibrium with respect to choices in all three stages. The first testable hypothesis is therefore:

Hypothesis 1: *The share of alliance players who choose internal conflict is the same for all treatments and equal to zero.*

As it is well known from other contest experiments, individuals exhibit quite some heterogeneity and follow other motives besides pure maximization of monetary payoffs. Although individuals do not have a monetary incentive to choose internal conflict, we suspect that, in all treatments, a significant share of alliance players prefers to fight internally, in contrast to Hypothesis 1. Moreover, the treatment variations could affect individual choices of "split" versus "fight". This leads to two alternative hypotheses on the treatment effects, drawing on evolutionary stability arguments and on arguments that have been developed on non-aggression treaties in political science.

Consider first the effect of an unequal peaceful split of the prize (the UNEQUAL treatment). If players maximize their material payoffs only, they choose the peaceful split in equilibrium even if they obtain only the smaller share of the prize (30%). If, however, players' behavior is shaped by evolutionary forces in the context of a finite population (as introduced by Schaffer 1988), relative rather than absolute material payoff of a player determines the evolutionary success of this player. In the context of contest theory this has two implications. First, evolutionarily stable strategies typically involve higher fighting effort than in the Nash equilibrium, a result which is in line with most of the experimental evidence. Second, players choose conflict as part of the evolutionarily stable strategy if their peaceful share is too small relative to others' shares, even if their expected material payoff from fighting is lower than their material payoff in the peaceful sharing regime.[24] This yields an alternative hypothesis on the treatment effect of UNEQUAL.

Hypothesis A1: *(i) The share of alliance players who choose internal conflict is lowest among players who would receive* 70 *percent of the prize in case of a peaceful split, larger for players in the BASE treatment (with an equal peaceful split), and highest for players who would receive only* 30 *percent of the prize. (ii) Overall, there is more internal conflict in the UNEQUAL treatment than in the BASE treatment.*

If players with a small peaceful share are more likely to oppose the peaceful split, this should result in significantly more internal conflict in the UNEQUAL treatment than in the BASE treatment, because one alliance player is typically sufficient to trigger internal conflict. In addition to the evolutionary reasoning, some behavioral theories may offer alternative explanations.[25]

Let us turn now to the impact of ex ante non-aggression declarations on the internal allocation of the prize (the PRIVATE and the PUBLIC treatment). Recall that in these treatments all players need to make a declaration and that a player's declaration is non-binding for the player's actual choice of whether to fight. If players care only about monetary payoff, such declarations have no informational value and should not affect the fighting intensity nor the likelihood of internal fight (see Hypothesis 1).[26] Alternatively, there are arguments discussed in political science about reputational or audience cost of breaking non-aggression promises.[27] If such declarations have a causal effect on subsequent behavior and if there is a cost of breaking such "agreements", then public declarations should help to reduce the probability of conflict. Moreover, there might even be a cost of breaking unobserved private declarations, in which case the likelihood of conflict in PRIVATE would be lower than in the BASE treatment.[28]

[24] We do not formalize these results here, as they are similar to the formal analysis in Konrad and Morath (2012a). Building on Schaffer's (1988) concept of evolutionary stability in finite populations, evolutionary stability in contests has also been considered, for instance, by Leininger (2003), Eaton and Eswaran (2003), Eaton et al. (2011), and Konrad and Morath (2012b).

[25] For instance, individuals who care about relative standing may prefer a symmetric contest with low payoff rather than a higher own payoff that, however, is smaller than the co-player's payoff. Similarly, spiteful attitudes, or even a subjective non-monetary benefit from participating in a contest or winning it may contribute to such a behavioral pattern.

[26] This is also in line with the historical evidence in the introduction.

[27] See Mattes and Vonnahme (2010) for a short review and an empirical assessment.

[28] There are many possible channels through which the private and the public announcement could matter. In the context of theories about taste for consistency (see, e.g., Festinger 1957, Cialdini et al. 1995, and the discussion in Guadagno and Cialdini 2010), self-image (Bénabou and Tirole 2006), a subjective cost of lying

Hypothesis A2: *(i) The share of alliance players who choose internal conflict is lowest in the PUBLIC treatment, higher in the PRIVATE treatment and highest in the BASE treatment. (ii) The declaration in the PRIVATE treatment is a better predictor of a player's actual choice of whether to fight than the declaration in the PUBLIC treatment.*

Our data allow us to test the alternative Hypothesis A2 against the prediction of an ineffectiveness of ex ante non-binding agreements and to study the correlation between declarations and actual fighting choices. In the PRIVATE treatment, we expect a larger share of players to truthfully state their prize sharing intention, in which case the correlation between declaration and actual choice should be stronger in PRIVATE than in PUBLIC. Moreover, we will study the relationship between declarations and effort choices and how the history of the game in terms of stage 1 effort choices affects the emergence of conflict.

While the main focus of our experiment is on the choice of conflict, treatment differences in the likelihood of internal conflict should affect effort contributions in stage 1 when the alliance competes with the outgroup player.[29] Hypothesis 1 predicts the same likelihood of internal conflict across all treatments and therefore the same stage 1 effort. If the treatments UNEQUAL, PRIVATE or PUBLIC have a higher (lower) likelihood of internal conflict this reduces (increases) the alliance's value of winning stage 1. Therefore, if Hypothesis A1 applies, then alliance effort in the UNEQUAL treatment should be lower than the BASE treatment. If Hypothesis A2 applies, alliance effort in the PUBLIC treatment should be higher than in the BASE treatment and should be between those two in the PRIVATE treatment, in line with the anticipated likelihood of internal conflict. Choices in the previous stages of the game might also influence effort expended in the internal fight (stage 3), and there may be self-selection into stage 3. While we do not expect that this causes systematic differences in stage 3 effort, we will also briefly discuss behavior in the internal fight.

(Gneezy 2005, Lundquist et al. 2009), and other factors that are not directly related to monetary payoff in the specific interaction, one could see a role for both private and public declarations.

[29] Ke et al. (2013) focused on this aspect in an experiment in which fighting or peaceful sharing was strictly exogenously imposed and found that internal fighting has a negative effect on alliance effort.

4 Results

We report the main results following the logic of backward induction in solving the game, starting therefore with stage 3 (the internal conflict).

Fighting after the break-up of the alliance (stage 3). On average, former alliance members expend an effort of 165 in the internal conflict (stage 3), compared to an equilibrium value of 112.5 for players who maximize their monetary payoffs, in line with the overdissipation results in most contest experiments. (See Figure A.1 in the appendix for time series of average stage 3 effort across treatments.) Turning to a more systematic and detailed analysis of effort in the fight between former alliance members, we use random effects Tobit models to examine whether treatment variations and history matter in case alliances end up fighting internally.[30]

In the simplest model where we only include treatment dummies (PRIVATE, PUBLIC, and UNEQUAL), we find no significant treatment difference for the average effort expended in the internal conflict (compare the first estimation in Table 2; the constant measures average effort in the BASE treatment).[31] In a second estimation, we further explore whether players' choices in stages 1 and 2 (and, in the treatments with declarations, also in stage 0) can explain stage 3 effort. In terms of decisions of whether to fight internally, the included variables are: own ex ante declaration ("Fight0") in treatments PRIVATE and PUBLIC, the co-player's declaration ("Fight0_partner") in PUBLIC, and the actual fighting choice in stage 2 ("Fight2", "Fight2_partner").[32] Moreover, we include the own effort choice in stage 1 ("Effort1"), the absolute difference between the two alliance members' stage 1 effort ("Effort1_diff_abs"), and,

[30] By using a Tobit specification, we take into consideration that effort choices are restricted between 0 and 250 and that in all treatments a number of choices lie on the boundary. Also, the estimation accounts for heterogeneity across players and high correlations within each individual player, by adding random effects to the Tobit model. We report results using data from periods 13 to 24, i.e., we consider more experienced behavior. Using the full dataset does not yield qualitatively different results.

[31] Two-sample Wilcoxon rank-sum tests at the level of matching groups (one observation is the average effort per matching group over periods 13-24) suggest that both UNEQUAL and PRIVATE are significantly different from BASE, at the 5%-level. The difference to the regression results in Table 2 might be caused by the very few observations in stage 3, considerable heterogeneity across individuals, and by not taking into account the significant proportion of observations lying on the upper bound (250); when controlling for individual heterogeneity and censored observations as in the random-effects Tobit estimations, the differences are no longer significant.

[32] "Fight0", "Fight0_partner", "Fight2", and "Fight2_partner" are dummy variables that are equal to 1 if one (or one's partner) chooses "fight" in stage 0 or stage 2, respectively, and zero otherwise.

Dependent Variable: effort of players A or B in stage 3 (periods 13-24)			
Independent variables:	xtTobit1	xtTobit2	xtTobit3
Constant	172.02***	183.61***	184.99***
	(16.56)	(31.15)	(33.77)
UNEQUAL	7.58	5.85	6.95
	(21.54)	(22.23)	(22.36)
UNEQUAL×30%		2.24	1.39
		(16.75)	(16.91)
PRIVATE	17.94	-9.06	-7.91
	(24.55)	(28.16)	(28.36)
PRIVATE×Fight0		57.21	59.99
		(34.92)	(36.03)
PUBLIC	-2.12	-2.03	-0.13
	(22.80)	(22.77)	(22.81)
PUBLIC×Fight0		12.93	10.70
		(36.07)	(38.05)
PUBLIC×Fight0_partner		55.03	50.02
		(35.94)	(36.02)
Effort1		-0.13	-0.14
		(0.10)	(0.10)
Effort1_diff_abs		0.34***	0.35***
		(0.08)	(0.08)
Fight2		-24.79	-25.69
		(26.92)	(26.94)
Fight2_partner		-29.79	-32.10
		(26.10)	(26.13)
Individual characteristics	No	No	Yes
Log-Likelihood	-1160.82	-1149.41	-1148.72
Wald $\chi^2()$	0.84	23.10**	24.41*

Note: 120 subjects, 256 observations. 4 left-censored obs., 194 uncensored obs., 58 right-censored obs. in Tobit models. ***(**,*) significant at 1%(5%,10%). The estimations include treatment dummies as well as interactions indicating the player with the smaller share in UNEQUAL ("30%") and the declaration on the fighting intention by themselves ("Fight0" in PRIVATE and PUBLIC) and their partners ("Fight0_partner" in PUBLIC), actual fighting decisions in stage 2 ("Fight2", "Fight2_partner"), effort choice in stage 1 ("Effort1"), the absolute difference between own and alliance partner's stage 1 effort ("Effort1_diff_abs"), and individual characteristics. Reference category is the BASE treatment.

Table 2: Alliance players' effort in the internal fight (stage 3).

for the UNEQUAL treatment, a dummy that indicates that a player would have received the smaller share in case of peace ("UNEQUAL×30%"). In a third estimation, we further control for individual characteristics, including age, gender, height, number of siblings, and a dummy for economics students; none of these variables is significant. The coefficient that measures absolute difference in the two alliance members' contributions to alliance effort ("Effort1_diff_abs") is highly significant and positive: More unequal effort contributions in stage 1 lead to more effort expended in the internal fight.[33]

Ex ante declarations on the prize sharing intentions have some weak impact on effort in the internal conflict. Alliance members who privately declared an intention to fight expend around 60 tokens more than those who, ex ante, declared an intention to split. This is shown by the coefficient of "PRIVATE×Fight0" in Table 2. When the initial declarations are displayed to other players, this effect of a declaration to fight is lower (compare the estimated coefficient of "PUBLIC×Fight0"). A publicly announced intention to fight, however, seems to have an impact on the alliance partner's effort choice and makes him expend around 50 tokens more, compared to an ex ante declaration of "split" (see the coefficient of "PUBLIC×Fight0_partner"). These large coefficients are not significant though (the p-values are larger than 0.1), potentially because the probability of internal fight in PRIVATE and PUBLIC is rather low. Finally, whether it is their own or their partner's actual choice to fight (i.e., stage 2 choices: "Fight2" and "Fight2_partner") does not significantly affect effort in the internal conflict.

Result 1: *a) Former alliance members of a victorious alliance fight heavily if they turn against each other; average effort does not significantly differ across treatments, in line with Hypothesis 1. b) More unequal effort contributions of the alliance players when fighting the outgroup player cause the subsequent internal fight (if reached) to be more intense.*

Initial intentions and actual choices to break-up (stage 2). We next turn to our main question on the determinants of the likelihood of a victorious alliance to break up and fight (Hypothesis 1 and alternative Hypotheses A1 and A2). For an overview of the results, consider first the right part of Table 3, which presents the alliance players' actual choices of "split" versus

[33] Notice that only the absolute value of the coefficient of the difference in stage 1 efforts has a significant impact on stage 3 effort; the simple difference between own and co-player's stage 1 effort is not significant. Similarly, own stage 1 effort ("Effort1") does not significantly explain stage 3 choices (even when excluding "Effort1_diff_abs"). Due to space constraints, we do not report these two specifications in Table 2. There are multiple explanations for why players who expended much in the alliance contest might also expend high effort in stage 3; but also players who expended comparatively little effort might increase their stage 3 effort, for instance because they are narrowly selfish players who free-ride in stage 1 and take their chances in stage 3.

Treatment	No. of obs. in stage 0	Ex ante declaration "split"	Ex ante declaration "fight"	No. of obs. in stage 2	Actual binding decision "split"	Actual binding decision "fight"	% internal fight if AB wins
BASE	1152	N/A	N/A	550	86.4%	13.6%	24.0%
UNEQUAL	1056	N/A	N/A	448	67.4%	32.6%	54.5%
PRIVATE	1152	80.9%	19.1%	560	83.2%	16.8%	27.5%
PUBLIC	1152	88.9%	11.1%	528	85.6%	14.4%	25.4%

Table 3: Alliance players' decisions on the prize sharing rule before and after the contest with the outgroup player.

"fight", conditional on reaching stage 2.[34] First, we observe a substantial amount of internal conflict in all of the treatments. In the UNEQUAL treatment, "fight" is chosen more frequently than in the BASE treatment. Also, the percentage of winning alliances ending up in internal fight is more than twice as high in UNEQUAL (54.5%) compared to BASE (24%). This considerable propensity to fight despite its negative material consequences contradicts Hypothesis 1, but is very much in line with Hypothesis A1. As suggested by evolutionary theory (and as we will see below), many players with a peaceful share of only 30 percent prefer fighting and impose this fight upon their co-players who would have obtained a 70 percent share and would have been satisfied with the peaceful settlement.

Second, the ex ante declaration stage does not have a noticeable impact on the average fighting propensity. In the three treatments with an equal split in case of peace (BASE, PRIVATE, PUBLIC), alliance players preferred to "fight" in $13.6\% - 16.8\%$ of cases. The left part of Table 3 shows that there are some alliance players who declare an intention to "fight" even when this declaration is publicly observed (11.1% in PUBLIC). The share of alliance players who declare to "fight" is almost twice as high (19.1%) when the declarations are not shown to others in the group. This difference in declarations contrasts with very small differences in the frequency of actual choices to fight between the PRIVATE and the PUBLIC treatment (16.8% compared to 14.4%).

To explore in detail what influences an alliance player's likelihood to choose the internal fight,

[34] Note that, in the experimental instructions, we did not use the word "fight". Instead, participants were asked to choose between a split of the prize (in predefined shares) and competing with their co-player about the entire prize value. Yet we will use the words "split" and "fight" for simplicity whenever we talk about this decision.

Dependent Variable: A dummy for the choice of players A or B between "fight" (=1) and "split" (=0) in stage 2 conditional on winning in stage 1 (periods 13-24)						
Independent Variables:	xtLogit1	AME1	xtLogit2	AME2	xtLogit3	AME3
Constant	-2.84		-2.81		-3.16	
	(0.41)		(0.49)		(0.55)	
UNEQUAL	1.44***	0.14**	-1.71**	-0.11**	-1.69**	-0.11**
	(0.52)	(0.07)	(0.77)	(0.05)	(0.77)	(0.05)
UNEQUAL×30%			4.76***	0.31***	4.71***	0.31***
			(0.78)	(0.05)	(0.78)	(0.05)
PRIVATE	-0.30	-0.02	-1.71**	-0.11**	-1.82***	-0.12***
	(0.54)	(0.03)	(0.69)	(0.05)	(0.68)	(0.05)
PRIVATE×Fight0			3.65***	0.24***	3.52***	0.23***
			(0.83)	(0.06)	(0.83)	(0.06)
PUBLIC	0.16	0.01	0.01	0.00	-0.03	-0.00
	(0.53)	(0.04)	(0.52)	(0.03)	(0.51)	(0.03)
PUBLIC×Fight0			1.73*	0.11*	1.62*	0.10*
			(1.00)	(0.07)	(0.97)	(0.08)
PUBLIC×Fight0_partner			1.70	0.11	1.56	0.10
			(1.17)	(0.07)	(1.16)	(0.08)
Effort1			-0.00	-0.00	-0.00	-0.00
			(0.004)	(0.00)	(0.00)	(0.00)
Effort1_diff			0.02***	0.001***	0.02***	0.001***
			(0.00)	(0.00)	(0.00)	(0.00)
Individual characteristics	No		No		Yes	
Log-likelihood	-352.39		-276.67		-275.00	
Wald $\chi^2()$	12.59***		88.48***		90.93***	

Note: 184 subjects, 940 observations. ***(**,*) significant at 1%(5%,10%). The estimations include treatment dummies as well as interactions indicating the player with the smaller share ("30%" in UNEQUAL), the declaration on their own and partner's fighting intention ("Fight0" in PRIVATE and PUBLIC; "Fight0_partner" in PUBLIC), effort choices in stage 1, the difference between own and alliance partner's stage 1 effort ("Effort1_diff"), and individual characteristics. Reference category is the BASE treatment.

Table 4: Alliance players' probability to choose the internal fight (after having defeated the outgroup player).

we report random-effects logistic regressions and average marginal effects (AME) in Table 4. The dependent variable is an alliance player's binary choice of whether or not to fight against the former ally in case of a victory against the outgroup player C; "1" indicates the choice of internal conflict and "0" the choice of the peaceful split. The reference category is the BASE treatment.

Subjects are 14 percentage points more likely to choose "fight" in the UNEQUAL treatment than in the BASE treatment (see the first estimation in Table 4). Separating the subjects in UNEQUAL according to the peaceful share they would receive, those who would get the large share are 11 percentage points *less* likely to choose to fight than those in BASE (now measured by "UNEQUAL" in xtLogit2 and xtLogit3); the estimated coefficient is significant at the 5%-level. The subjects with the smaller share (30%), however, are 31 percentage points *more* likely to choose to fight than those with the larger share (measured by "UNEQUAL×30%") and therefore about 20 percentage points more likely to choose to fight than the subjects in the BASE treatment. Hence, the regression results further corroborate the results obtained from descriptive data and are in line with Hypothesis A1.

Further, consider the treatments with ex ante declarations. These are on average not significantly different from the BASE treatment in terms of fighting probability (compare the coefficients of "PRIVATE" and "PUBLIC" in xtLogit1). Private or public announcements do not significantly affect the probability that a victorious alliance breaks up and ends up in an internal fight, in contrast to Hypothesis A2.[35]

Finally, the difference between own effort and partner's effort in the stage 1 contest with the outgroup player ("Effort1_diff") has a strongly significant impact on the likelihood to fight (0.1 percentage points per 1 unit of contribution gap).[36] The more a player expended in stage 1 and the less his co-player contributed to alliance effort, the higher the likelihood that the player triggers internal fight.[37] While we did not formulate a hypothesis on this relationship, possible explanations for this effect include a sunk-cost interpretation, a feeling of entitlement to the prize, and a tendency to punish the co-player for having contributed too little in stage 1.

Result 2: *a) The individual probability of choosing to fight is highest among players who would*

[35] Using two-sample Wilcoxon rank-sum (Mann-Whitney) tests at the level of matching groups, the treatments PRIVATE and PUBLIC do not significantly differ from BASE with respect to individual choices to fight and overall occurrence of an internal fight; in UNEQUAL, however, both the share of individuals who choose to fight and the overall likelihood of an internal fight significantly differ from BASE (p-value < 0.001).

[36] In the UNEQUAL treatment, this effect holds independently of the peaceful share a player would have received; an additional interaction term of "Effort1_diff" and a dummy indicating the subjects with the smaller share is insignificant.

[37] "Effort1" and "Effort1_diff" are obviously correlated; hence, the coefficient of "Effort1" becomes significant when dropping the "Effort1_diff". Moreover, contrary to the estimations of stage 3 effort, the simple difference between own and co-player's stage 1 effort matters for a player's willingness to fight; the estimated coefficient of the *absolute* difference is not significant (we do not report this specification due to space constraints). Finally, the intensity of the contest in stage 1 (measured by the outgroup player C's effort or by total effort expended by all three players) does not explain an alliance player's choice of internal conflict.

get the smaller share (30%) in the UNEQUAL treatment and lowest among those who would get the larger share (70%) in the UNEQUAL treatment; overall, the likelihood of internal fight is more than twice as high in UNEQUAL compared to BASE.
b) Ex ante declarations on prize sharing intentions do not have a significant effect on the likelihood of internal conflict. c) The probability of a choice to fight is increasing in the difference between own and the co-player's effort contribution when fighting the outgroup player.

Even if there is no average effect of the ex ante declarations for the decision of whether to fight, players may try to make use of the public declarations and react to whether the declarations are public or private. The results in Table 4 show that those alliance players who have *privately* declared that they intend to choose "fight" are 24 percentage points more likely to actually initiate a fight in stage 2 than those who have declared to "split" (compare "PRIVATE×Fight0" in xtLogit2). This suggests that private declarations are not random, but reveal true intentions. Also in the PUBLIC treatment, a player is more likely to initiate internal fight if he declared an intention to choose "fight" in stage 0 (see the coefficients for "PUBLIC×Fight0" in Table 4); the significance level, however, is much weaker (and the marginal effect is smaller) than in the PRIVATE treatment. The declarations in the PUBLIC treatment may not fully reveal a player's true preference, but alliance players may choose a declaration which they expect to have a strategic effect on the alliance partner's behavior or on the joint enemy. This can explain why the private non-aggression declarations (in PRIVATE) may be a better predictor of players' actual choice than the public non-aggression declaration (in PUBLIC). Also, it can explain why the effort contribution to the fight against the outgroup does not significantly react to a non-aggression declaration in the PUBLIC treatment (as we will examine in the next section).

As an additional remark on the declarations, Table 5 contrasts declarations on the prize sharing intention with actual choices. In both the PRIVATE and the PUBLIC treatment, a large share of players of a victorious alliance make a choice in stage 2 which is consistent with their ex ante declaration (86% in the treatment with private declarations and 84.9% if the declaration was publicly announced).[38] The two treatments, however, differ when analyzing *deviations* of

[38] This reveals the endogeneity problem that may loom in empirical work which Mattes and Vonnahme (2010)

Treatment	PRIVATE	PUBLIC
Declaration = Choice	86%	84.9%
Declaration ≠ Choice	14%	15.2%
of which:		
Declaration="fight" & Choice="split"	60.3%	31.3%
Declaration="split" & Choice="fight"	39.7%	68.8%

Table 5: Comparison of ex ante declarations and actual fighting decisions in the PRIVATE and the PUBLIC treatment.

victorious alliance players from their initial declaration. In the PRIVATE treatment, a majority of players switch from declaring "fight" to an actual choice of "split" (60.3%); in the PUBLIC treatment, this share is only 31.3%, but here the majority of switches is from declaring "split" to an actual choice of "fight" (68.8%). This observation is in line with Hypothesis A2. In the PRIVATE treatment, alliance players could state their true fighting intention without fearing any consequences on the co-player's actions. (We can speculate about what causes players sometimes to switch to "split"; one possible reason is the partner's effort contribution.) In the PUBLIC treatment, the low share of alliance players who declared "fight" and the substantial share of players who deviated from a declaration of "split" suggests an attempt to make strategic use of the (non-binding) public announcement by declaring "split" despite an intention to fight.[39]

Effort in the conflict with the stand-alone player (stage 1). We finally turn to the contest between the alliance and the outgroup player (stage 1) to test whether the likelihood of internal fighting affects the alliance's ability to mobilize joint effort. The time series of efforts is shown in Figure A.2 in the appendix.[40] We again use random effects Tobit models to estimate an alliance player's effort in stage 1 (see Table 6). The included explanatory variables are treatment dummies ("UNEQUAL", "PRIVATE", "PUBLIC"), and, from estimation 2 onwards, variables

also hint at: Using non-aggression declarations as an exogenous explanatory variable might overestimate the effect of such promises for the probability of violent conflict.

[39] Further support for this result can be found in Table A.1 in the appendix where we re-examine the probability of choosing "fight" in stage 2, separating the observations according to the ex ante declaration.

[40] The summary statistics in Figure A.2 in the appendix also suggest that the strategic reaction of the outgroup player C to the treatment variations is very small; this can be confirmed by running random effects Tobit regressions similar to the estimated equation in Table 6.

Dependent Variable: effort of players A or B in stage 1 (periods 13-24)			
Independent variables:	xtTobit1	xtTobit2	xtTobit3
Constant	54.10***	54.13***	59.28***
	(5.43)	(5.36)	(6.14)
UNEQUAL	-17.06**	-17.00**	-15.26**
	(7.88)	(7.78)	(7.39)
PRIVATE	-0.59	-7.13	-4.26
	(7.68)	(7.67)	(7.34)
PRIVATE×Fight0		39.78***	41.63***
		(6.87)	(6.84)
PUBLIC	-0.84	0.80	2.44
	(7.68)	(7.64)	(7.24)
PUBLIC×Fight0		10.63	11.03
		(7.74)	(7.71)
PUBLIC×Fight0_partner		-30.55***	-30.85***
		(6.73)	(6.72)
Characteristics	No	No	Yes
Log-likelihood	-10427	-10339	-10389
Wald $\chi^2()$	6.49*	62.59***	84.88***

Note: 188 subjects, 2256 observations. There are 281 left-censored obs., 1961 uncensored obs., 14 right-censored obs. in Tobit models. ***(**,*) significant at 1%(5%,10%). The estimations include treatment dummies as well as interactions indicating their own and their partner's declaration on the fighting intentions ("Fight0" in PRIVATE and PUBLIC; "Fight0_partner" in PUBLIC) and individual characteristics. Reference category is the BASE treatment.

Table 6: Alliance players' effort in the conflict with the outgroup player (stage 1).

indicating the players' declarations in stage 0 ("Fight0", "Fight0_partner") and individual-specific characteristics as obtained from the exit questionnaire.[41]

In all estimations in Table 6, we find that average effort in stage 1 in UNEQUAL is around 17 points lower than average effort expended in BASE; the difference is statistically significant at the 5% level. Hence, in line with the results on the likelihood of internal conflict (Hypothesis A1), the unequal split of the prize reduces the alliance players' value of winning stage 1 and makes the hold-up problem more severe. Moreover, the coefficients on the treatment dummies PRIVATE and PUBLIC are both not significantly different from zero (compare the first estimation in Table 6), which suggests that adopting a stage 0 with private or public declarations does not help to

[41] Among the individual characteristics included, the only significant coefficient is obtained for economics students who expend around 18 points less effort in stage 1.

mitigate the hold-up problem. This result is in line with the non-binding nature of the ex ante declarations and with the previous observation that ex ante declarations do not help to reduce the likelihood of internal fight (in contrast to the alternative Hypothesis A2).[42]

Result 3: *a) Anticipation of the higher internal fighting frequency in the UNEQUAL treatment leads to lower stage 1 effort of the alliance members, compared to the BASE treatment. b) Non-binding ex ante declarations do not help to mitigate the hold-up problem.*

Even if there is no effect of ex ante declarations on average, estimations 2 and 3 in Table 6 show that, in the PRIVATE treatment, alliance players who have secretly declared that they intend to choose "fight" expend significantly more effort (around 40 points) than those who have declared to "split" (compare the coefficient of "PRIVATE×Fight0"). Second, in the PUBLIC treatment, players who have declared that they intend to choose "fight" do not expend significantly more effort (although the estimated coefficient of "PUBLIC×Fight0" is positive). Third, if in the PUBLIC treatment the partner in the alliance has declared an intention to fight ("Fight0_partner"), a player expends much less effort (the coefficient is -30.55 and highly significant). Even if the declarations are non-binding, they may have informational value in a world in which players do not like to deviate from their previous declaration.[43] Consequently, players may take this declaration quite seriously and anticipate that a public announcement to fight implies a higher likelihood of internal conflict.

5 Conclusions

Members of a victorious alliance may decide to peacefully share the prize that they have jointly won, or they may decide to enter into a resource-wasteful conflict about the spoils of victory. Whether or not they fight will depend on the institutional setup and existing norms and rules

[42] The results on the treatment differences are supported by two-sample Wilcoxon rank-sum (Mann-Whitney) tests at the level of matching groups: PRIVATE and PUBLIC do not significantly differ from BASE, but we can reject (at the 5%-level) that average effort in UNEQUAL is the same as in BASE.

[43] If players differ in their cost of making a false or inconsistent declaration and if this cost is strictly positive, then the set of players who declare "split" is a mixture of people who eventually choose "split" and who eventually choose "fight". But only those who eventually choose "fight" and have a high cost of declaring "split" under these circumstances will declare "fight". Accordingly, a declaration of "fight" is revealing in such a context.

about how to split the prize if the division of the prize takes place peacefully. While many factors may play a role in the historical examples discussed in the introduction, the sharing rules in case of a peaceful settlement are potentially important factors that may yield different probabilities of the emergence of internal distributional conflict. We study experimentally how different institutional environments affect the emergence of internal conflict and how variations in the threat of internal conflict influence the alliance members' willingness to contribute to the fight against an outgroup player.

As our first main result, we find that an imbalance between the alliance members' 'strength' in the internal conflict and the rent distribution in case of a peaceful settlement matters for the likelihood of internal conflict. Players are more likely to fight internally the more unequal the division of the prize is, even if the peaceful settlement yields a higher material payoff to both alliance players. This result is in line with the prediction based on evolutionary arguments where players care about their relative material payoff. Moreover, players contribute less effort to the contest against the outgroup player if the probability of a break-up of the alliance in a fight about the prize is higher.

As a second dimension, we study the role of non-aggression declarations at the onset of the conflict between the alliance and its adversary: Alliance players may make non-binding declarations at this point about their intention of whether to fight internally or to share peacefully with their alliance partner. We analyze whether the opportunity to make such declarations affects the actual fighting probability and the effort contribution to the contest against the outgroup player, and we consider both publicly revealed declarations and private declarations which are not shown to the co-players but may convey information about the initial prize sharing intention. In line with the prediction for players who maximize their monetary payoff, the opportunity to make declarations neither changes the actual probability to fight nor the effort contributions of alliance players in the conflict with the outgroup player in a significant way, compared to a situation where declarations of this type are not possible. Even if alliance players may intend to make use of a public declaration of an intention to share the prize peacefully with the former ally, the opportunity to make such non-aggression declarations does not help to mitigate the hold-up problem: it does not reduce actual fighting within victorious alliances.

In all treatments that we consider, the likelihood of internal distributional conflict is higher the more unequally the alliance members have contributed to the fight against the joint enemy. Former alliance members who expended more effort than their alliance partner when fighting the outgroup player are more likely to oppose the peaceful settlement. Moreover, higher asymmetries in the alliance partners' contributions to alliance effort cause a subsequent internal fight to be more intense.

To summarize, our experiment confirms the emergence of resource-wasteful conflict even in situations where peaceful settlement leads to a Pareto superior outcome. By opposing the peaceful split, players are willing to sacrifice a substantial share of the material payoff in case of peace and to accept an expected material payoff which, taking into account the very intense fighting in the internal conflict and the overdissipation of resources in this conflict, is much lower than what they would receive in case of peace. The setup of our experiment allows to identify two main reasons for why, despite this dissipation of resources, alliances break up in a violent conflict: Former allies are not willing to accept a distribution of resources that does not coincide with their relative strength and with their relative contribution to alliance success. In turn, alliances that are more "symmetric" in terms of peaceful arrangements and in terms of willingness to contribute to alliance effort are more successful in keeping internal peace and, as a direct effect, they are more successful in the conflict with the joint enemy. Our experiment also shows that institutions that allow players to declare their intention to keep peace (or fight) do not effectively improve alliance success, even if some players may feel a cost of breaking such initial arrangements. Instead, in our context, similarity between players with respect to their willingness to contribute to the joint cause and their shares in the prize seems to be more promising for the formation of alliances in the shadow of conflict.

The results have been obtained in a laboratory environment. This is useful to isolate and simplify the decision situation, compared to the complex and dynamic setup of peace negotiations between countries. Experimental results about fighting propensity or the commitment value of non-aggression declarations in international politics should, of course, not be used to make direct probability predictions about peace negotiations. However, the difference in behavior that is triggered by changes in the negotiation setup is indicative also of the direction of behavioral differences that emerge from a change in the institutional environment in which peace negotiations between country leaders take place. For instance, we expect country leaders to be less willing to settle on an asymmetric peace agreement that is disadvantageous for them, and to be more willing to escalate conflict even if conflict is wasteful. In international politics, this result may even be magnified by concerns about what a disadvantageous agreement implies for future interactions, by reputational concerns, and by concerns about how such disadvantageous outcomes affect a country leader's likelihood of staying in power.

References

[1] Abbink, K., and J. Brandts, 2009, Political autonomy and independence: theory and experimental evidence, Working Paper.

[2] Abbink, K., J. Brandts, B. Herrmann, and H. Orzen, 2010, Intergroup conflict and intragroup punishment in an experimental contest game, American Economic Review, 100(1), 420-447.

[3] Abbink, K., J. Brandts, B. Herrmann and H. Orzen, 2012, Parochial altruism in inter-group conflicts, Economics Letters, 117(1), 45-48.

[4] Ahn, T.K., R.M. Isaac, and T.C. Salmon, 2011, Rent seeking in groups, International Journal of Industrial Organization, 29(1), 116-125.

[5] Amegashie, J.A., C.B. Cadsby, and Y. Song, 2007, Competitive burnout: theory and experimental evidence, Games and Economic Behavior, 59(2), 213-239.

[6] Amaldoss, W., and A. Rapoport, 2009, Excessive expenditure in two-stage contests: theory and experimental evidence, in: I.N. Haugen and A.S. Nilsen (eds.), Game Theory: strategies, equilibria, and theorems, Nova Science Publishers, NY, 241-266.

[7] Anbarci, N., S. Skaperdas, and C. Syropoulos, 2002, Comparing bargaining solutions in the shadow of conflict: how norms against threats can have real effects, Journal of Economic Theory, 106(1), 1-16.

[8] Anderson, L.R., and S.L. Stafford, 2003, An experimental analysis of rent seeking under varying competitive conditions, Public Choice, 115, 199-216.

[9] Balafoutas, L., R. Kerschbamer, and M. Sutter, 2012, Distributional preferences and competitive behavior, Journal of Economics Behavior and Organization, 83(1), 125-135.

[10] Beilenson, L.W., 1969, The treaty trap: a history of the performance of political treaties by the United States and European Nations, Washington D.C., Public Affairs Press.

[11] Bénabou, R., and J. Tirole, 2006, Incentives and prosocial behavior, American Economic Review, 96(5), 1652-1678.

[12] Beviá, C., and L.C. Corchón, 2010, Peace agreements without commitment, Games and Economic Behavior, 68(2), 469-487.

[13] Blainey, G., 1988, The causes of war, 3rd edition, Free Press, New York.

[14] Bornstein, G., 2003, Intergroup conflict: individual, group, and collective interests, Personality and Social Psychology Review, 7(2), 129-145.

[15] Bornstein, G., U. Gneezy, and R. Nagel, 2002, The effect of intergroup competition on group coordination: an experimental study, Games and Economic Behavior, 41(1), 1-25.

[16] Bunselmeyer, R.E., 1975, The cost of the war 1914-1919, British Economic War Aims and the Origins of Reparation, Archon Books.

[17] Cason, T.N., R.M. Sheremeta, and J. Zhang, 2012, Communication and efficiency in competitive coordination games, Games and Economic Behavior, 76(1), 26-43.

[18] Charness, G., L. Rigotti, and A. Rustichini, 2007, Individual Behavior and Group Membership, American Economic Review, 97(4), 1340-1352.

[19] Chen, Y., and S. Xin Li, 2009, Group identity and social preferences, American Economic Review, 99(1), 431-457.

[20] Chen H., S.H. Ham, and N. Lim, 2011, Designing multiperson tournaments with asymmetric contestants: an experimental study, Management Science, 57(5), 864-883.

[21] Cialdini, R.B., M.R. Trost, and J.T. Newsom, 1995, Preference for consistency: the development of a valid measure and the discovery of surprising behavioral implications, Journal of Personality and Social Psychology, 69, 318-328.

[22] Claude, I.L., 1962, Power and International Relations, Random House, New York.

[23] Davis, D.D., and R.J. Reilly, 1998, Do too many cooks always spoil the stew? An experimental analysis of rent-seeking and the role of a strategic buyer, Public Choice, 95, 89-115.

[24] Dechenaux, E., D. Kovenock, and R. Sheremeta, 2012, A survey of experimental research on contests, all-pay auctions and tournaments, Working Papers 12-22, Chapman University, Economic Science Institute.

[25] Eaton, B.C., and M. Eswaran, 2003, The evolution of preferences and competition: a rationalization of Veblen's theory of invidious comparisons, Canadian Journal of Economics, 36(4), 832-859.

[26] Eaton, B.C., M. Eswaran, and R.J. Oxoby, 2011, 'Us' and 'them': the origin of identity, and its economic implications, Canadian Journal of Economics, 44(3), 719-748.

[27] Esteban, J.M., and J. Sákovics, 2003, Olson vs. Coase: coalitional worth in conflict, Theory and Decision, 55(4), 339-357.

[28] Fearon, J.D., 1995, Rationalist explanations for war, International Organization, 49(3), 379-414.

[29] Festinger, L., 1957, A theory of cognitive dissonance, Stanford, CA: Stanford University Press.

[30] Fischbacher, U., 2007, z-Tree: Zurich Toolbox for Ready-made Economic Experiments, Experimental Economics, 10(2), 171-178.

[31] Garfinkel, M., and S. Skaperdas, 2000, Conflict without misperceptions or incomplete information: how the future matters, Journal of Conflict Resolution, 44(6), 793-807.

[32] Gneezy, U., 2005, Deception: the role of consequences, American Economic Review, 95(1), 384-394.

[33] Greiner, B., 2004, An online recruitment system for economic experiments, in: K. Kremer, V. Macho (eds.): Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Ges. für Wiss. Datenverarbeitung, 79-93.

[34] Guadagno, R.E., and R.B. Cialdini, 2010, Preference for consistency and social influence: a review of current research findings, Social Influence, 5(3), 152-163.

[35] Gunnthorsdottir, A., and A. Rapoport, 2006, Embedding social dilemmas in intergroup competition reduces free-riding, Organizational Behavior and Human Decision Processes, 101(2), 184-199.

[36] Jackson, M.O., and M. Morelli, 2011, The reasons for wars: an updated survey, in: C.J. Coyne, R.L. Mathers (eds.): The Handbook on the Political Economy of War, Cheltenham, UK: Edward Elgar, 34-57.

[37] Johnson, J.C., and B.A. Leeds, 2011, Defense pacts: a prescription for peace?, Foreign Policy Analysis, 7(1), 45-65.

[38] Katz, E., and J. Tokatlidu, 1996, Group competition for rents, European Journal of Political Economy, 12(4), 599-607.

[39] Ke, C., K.A. Konrad, and F. Morath, 2013, Brothers in arms - an experiment on the alliance puzzle, Games and Economic Behavior, 77, 61-76.

[40] Kimball, Anessa L., 2006, Alliance formation an conflict initiation: the missing link, Journal of Peace Research, 43(4), 371-389.

[41] Konrad, K.A., 2009, Strategy and Dynamics in Contests, Oxford University Press.

[42] Konrad, K.A., and W. Leininger, 2011, Self-enforcing norms and efficient non-cooperative collective action in the provision of public goods, Public Choice, 146(3-4), 501-520.

[43] Konrad, K.A., and F. Morath, 2012a, Evolutionary determinants of war, Max Planck Institute for Tax Law and Public Finance Working Paper No. 2012-12.

[44] Konrad, K.A., and F. Morath, 2012b, Evolutionarily stable in-group favoritism and out-group spite in intergroup conflict, Journal of Theoretical Biology, 306, 61-67.

[45] Kugler, T., A. Rapoport, and A. Pazy, 2010, Public good provision in inter-team conflicts: effects of asymmetry and profit-sharing rule, Journal of Behavioral Decision Making, 23(4), 421-438.

[46] Lacomba, J.A., F. Lagos, E. Reuben, and F. van Winden, 2014, On the escalation and de-escalation of conflict, Games and Economic Behavior 86, 40-57.

[47] Leininger, W., 2003, On evolutionarily stable behavior in contests, Economics of Governance, 4(3), 177-186.

[48] Levy, J.S., 1981, Alliance formation and war behavior: an analysis of the Great Powers, 1495-1975, Journal of Conflict Resolution, 25(4), 581-613.

[49] Lim, W., A. Matros, and T.L. Turocy, 2014, Bounded rationality and group size in Tullock contests: experimental evidence, Journal of Economic Behavior & Organization, 99, 155-167.

[50] Leeds, B.A., and B. Savun, 2007, Terminating alliances: why do states abrogate agreements?, Journal of Politics, 69(4), 1118-1132.

[51] Long, A.G., T. Nordstrom, and K. Baek, 2007, Allying for peace: treaty obligations and conflict between allies, Journal of Politics, 69(4), 1103-1117.

[52] Lundquist, T., T. Ellingsen, E. Gribbe, and M. Johannesson, 2009, The aversion to lying, Journal of Economic Behavior and Organization, 70(1-2), 81-92.

[53] Mattes, M., 2008, The effect of changing conditions and agreement provisions on conflict and renegotiation between states with competing claims, International Studies Quarterly, 52(2), 315-334.

[54] Mattes, M., and G. Vonnahme, 2010, Contracting for peace: do nonaggression pacts reduce conflict?, Journal of Politics, 72(4), 925-938.

[55] McBride, M., and S. Skaperdas, 2009, Conflict, settlement, and the shadow of the future, CESifo Working Paper No. 2897.

[56] Mearsheimer, J.J., 1994-1995, The false promise of international institutions, International Security, 19(3), 5-49.

[57] Millner, E.L., and M.D. Pratt, 1989, An experimental investigation of efficient rent-seeking, Public Choice, 62, 139-151.

[58] Millner, E.L., and M.D. Pratt, 1991, Risk-aversion and rent-seeking: an extension and some experimental evidence, Public Choice, 69, 81-92.

[59] Morgan, J., H. Orzen, and M. Sefton, 2012, Endogenous entry in contests, Economic Theory, 51(2), 435-463.

[60] Morrow, J.D., 1991, Alliances and asymmetry - an alternative to the capability aggregation model of alliances, American Journal of Political Science, 35(4), 904-933.

[61] Morrow, J.D., 1994, Alliances, credibility, and peacetime costs, Journal of Conflict Resolution, 38(2), 270-297.

[62] Morrow, J.D., 2000, Alliances: why write them down?, Annual Review of Political Science, 3, 63-83.

[63] Niou, E.M.S., and P.C. Ordeshook, 1994, Alliances in anarchic international systems, International Studies Quarterly, 38(2), 167-191.

[64] Nitzan, S., 1991, Collective rent dissipation, Economic Journal, 101, 1522-1534.

[65] O'Connor, R.G., 1969, Victory in modern war, Journal of Peace Research, 6(4), 367-384.

[66] Organski, A.F.K., 1958, World politics, Knopf, New York.

[67] Parco, J.E., W. Amaldoss, and A. Rapoport, 2005, Two-stage contests with budget constraints: an experimental study, Journal of Mathematical Psychology, 49(4), 320-338.

[68] Potters, J., C.G. De Vries, and F. van Winden, 1998, An experimental examination of rational rent-seeking, European Journal of Political Economy, 14(4), 783-800.

[69] Price, C.R., and R.M. Sheremeta, 2011, Endowment effects in contests, Economics Letters, 111(3), 217-219.

[70] Price, C.R., and R.M. Sheremeta, 2014, Endowment origin, demographic effects and individual preferences in contests, Journal of Economics and Management Strategy, forthcoming.

[71] Reynolds, D., 2002, From world war to cold war: the wartime alliance and post-war transitions, 1941-1947, The Historical Journal, 45(1), 211-227.

[72] Schaffer, M.E., 1988, Evolutionary stable strategies for a finite population and a variable contest size, Journal of Theoretical Biology, 132(4), 469-478.

[73] Schmidt, D., R.M. Sheremeta, R. Shupp, and J.M. Walker, 2013, Resource allocation contests: experimental evidence, Journal of Economic Psychology, 39, 257-267.

[74] Sheremeta, R.M., 2010, Experimental comparison of multi-stage and one-stage contests, Games and Economic Behavior, 68(2), 731-747.

[75] Sheremeta, R.M., 2011, Contest design: an experimental investigation, Economic Inquiry, 49, 573-590.

[76] Sheremeta, R.M., and J. Zhang, 2010, Can groups solve the problem of over-bidding in contests? Social Choice and Welfare, 35, 175-197.

[77] Skaperdas, S., 1996, Contest success functions, Economic Theory 7(2), 283-290.

[78] Skaperdas, S., and C. Syropoulos, 1996, Can the shadow of the future harm cooperation?, Journal of Economic Behavior & Organization, 29(3), 355-372.

[79] Slantchev, B.L., 2003, The power to hurt: costly conflict with completely informed states, American Political Science Review, 97(1), 123-133.

[80] Sutter, M., 2009, Individual behavior and group membership: comment, American Economic Review, 99(5), 2247-2257.

[81] Tullock, G., 1980, Efficient rent seeking, in: J.M. Buchanan, R.D. Tollison, and G. Tullock (eds.), Towards a Theory of the Rent-Seeking Society, Texas A&M University Press, College Station, 97-112.

[82] Wärneryd, K., 1998, Distributional conflict and jurisdictional organization, Journal of Public Economics, 69(3), 435-450.

[83] Wagner, R.H., 1994, Peace, war, and the balance of power, American Political Science Review, 88(3), 593-607.

[84] Wittman, D., 1979, How a war ends: a rational model approach, Journal of Conflict Resolution, 23(4), 743-763.

A Supplementary appendix

A.1 Choices of "fight" versus "split" conditional on the initial declaration

	Dependent Variable: A dummy for alliance player's actual choice between "fight" (=1) and "split" (=0) in stage 2 conditional on winning in stage 1			
Data	Subsample 1 (Declaration="split")		Subsample 2 (Declaration="Fight)	
Model	xtLogit1	AME1	xtLogit2	AME2
Constant	-3.68***		0.96**	
	(0.50)		(0.46)	
PUBLIC	0.94**	0.05*	-0.45	-0.10
	(0.48)	(0.03)	(0.45)	(0.09)
Effort1_diff	0.01***	0.001***	0.01***	0.001***
	(0.00)	(0.00)	(0.00)	(0.00)
PUBLIC×Effort1_diff	-0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Characteristics	Yes		Yes	
Log-likelihood	-235.95		-93.90	
Wald $\chi^2()$	35.53***		20.76***	

Note: Subsample 1: 94 subjects, 930 observations. Subsample 2: 37subjects, 156 observations. ***(**,*) significant at 1%(5%,10%). The estimations include a treatment dummy ("PUBLIC"), the difference between own and alliance partner's stage 1 effort ("Effort1_diff") as well as their interaction term, and individual characteristics as obtained from the exit questionnaire. Reference category is the PRIVATE treatment.

Table A.1: Relation between initial fighting intentions and actual fighting choice.

The first subsample only includes observations where victorious alliance members had ex ante declared an intention to split. After controlling for differences in stage 1 effort ("Effort1_diff") and individual characteristics, the likelihood of a switch from "split" to "fight" is still significantly higher for players in the PUBLIC treatment than in the PRIVATE treatment (compare the first estimation in Table A.1). On the other hand, since ex ante declarations of "fight" should reveal players' true intention in both the PRIVATE and the PUBLIC treatment, this treatment difference disappears in the estimation restricted to the subsample of players who initially declared an intention to fight (compare the second estimation in Table A.1).

A.2 Average effort in the internal conflict

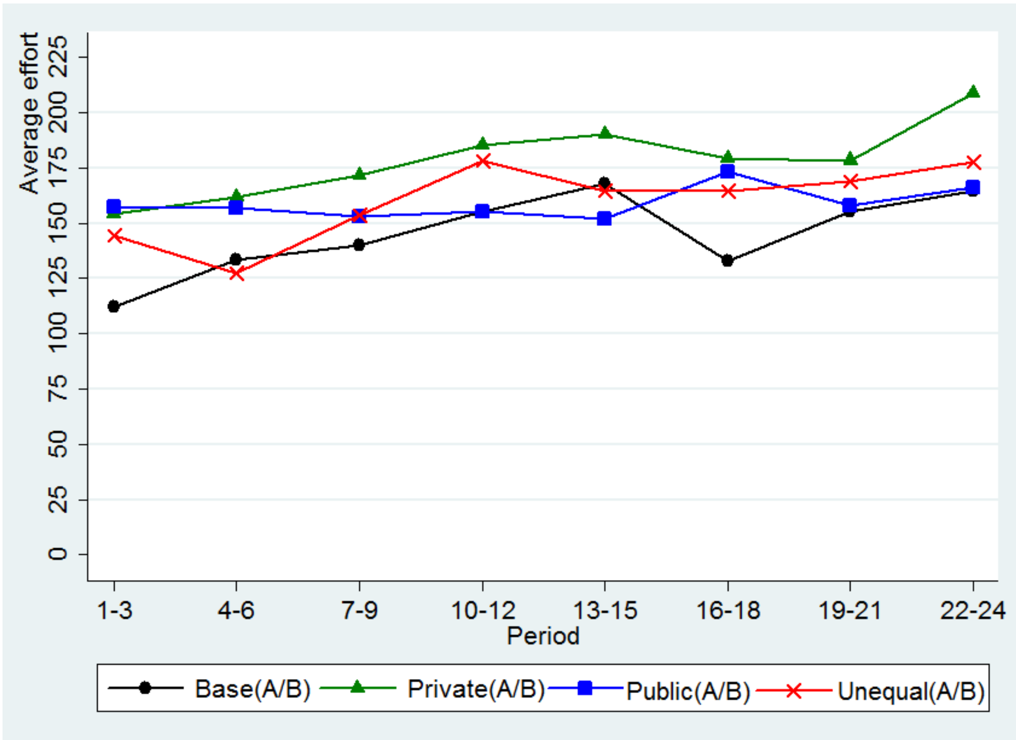

Figure A.1: Average effort of players A or B in stage 3 (by treatment).

A.3 Average effort in the conflict between the alliance and the stand-alone player

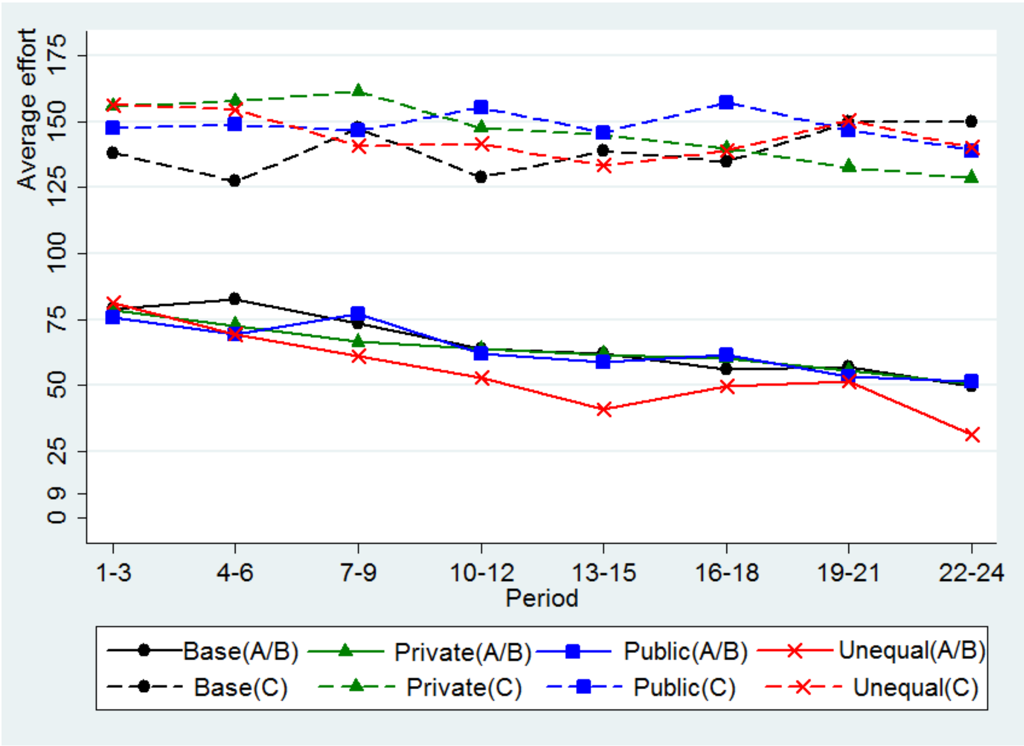

Figure A.2: Average effort in stage 1 for alliance players (A or B) and player C.

A.4 Experimental Instructions (a sample for the PUBLIC treatment)[44]

Welcome to this experiment! Please read this instruction carefully and completely. Properly understanding the instruction will help you to make better decisions and hence earn more money.

Your earnings in this experiment will be measured in Talers. At the end of the experiment we will convert the Talers you have earned to cash and pay you in private. For each 45 Talers you earn you will be paid 1 Euro in cash. Therefore, the more Talers you earn, the more cash you will gain at the end of today's experiment. In addition to the Talers earned during the experiment, each participant will receive a show-up fee of 4 Euros.

Please keep in mind that you are not allowed to communicate with other participants during the experiment. If you do not obey this rule you will be asked to leave the laboratory without getting paid. Whenever you have a question, please raise your hand; an experimenter will come to you.

1. Your task

This experiment will consist of 24 rounds. Before the actual experiment starts, you will first have to answer a few questions related to the experiment. The questions will be presented to you through the computer screen.

In the experiment, groups consisting of three people are formed. These groups are randomly composed in each round. Your task in each round is to make some decisions. The money you earn depends on your decision and the decisions of the two other players in your group.

Let the three players in one group be called A, B, and C. In each round, players A, B, and C compete for a prize of 450 Talers. The competition in each round works as follows:

1. Two players A and B form an "alliance". Player C is playing on his own.
2. If player C wins the competition, he will gain the whole prize of 450 Talers.
3. In case the alliance of A and B wins the competition, then players A and B have to choose between two possible options about how the prize should be divided among them:
    - Option 1: A and B each obtain half of the prize, i.e., 225 Talers each.

[44] The original instructions distributed to the participants were in German.

- Option 2: A and B compete for the whole prize of 450 Talers.

4. Your role in the experiment will be either that of player A, B, or C. This role will be randomly assigned to you. Each participant will keep his role throughout the entire experiment.

5. At the beginning of each round, players A and B declare how they prefer to divide the prize if they, as an alliance, win the competition against player C (option 1 or option 2). Afterwards, this decision will be displayed on the screen to all players of a group. There will be the possibility to change the own choice if the alliance of A and B wins the competition with C.

6. Then, all players will simultaneously choose an "expenditure". Each player decides independently on his own expenditure. A player's expenditure is chosen as an integer between 0 and 250, and it corresponds to the amount of Talers the player would like to expend in the competition to win the prize. You will have to pay this amount of Talers to the lab, whether or not you win the competition. In the following, player A's expenditure will be denoted by X_A, player B's expenditure will be denoted by X_B, and player C's expenditure will be denoted by X_C.

7. Afterwards, you will be shown the amount of Talers that the other players in your group have expended. The expenditures of players A and B will be added up, and the sum of X_A and X_B corresponds to the expenditure that the alliance of A and B spends on the competition. The total expense within a group is equal to the sum of all players expenditures: $X_A + X_B + X_C$.

8. Now a "fortune wheel" will turn and decide whether the alliance consisting of A and B or whether player C wins the 450-Taler-prize. As you will see, the fortune wheel is divided into two colors - red and blue. The red color represents the total Talers spent by player A and B (i.e., $X_A + X_B$). The blue color represents the Talers spent by player C (i.e., X_C). The size of the two colored areas on the wheel represents exactly their shares in the total expense ($X_A + X_B + X_C$).

9. At the centre of the fortune wheel there is an arrow initially pointing to the top. After some time the arrow starts to rotate and then stops randomly. If the arrow stops in the red-colored area, players A and B win the prize. If the arrow stops in the blue-colored area, player C wins the

prize. This means that the probability that players A and B win the prize is equal to their share of their joint expenditure in the total expense, hence

$$\text{probability that } A \text{ and } B \text{ win} = \frac{\text{expenditure } X_A + \text{expenditure } X_B}{\text{total expense } X_A+X_B+X_C}$$

Equivalently, the probability that player C wins the prize is equal to the share of his expenditure in the total expense:

$$\text{probability that } C \text{ wins} = \frac{\text{expenditure } X_C}{\text{total expense } X_A+X_B+X_C}$$

For your information, the probabilities that either the alliance of A and B or player C wins the prize will be displayed to you.

Therefore, each player's probability of winning depends not only on his own expenditure in the competition but also on the expenditures of the other players in the group. Note that the more Talers a player spends, the more likely it is that he wins the competition. More Talers expended, however, means that a player has to pay more Talers to the lab.

10. If none of the players expends any Taler, i.e., $X_A = X_B = X_C = 0$, then it is equally likely that either the alliance of A and B or player C wins. If A and B both do not expend any Taler, but C expends at least 1 Taler, player C wins the competition. If player C does not expend any Taler, but either player A or player B (or both) expends at least 1 Taler, the alliance of A and B wins the competition.

11. Every player has to pay his expenditure (in Taler) to the lab, irrespective of the outcome of the fortune wheel. Therefore, your earnings per round will be calculated as your gain in the competition minus your expenditure: earnings = gain − expenditure.

12. In case player C wins, the competition ends. Player C gets the 450-Taler-prize; players A and B will gain nothing. While players A and B do not have any gain, but have to pay their expenditures, the earnings of player C are calculated as follows: C's earnings $= 450 - X_C$.

13. In case the alliance of A and B wins the competition, then player C will receive nothing, but he has to pay his expenditure. Players A and B have to choose independently between two options about how to divide the prize among them:

- Option 1: The prize will be split between A and B: Each of the players A and B obtains exactly half of the prize, i.e., 225 Talers. This means for A and B: Earnings of $A = 225-$ expenditure X_A and earnings of $B = 225 - X_B$.

- Option 2: Players A and B again compete with each other for the prize of 450 Talers. The procedure of this competition between A and B is basically the same as in the competition between the alliance of A and B and player C. First, A and B decide simultaneously and independently about the amount of Talers they would like to expend in order to win the prize of 450 Taler. This expenditure is again chosen as an integer between 0 and 250, and it has to be paid to the lab in addition to the expenditures already paid (X_A and X_B), whether or not the player wins the competition.

  In the following these new expenditures of A and B are denoted by Y_A and Y_B. (Note that these expenditures are only chosen if the alliance of A and B has won against player C.) Again a fortune wheel will determine the winner. The probability that A wins the prize of 450 Taler will be:

  $$\text{probability that } A \text{ wins} = \frac{\text{expenditure } Y_A}{\text{total expense } Y_A + Y_B}$$

  Equivalently, the probability that player B wins, will be:

  $$\text{probability that } B \text{ wins} = \frac{\text{expenditure } Y_B}{\text{total expense } Y_A + Y_B}$$

  Therefore, each player's probability of winning now depends only on the expenditures in this new competition. A yellow-colored area on the fortune wheel will represent the share of A's expenditure in total expense $Y_A + Y_B$, and a green-colored area will represent the share of B's expenditure in total expense. Again the arrow will rotate to decide whether A or B wins the prize.

  Hence, in case players A and B have won the competition with player C and Option 2 has been selected, the earnings of players A and B are calculated as follows.

  - If A wins against B, player B has to pay both his expenditures X_B and Y_B but does not receive any gain. A's earnings in this case will be: A's earnings $= 450- X_A - Y_A$.
  - If A loses against B, player A has to pay both his expenditures X_A and Y_A but does

not receive any gain. B's earnings will be: B's earnings $= 450 - X_B - Y_B$.

- The decision between options 1 and 2 will be made separately in each round, and it will be valid for both player A and B of a group. Both players A and B decide simultaneously and independently which option to choose.
  - If both players A and B of a group choose option 1 (half of the prize) , then this option will be selected in this round.
  - If one of two players A and B chooses option 1 and the other player chooses option 2, then in 9 out of 10 cases option 2 (competition about the whole prize) will be selected for both players and in 1 out of 10 cases option 1 (half of the prize) will be selected for both players.
  - If both players A and B choose option 2 (competition for the whole prize), then option 2 will be selected in this round.
- Please note, that you can change the choice, which you made at the beginning of the respective round. Even if, at the beginning of the round, you have declared that you would choose option 1, you can now choose option 2 and vice versa in case you have won the competition with player C.

2. Procedure

The experiment will consist of 24 identical rounds. In each round, you will have the same role (player A, B, or C). The other two players in your group will be randomly assigned to you in each round.

You will not know who the other players in your group are. All the decisions you make will remain anonymous, and any attempt to reveal your identity to anyone is prohibited. After the experiment, you will be asked to answer some questions, including some personal information (e.g., gender, age, major...). All the information you provide will be kept anonymous and strictly confidential.

At the end of today's experiment, we will randomly select 6 out of the 24 rounds to pay you. Your total earnings in those 6 rounds will be added up, converted to euros and paid to you in cash. This means that the earnings of all other rounds will not be paid to you and that you do not have to pay the expenditures of these rounds either. You will get to know which 6 out of the 24 rounds will be chosen only after finishing these 24 rounds.

Additionally to your earnings in these 6 selected rounds, you will receive 0.60 euros for each of the 24 rounds you have played.

Before the experiment starts, we will ask you some questions (which are related to the actions in the experiment) through the computer screen.