Cooperating Through Leaders☆

David K. Levine[1], Salvatore Modica[2], Aldo Rustichini[3]

Abstract

We study two by two games among homogeneous groups where players can choose to fight or cooperate and the average payoff from outcomes in which conflict occurs is smaller than the cooperation outcome. The novelty of our approach is that group choices are made under guidance of leaders who offer proposals to passive followers on the best course of action. *Accountability* of leaders is possible because of ex-post punishment which can be imposed by the groups, when the realized utility is smaller than that implicitly promised. *Competition* among leaders is possible if groups are willing to listen to more than one leader. We prove that in all games the equilibrium probability of cooperation is close to one if leaders compete and can be adequately punished.

Keywords: Plural Societies, Polarization, Social Conflict, Accountability, Political Equilibria.

☆First Version: January 6, 2022. We would like to thank Luigi Balletta, Rohan Dutta, and especially Sandeep Baliga for a short but illuminating exchange. DKL and SM gratefully acknowledge support from the MIUR PRIN 2017 n. 2017H5KPLL_01 and the Leverhulme Trust. AR thanks the U.S. Department of Defense, contract W911NF2010242

Email addresses: `david@dklevine.com` (David K. Levine), `salvatore.modica@unipa.it` (Salvatore Modica), `aldo.rustichini@gmail.com` (Aldo Rustichini)

[1]Department of Economics, RHUL and WUSTL

[2]Università di Palermo, Dipartimento SEAS

[3]Department of Economics, University of Minnesota

1. Introduction

We study two by two games played by groups through leaders; the strategic interaction is in the game played by the leaders. There are two types of leaders: *group* leaders, whose utility from the outcomes in the underlying game is identical to that of their group; and a *common* leader, whose utility is the average of the utilities of the two groups. The group leaders address only their own group; the common leader talks to both groups. Groups act on the basis of the suggestions given by the leaders. The main finding of the paper is that the presence of a common leader can lead groups to cooperate even when groups play the prisoners dilemma. The condition that favors cooperative outcomes is that the group leaders be adequately accountable to the groups.

We start with a symmetric two by two game played by homogeneous groups, where choices for each group are C (for cooperate) or F (for fight); the prisoner's dilemma is an example of such games. As in Baliga et al. (2011) the game is played through leaders, indeed by leaders; groups follow and can punish the leaders; and different groups may follow different leaders (unlike in elections, where there is a winning leader). The essential difference is that in this paper there may be several competing leaders addressing each group, and therefore each group must assess the competing leaders' proposals and choose which leader to follow.[4]

Although our model is one of full information and multiple groups, the specification of the leaders' proposals arises from the same idea of expectation shaping as in Hermalin (1998), where a single leader of a single group is the only one who comes to learn a payoff relevant signal, and acting on the basis of the signal shapes the followers' expectations (in such a way that leader imitation by followers is an equilibrium).[5]

In the present context expectations concern what the other group does. Therefore it makes sense to study proposals to group 1 of the form "Let us fight, the other group will surrender" - a proposal formally described below as FC. The proposal may be interpreted as consisting of a suggestion and an expectation.[6] Thus formally a leader's strategy in the leaders' game will be one of the four action profiles of the underlying game. In fact, think of the prisoners dilemma, a proposal like FC to group 1 also contains the implicit promise of a high payoff. Whence the idea of punishing the leader if such a payoff does not materialize.

To see how the leaders game is played consider a prisoners dilemma. Suppose that the common leader proposes cooperation to both group, "Cooperate, the other group will too", formally CC; suppose that this is the best outstanding proposal for group 2; and suppose that the group leader of group 1 proposes FC, "Let us fight, they will submit" to her group. Then group 1 chooses the

[4]In Baliga et al. leaders take actions on the underlying game, while in our case action C or F is taken by the groups; but this difference is inessential (the chosen leader could be delegated to take the chosen action). Also, in Baliga et al. groups are heterogeneous in that different members of the same group have different preferences on action profiles; but this only affects the way groups arrive at the decision.

[5]We thank the associate editor for pointing this out. The expectation shaping characteristic, the editor observed, suggests the term "spiritual" leaders.

[6]It turns out that any of the four possible outcomes in the underlying two by two game may occur as equilibrium outcomes of the leaders' game, so this interpretation is consistent.

group leader's proposal (because for them FC is better than CC) and fights; group 2 plays C (they chose the common leader so play as she recommends); therefore the outcome is FC; group members and leaders receive the game's payoffs, but the common leader is also punished by group 2 (because the realized payoff of group 2 is lower than that implicitly promised by CC).[7] By the same rules, if the common leader plays CC and the two group leaders mix then all profiles of the underlying game have positive probability in equilibrium. Notice that the leaders game is a $4 \times 4 \times 4$ game, not the same as the underlying two by two game.

Scope of the Model

One field of application of the model is civil wars, where the obvious candidate to play the role of common leader is the United Nations (UN). Doyle and Sambanis (2006) analyze 121 civil wars between 1945 and 1999. Of these, 99 ended with a military victory or a truce - that is, there was not a successful third parties intervention. Of the remaining 22, 14 ended with a negotiated settlement mediated by the UN. Our model suggests that the conflicts ending in fights may be Prisoner Dilemmas with low punishments for the group leaders.[8] The cases of successful peacebuilding on the other hand, as also Doyle and Sambanis reckon, my be interpreted as Stag Hunt games, where the common leader drives the parties to the existing good equilibrium.

We find cases of fruitful intervention by a figure which can be identified with a common leader in the context of internal politics as well. As an example we may think of Nelson Mandela, who helped bring the parties together in the racial conflict in South Africa. Mandela was jailed from 1964 to 1982 for opposing the apartheid system. The group leaders, both on the white and black side, were radical. In his inaugural speech as President in 1994 he declared: "The time for the healing of the wounds has come. The moment to bridge the chasms that divide us has come." His presidency gave a serious blow to the apartheid, and his government's results were positive on the economic side as well.[9]

On a more abstract level, the issue of polarization and potential conflict among groups is at the center of the book by Rabushka and Shepsle (1971). Observing that some societies are conflictual and others are not, they call for "a theory of political entrepreneurship." This paper hopefully contributes to building such a theory.

Related Literature

We are not the first to point out the possible merits of third-party intermediation in conflicts. Meirowitz et al. (2019) is a remarkable paper where a "neutral broker who does not favor either of the players" increases the chance that the conflicting parties achieve desirable outcomes. But in that

[7] This is actually an equilibrium of the leaders game for some parameter range.

[8] In civil wars one may think that the punishment, seen as cost of failure, is particularly high (in the limit, death); but in a war life is at risk whether you have promised victory or not, so that the additional punishment inflicted by followers is actually small.

[9] The quotation is from Mandela's Reuters obituary, taken from reuters.com/article/uk-mandela-obituary-idUKBRE9B417G20131206. Economy data: per capita GDP fell at the rate of 1.35% per year in the decade preceding 1994, and rose by 1.4% per year in the following decade, data from FRED.

case the third party is really just a mediator, to whom the parties are willing to reveal their private information. Our common leader acts on the basis of her preferences over the game's outcomes.

We have already said that our model is closely related to Baliga et al. (2011)[10]. We add the observation that if the group in that paper were homogeneous the model would reduce to our case with only group leaders. The same goes for Hermalin (1998) to which we owe the idea of expectation shaping.

In the political economy literature there are competing leaders, see for example Dewan and Squintani (2018). But competition is electoral; one of a number of competing leaders is elected and chooses for everyone. These models deal with different aspects of collective decision making.

There are other studies where delegation and/or leadership has a role. In Eliaz and Spiegler (2020), as here, a representative agent chooses among policy proposals and then selects and implements the one with the highest expected payoff (we explicitly model the proposers and their incentives, and allow each of them to address several representative followers).

Like us, Dutta et al. (2018) consider punishment of leaders, but their punishment is based *ex ante* considerations and there are no common leaders. Prat and Rustichini (2003) explore the idea that games among principals can be played through the mediation of agents who receive transfers conditional on the action chosen, to induce them to play one action rather than another.

Loosely related to the present context, Esteban and Ray (1994), Esteban et al. (2012) and Duclos et al. (2004), construct a general, well founded measure of polarization. The *salience* of ethnic conflict is analyzed in Esteban and Ray (2008). These models are tested against data in several follow up studies (for example in Esteban et al. (2012)).

Outline of the Paper

In the next two sections we set up the model. In Section 4 we present the case where the underlying two by two game is a Prisoners Dilemma. Section 5 is devoted to other games, and Section 6 concludes.

2. The Game Between Leaders

Interpreting players as homogeneous groups, we focus on the role of leaders in the collective decision making process. Leaders propose an action profile, like in "Follow me and fight, the enemy will surrender," or in "Follow me and cooperate, the other group will also cooperate." In each group followers choose a leader with the best offer for them and play according to the leader's suggestion; but they observe ex post their realized utility, and if it is less then what implicitly promised by the leader they punish her.

We turn to the formal model. There is a symmetric 2×2 game between groups. The groups are denoted by $k \in \{1, 2\}$, where each group has a representative follower. The followers choose actions $a_k \in \{C, F\}$, where C means cooperation and F means fight. Action profiles are denoted by $a \in A$,

[10] In turn in the tradition of Barro (1973)

and all group k members receive utility $u_k(a_k, a_{-k})$ where $-k$ denote the other group. We assume that payoffs are distinct:

$$\text{for all } k, a \neq a' \text{ implies } u_k(a) \neq u_k(a'). \tag{1}$$

These utility functions give rise to the *underlying game.* The family of underlying games studied in the paper is presented in the next section.

We now describe the leaders' game built upon the underlying game. There are three leaders $\ell \in \{0, 1, 2\}$: two *group leaders* $k = 1, 2$ who have the same interest as group $k = \ell$, and a *common leader* $k = 0$ who cares about both groups. Denoting by $U^\ell(a)$ the utility leader ℓ obtains from profile a, we have $U^\ell(a) = u_\ell(a)$ for the group leaders $\ell = 1, 2$, and we take $U^0(a) = (u_1(a)+u_2(a))/2$ for the common leader $\ell = 0$ who thus shares the preferences of both groups.

As indicated, each leader presents his plan of action to their potential followers. The group leaders make offers only to their own group, the common leader to both groups. Specifically, a leader strategy is an $s^\ell \in A$, that is, an action profile in the underlying game. This represents an offer and a promise to the potential followers. The common leader presents his offer to both groups: each group is asked to play s^0_k and promised if they do so that the other group will play s^0_{-k}. Group leaders address only their own group: "follow me and play s^ℓ_ℓ, the other group will play $s^\ell_{-\ell}$." The profile of leaders' strategies is $s \equiv (s^\ell)_{\ell\in\{0,1,2\}}$.

In addition to receiving direct utility the leaders may lose utility due to punishments by the followers. In each group they have the ability to impose a utility penalty P on their group leader, and $Q/2$ on the common leader (who then loses Q if punished by both groups).

In the bulk of the paper we assume that the follower of group k considers the proposal of the corresponding group leader and the one by the common leader, but as a benchmark we also analyze the case in which followers ignore the proposal of the common leader; in this case each group just follows their own group leader, there is no competition among leaders.

Among the proposals they consider the followers choose the one promising them the highest utility. That is, given a strategy profile s of the leaders, follower k chooses the proposal that maximizes $u_k(s^\ell)$ over the proposals they consider. The maximizer for group k is unique by assumption (1), though it may be proposed by more than one leader; denote it by $g^k(s) \in A$. Utility $u_k(g^k(s))$ is the utility group k expects. Group k then implement their part in the chosen strategy, that is they play $g^k(s)_k$. Therefore, given a profile of leaders' strategies s, the *implemented action profile* will be $g(s) \equiv \left(g^k(s)_k\right)_{k=1,2} \in A$. This determines the utility of the groups, $u_k(g(s))$, and the direct utility of the leaders $U^\ell(g(s))$.

After actions are implemented and direct utility accrues, followers of group k impose a punishment to the followed leaders when the obtained utility is less than the one promised. Note that no counterfactual reasoning is required by the followers: they simply compare the promised utility to the actual utility. Precisely, if $u_k(g^k(s)) < u_k(g(s))$ then group k punishes $\ell \in \{0, k\}$ such that $s^\ell = g^k(s)$, where the punishment is P if $\ell = k$ and $Q/2$ if $\ell = 0$.

The sum of the direct utility and the punishments obtained as we have just described determine the payoff of leader ℓ, $V^\ell(s)$, for any strategy profile s. We let $\mathbf{1}\{\mathfrak{c}\} = 1$ if condition $\mathfrak{c}$ is true and zero otherwise. Then the payoff of a group leader $\ell = 1, 2$ is

$$V^\ell(s) = U^\ell(g(s)) - P \cdot \mathbf{1}\{\ell = k \ \&\ g^k(s) = s^\ell \ \&\ u_k(s^\ell) < u_k(g(s))\} \tag{2}$$

and of the common leader

$$V^0(s) = U^0(g(s)) - (Q/2) \cdot \sum\nolimits_{k=1,2} \mathbf{1}\{g^k(s) = s^0 \ \&\ u_k(s^0) < u_k(g(s))\}. \tag{3}$$

We call the game played by the leaders $\ell \in \{0, 1, 2\}$, with $S^\ell = A$ and the utilities V^ℓ just defined, a *leaders game*. It is a finite game, hence an equilibrium in mixed strategies exists. We are interested in Nash equilibria in weakly undominated strategies of the leaders game. We call this a *leaders equilibrium*.

2.1. Competition Among Leaders Is Necessary for Their Effectiveness

If each group only considers proposals from their own group leader the outcomes of the leaders game are the same as in the underlying game:

Proposition 1. *For any leaders game, if each group only considers the proposal of their own group leader, then at the Nash equilibria of the leaders game the distributions of action profiles chosen by the groups are the same as those induced by the Nash equilibria of the corresponding underlying game.*

The proof, with all the others omitted in the text, is in Appendix A. Thus, without competition among leaders there are no improvements over the outcomes of the underlying game.[11] In the sequel competition among leaders will be always assumed: each group considers the proposals by their group leader and by the common leader.

3. The Underlying Games

As indicated, we restrict attention to symmetric two-by-two underlying games with representative followers $k = 1, 2$. As indicated, each player has two possible actions, C and F. We assume that if both play C they get a higher utility than if they both play F, that is $u_k(CC) > u_k(FF)$ for both k, and without loss of generality we set $u_k(CC) = 1$ and $u_k(FF) = 0$. Therefore, with $\lambda, \xi \in \mathbb{R}$, the family of underlying games we consider is the following (with $\lambda, \xi \notin \{0, 1\}$ by (1)):

	C	F
C	$1, 1$	ξ, λ
F	λ, ξ	$0, 0$

[11] The equilibrium strategy profile in the leaders game implementing a Nash equilibrium of the underlying game is not necessarily unique; but for any equilibrium in the leaders game the induced mixed action profile in the underlying game is unique.

We are interested in the conditions on the leadership structure that, when cooperation is desirable, make it an equilibrium. Observe that in the leaders game each leader has four possible strategies (the action profile of the underlying game) so the leaders game is $4 \times 4 \times 4$. In the next section we focus on the case where the underlying game is a Prisoners Dilemma: $\lambda > 1$ and $\xi < 0$. We take up the other cases in Section 5.

The "aggressive" proposal by group leader k "we play F and they play C" will be denoted by $F^k C^{-k}$. This is FC for leader 1 and CF for leader 2.

4. Prisoners Dilemma as Underlying Game

We compute a couple of payoff entries in the $4 \times 4 \times 4$ leaders game for illustration. Suppose the leaders play (CC, CF, FC): in this case the common leader's proposal prevails in both groups, the implemented profile is CC and all leaders get 1. Consider now the profile (CC, FC, CC). Here leader 1 wins group 1 because $u_1(FC) > u_1(CC)$; group 2 receives the proposal CC from both their leader and the common leader; so group 1 plays F and group 2 plays C, that is the implemented action profile is FC; leader 1 gets λ; the common leader and leader 2 are followed and punished by group 2 (since the group gets ξ against a promise of 1); so the common leader gets utility of $(\lambda + \xi)/2 - Q/2$ and leader 2 gets $\xi - P$.

Reduction

In the PD case the leaders game can be considerably simplified. For the group leaders, the strategies CC and $C^k F^{-k}$ are weakly dominated by FF. For the common leader, the strategies CF and FC are then weakly dominated by FF for all $P > 0$. So the analysis is reduced to the game where the group leaders only play $F^k C^{-k}$ or FF and the common leader plays only CC or FF. This is proved in Lemmas 1 and 2 in Appendix A. In summary, the game is reduced to a simpler game with three players, each player with two actions. The simplified game is presented in table 1, where the three payoffs in each entry are naturally ordered with the leaders' index (first common then the other two).

Table 1: The game after elimination of weakly dominated strategies. The left panel shows utilities when the common leader plays CC; in the right panel are utilities when the common leader plays FF.

CC	CF	FF
FC	$0, -P, -P$	$\frac{\lambda+\xi-Q}{2}, \lambda, \xi$
FF	$\frac{\lambda+\xi-Q}{2}, \xi, \lambda$	$1, 1, 1$

FF	CF	FF
FC	$0, -P, -P$	$0, -P, 0$
FF	$0, 0, -P$	$0, 0, 0$

Nash Equilibria of the leaders game

So when the underlying game is a PD the leaders game reduces to the $2 \times 2 \times 2$ game where a strategy profile of the three players is a vector of the form (q, p_1, p_2), q being the probability that the common leader plays CC ($(1-q)$ that he plays FF), and p_k the probability that the k group leader plays $F^k C^{-k}$ (FF played with probability $1 - p_k$).

The next result characterizes the equilibria of the leaders' game. As indicated some of the equilibria are mixed, and the mixing probabilities are given in the next two displayed equations. Equation (4) below describes a mixture in the reduced game, where the common leader plays CC for sure and each of the group leaders plays F^kC^{-k} with probability $\tilde{p}$.

$$\tilde{q} = 1\,, \qquad \tilde{p} \equiv \frac{\lambda - 1}{\lambda - 1 + P + \xi} \tag{4}$$

Note that $\tilde{p}$ converges to 0 as P becomes large so the induced outcome converges to cooperation as P becomes large. The pair $(\hat{q}, \hat{p})$ in (5) below describes a fully mixed profile (again the the group leaders play the same strategy).

$$\hat{q} \equiv \frac{P}{P + \lambda - 1 - \hat{p}(P + \lambda + \xi - 1)}\,, \qquad \hat{p} \equiv \frac{1}{1 + Q - (\lambda + \xi)} \tag{5}$$

The formal result is the following:

Proposition 2. *In the leaders game with prisoners dilemma underlying game:*
If $P < -\xi$ *the equilibria are all* $(q, 1, 1)$ *for* $-\frac{P}{\xi} < q \leq 1$*, with outcome* FF*.*
If $P > -\xi$*:*
If $Q < \lambda + \xi$ *the equilibria are* $(1, 0, 1), (1, 1, 0)$*, with outcome* F^kC^{-k}*, and* $(1, \tilde{p}, \tilde{p})$
If $Q > \lambda + \xi$ *and* $P + \xi < (\lambda - 1)(Q - (\lambda + \xi))$ *the unique equilibrium is* $(\hat{q}, \hat{p}, \hat{p})$
If $Q > \lambda + \xi$ *and and* $P + \xi > (\lambda - 1)(Q - (\lambda + \xi))$ *the unique equilibrium is* $(1, \tilde{p}, \tilde{p})$*.*

We shall show how the fully mixed $(\hat{q}, \hat{p}, \hat{p})$ equilibrium arises in the next proposition. In Appendix A we identify all the remaining equilibria in the game in a sequence of lemmas; the analysis is organized considering three possible cases for the value of q, namely $q = 0$, $q = 1$ and then $q \in (0, 1)$. We now comment the result just stated.

If the punishment of the group leaders is small, the only possible equilibrium outcome is conflict, just as in the underlying game. Consider now what happens for $P > -\xi$. If the common leader's punishment Q is small there are two pure strategy asymmetric equilibria with outcome F^kC^{-k} which are robust to increasing the value of P: the common leader plays CC for sure because she is not afraid of being punished; and given this, if group leader k plays F^kC^{-k} the other group leader cannot do better than play FF (to beat the common leader she would have to play F^kC^{-k} as well which would end up in outcome FF and punishment). There is also the mixed equilibrium $(1, \tilde{p}, \tilde{p})$, where the probability of the group leaders playing F^kC^{-k} go to zero as P grows and the equilibrium probability of cooperation goes to 1.

The same "good" $(1, \tilde{p}, \tilde{p})$ equilibrium obtains uniquely for $Q > \lambda + \xi$ if P is large enough, that is, if the group leaders bear adequate responsibility for their actions.

Punishing the common leader harshly and leaving the group leaders relatively free - $Q > \lambda + \xi$ and $0 < P + \xi < (\lambda - 1)(Q - (\lambda + \xi))$ - is not as good. In the $(\hat{q}, \hat{p}, \hat{p})$ equilibrium the group leaders will play FF with high probability, but the will also play FF with positive probability, indeed larger than $(\lambda - 1)/(\lambda - 1) + P$.

The mixed equilibria arise naturally in these leaders games. To see why observe that clearly the proposal FF by both group leaders is beaten by the CC proposal of the common leader, and this proposal is in turn easily beaten by the aggressive proposals F^kC^{-k} of the group leaders. However the two group leaders cannot both play F^kC^{-k} for sure, because they would anticipate the FF outcome and the consequent punishment imposed by followers. This gives rise to a mixed equilibrium: group leaders randomize between aggressive play F^kC^{-k} and a conservative FF; and the common leader may mix too, between proposing cooperation CC and effectively opting out by playing FF. In particular this is the case when $Q > \lambda + \xi$ and $P + \xi < (\lambda - 1)(Q - (\lambda + \xi))$, as we now show.

Proposition 3. *An equilibrium with* $\min\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\} < q < \min\{\max\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\}, 1\}$ *exists, with the mixed strategy* $(\hat{q}, \hat{p}, \hat{p})$ *defined in equation* (5)*, if and only if* $Q > \lambda + \xi$ *and* $P + \xi < (\lambda - 1)(Q - (\lambda + \xi))$.

Proof. For q to be part of a fully mixed equilibrium the common leader has to be indifferent between CC and FF, which is true if and only if $(1-p)^2 + 2p(1-p)\frac{\lambda+\xi-Q}{2} = 0$ that is if

$$p = \frac{1}{1 + Q - (\lambda + \xi)} \equiv \hat{p} \tag{6}$$

Note that $0 \leq \hat{p} \leq 1$ if and only if $Q \geq \lambda + \xi$. On the other hand the indifference for group leader 1 (for example) between FC and FF requires:

$$-pP + (1-p)(q\lambda - (1-q)P) = pq\xi + (1-p)q$$

which is rewritten as:

$$p = \frac{P + \lambda - 1 - P/q}{P + \lambda + \xi - 1} \equiv f(q) \tag{7}$$

Combining equations 6 and 7 we conclude that an equilibrium with q in the range exists if $0 < q < 1$, $f(q) = \hat{p}$ and

$$\min\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\} < q < \max\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\}.$$

Observe that $\frac{P}{P+\lambda-1} < -\frac{P}{\xi}$ if and only if $P + \lambda + \xi - 1 > 0$, in which case f is strictly increasing; and f is strictly decreasing if the inequalities are reversed. Since $f(\frac{P}{P+\lambda-1}) = 0$ and $f(-\frac{P}{\xi}) = 1$, there is unique $\hat{q}$ in the given range such that

$$f(\hat{q}) = \hat{p}. \tag{8}$$

it is easy to check that this $\hat{q}$ is indeed the value in equation (5).

If $\max\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\} < 1$ - that is if $P < -\xi$ - we are done. For $P > -\xi$ we have $\max\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\} = -\frac{P}{\xi} > 1$, hence in this case we must check whether an equilibrium exists with $\frac{P}{P+\lambda-1} < q < 1$. Since

$f(\frac{P}{P+\lambda-1}) = 0$ and in this case f is increasing, the equilibrium exists if $f(1) > \hat{p}$, that is if

$$\frac{\lambda-1}{P+\lambda+\xi-1} > \frac{1}{Q-(\lambda+\xi-1)}$$
$$P+\xi < (\lambda-1)(Q-(\lambda+\xi))$$

where we have used that (since $P > -\xi$) $P+\lambda+\xi-1 > 0$ and that $Q-(\lambda+\xi-1) > 0$ because $\hat{p} \in (0,1)$. If the inequality is false then the conditions for the equilibrium $(1, \tilde{p}, \tilde{p})$ are met. □

Comparison with Correlated Equilibria

The leaders game built over an underlying game shares important features with the correlated equilibria of that underlying game: in both cases, thanks to a form of mediation, better outcomes than Nash equilibria can obtain; and in both solution concepts, leaders or the mediator suggest to followers an action profile, and followers respond. But the differences are deeper than the similarities.

In correlated equilibria the single mediator has no direct interest in the outcome; followers respond strategically to the action suggested privately to each, by updating the posterior on the action profile played by others, and would never want to punish the mediator. In the leaders game there are competing leaders with a direct interest in the outcome, so that their utility is affected by the action of the followers; the latter respond to the leaders' suggestions by choosing the best action profile from their point of view, and typically punish the chosen leaders with positive probability in equilibrium. Most importantly, although action profiles are implemented by the groups, the strategic interaction is among the leaders, not between the players of the underlying games.

We compare the sets of equilibrium action profiles taking as measurement of welfare the average utility of players in the underlying game (ignoring the welfare of the leaders which may include punishments).

Proposition 4. *If $P < -\xi$ the leaders equilibrium payoff is the same as the correlated payoff; otherwise it is strictly higher.*

Proof. The correlated payoff is zero. For $P < -\xi$ the leaders equilibrium payoff is zero too. For $P > -\xi$, consider first $Q < \lambda+\xi$; this implies $\lambda+\xi > 0$. In the asymmetric leaders equilibria the outcome is $F^k C^{-k}$ with average payoff $(\lambda+\xi)/2 > 0$.

The condition for average payoff in $(1, \tilde{p}, \tilde{p})$ to be positive is $\tilde{p}^2 * 0 + \tilde{p}(1-\tilde{p})(\lambda+\xi) + (1-\tilde{p})^2 * 1 > 0$, equivalently $P+\xi > -(\lambda-1)(\lambda+\xi)$. But the $(1, \tilde{p}, \tilde{p})$ equilibrium is either with $Q < \lambda+\xi$ whence $\lambda+\xi > 0$ and we are done; or $P+\xi > (\lambda-1)(Q-(\lambda+\xi))$ which implies the condition above.

Consider lastly $(\hat{q}, \hat{p}, \hat{p})$. If the common leader plays FF the possible outcomes are FC and FF both with positive probability hence average payoff is positive. If the common leader plays CC then the condition becomes as above $\hat{p}(\lambda+\xi) + (1-\hat{p}) > 0$ which is $Q > 0$. □

5. The Other Games

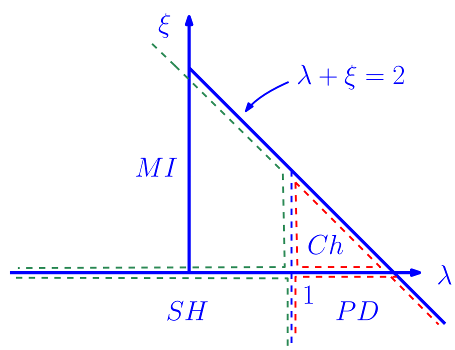
Let us go back to the family of two by two games with which we started in Section 3:

	C	F
C	$1,1$	ξ,λ
F	λ,ξ	$0,0$

Considering the combination of the two possible inequalities between λ and 1 on the one hand and ξ and 0 on the other, we have two sets of possible games. One has $\lambda > 1$, so the choice of F against C of the opponent is better than the choice of C: these are Prisoner's Dilemma if $\xi < 0$ and Chicken if $\xi > 0$. We call these *conflict games*, because CC is not a Nash equilibrium of the game. The other set of possible games has $\lambda < 1$, so the choice of C against C of the opponent is better than the choice of F: they are Stag Hunt if $\xi < 0$ and Mutual Interest if $\xi > 0$. We call them *cooperation games*, because they admit CC as an equilibrium.

Assume that unilateral deviations from the best common action profile reduces average group welfare: $u_k(CC) > [u_1(FC) + u_2(FC)]/2$ for both k, that is $\lambda + \xi < 2$. Together with $u_k(CC) > u_k(FF)$, this characterizes games where average players' payoff is highest at outcome CC. In this sense these are the games where conflict is detrimental. The family of underlying games can be visualized in (λ, ξ) space as in Figure 1.

Figure 1: **The family of games.** All payoffs are below the $\lambda + \xi = 2$ line. To the left of the $\lambda = 1$ line: above the horizontal axis ($\xi > 0$) there is Mutual Interest and below it is Stag Hunt. These are the cooperation games. To the right of $\lambda = 1$: above the axis we have Chicken, below it is Prisoners Dilemma. These are the conflict games.



The family covers models of conflict over a public good in the spirit of Esteban and Ray (2011) (see also Esteban and Ray (1994) and Esteban et al. (2012)); and of strategic complementarity as in Baliga and Sjöström (2020) (see also Baliga et al. (2011)). Details are in the WP version.

We have not covered the cooperation games and Chicken, to which we turn.

5.1. Leaders Equilibria in Cooperation Games

With underlying cooperation games the leaders games are necessarily efficient:

Proposition 5. *With a common leader, in the Mutual Interest and Stag Hunt games there is a unique leadership equilibrium for any value of P, with implemented action profile CC.*

Proof. The CC outcome is the most preferred by the common leader and she can guarantee that outcome by proposing it, because $u_k(F^kC^{-k}), u_k(F^kF^{-k}) < 1$ so the group leaders best response to CC by the common leader is to propose C to their group. □

5.2. The Chicken Game

It follows from Proposition 1 that the pure equilibria of the underlying chicken game survive as leadership equilibria when there is no common leader. Competition with the common leader is not sufficient to change this fact:

Proposition 6. *The outcomes FC and CF of the underlying game are equilibrium outcomes of the leaders game for all P and Q.*

However, for moderate Q and larger P the cooperative mixed equilibrium emerges:

Proposition 7. *For $Q \leq \min\{\lambda+\xi, (P+\lambda(\lambda-1+\xi))/(\lambda-1)\}$ the profile $(1, \tilde{p}, \tilde{p})$ is an equilibrium.*

Proofs of these statements are in Appendix A, and more on this game are contained in the WP version of the paper. The full analysis of equilibria and of the relation to the correlated equilibria is a bit more involved than in the PD case, but its results do not add to nor contradict those we already know.[12]

6. Conclusions

We have studied how political leadership can fundamentally alter outcomes in societies with group conflict when leaders are accountable to groups. We rely on a model of leadership which may be useful in general environments: given an underlying game among players, we construct a game among leaders in which the leaders' strategies are action profiles proposed by each leader to the society of players-followers. Followers choose among the proposals to maximize their utility.

The main insight derived from the analysis of our model is that conflict in polarized societies can be substantially reduced, *under appropriate conditions*, thanks to the mediation of interested leaders. The existence of leaders by itself cannot accomplish anything useful: the equilibrium outcomes are the same as in the game with no leaders (Proposition 1). With common leaders, our analysis has identified two main forces: competition among leaders and accountability. If there is competition among leaders, then in general cooperation and good outcomes are possible when the accountability of leaders is sufficiently large. In the limit of high accountability, full cooperation may be realized.

Our setup relies on simplifying assumptions, and some of these assumptions may be in contrast with important real world regularities. In the model, leaders share precisely the utility of their constituencies, so their incentives are perfectly in line with those of the groups. Leaders do not have a political career to pursue, nor derive utility from being leaders. Leaders cannot profit directly or indirectly on their position. The common leader in particular is assumed to share the interests

[12]We have $Q = P$ in the WP; but the nature of equilibrium for large P and Q does not change. The comparison with correlated payoff (which goes through unaltered with $Q \neq P$ is more delicate because there are high correlated payoffs. Incidentally, for the other games. The comparison is trivial in the case of the mutual interest game: both solution concepts uniquely predict the efficient outcome (where both groups get utility of 1). In the stag hunt, the outcome of the leaders' game is the efficient outcome for any value of P, equal to the highest correlated payoff.

of society as a whole. Followers, on their part, make the task of the leaders as easy as possible: they hear what the leaders say, and take their promises at face value, with the understanding that punishment will follow if the leader does not deliver. Finally, punishment must be sufficiently high for cooperation to arise. Fortunately, our analysis makes clear the leaders' role, so it can be taken to provide the best case scenario for possible positive effects of mediation in group conflict. Systematic empirical research will have to decide which are the realistic ranges of the losses groups can impose on leaders.

The behavior of followers in our model is extremely simplified, but we do not consider assumption of unsophisticated behavior completely unrealistic: in large and complex societies, understanding the structure of payoffs from social actions is at the same time very hard (because societies are complex) and unrewarding (because the action of each player - even when he has acquired enough information to evaluate the best choice - is in itself irrelevant). Thus a first simple approximation is to assume, as we do, that followers just consider the promised utility, and choose the highest.

A natural extension of the model presented here, providing a more realistic behavior of followers, would provide a foundation of their behavior on a model of information acquisition on relevant parameters affecting the utility of players. This information is hard to gather, so in our model it is delegated to leaders or parties, which can do that through costly effort, and then send messages (for example, political programs) to the entire society. Followers may then interpret the signals sent in the light of what they know and choose rationally the best action. In a different context, a similar idea is presented in Matějka and Tabellini (2021).

Appendix A. Omitted Proofs

We collect proofs omitted from the text, including the relative statement.

Proof of Proposition 1

This result is actually true for any leaders game, with any number of groups, and even without the assumption (1). Observe that the model trivially extends to the case of K groups: just take $k, \ell \in \{1, 2, \ldots, K\}$ instead of $k, \ell \in \{1, 2\}$. Proving the statement for this more general case requires no additional effort, so we state it for this case:

Statement. *For any leaders game, if each group only considers the proposal of their own group leader, then at the Nash equilibria of the leaders game the distributions of action profiles chosen by the groups are the same as those induced by the Nash equilibria of the corresponding underlying game.*

Proof. For a mixed strategy $\hat{\sigma}^k$ of leader k we let $\hat{\sigma}^k_{A_k}$ the induced distribution on A_k. Our first claim is that

$$\forall \hat{\alpha} \in NE(UG) \exists \hat{\sigma} \in NE(LG) : \forall k,\ \hat{\sigma}^k_{A_k} = \hat{\alpha}_k, \tag{A.1}$$

where $NE(UG)$ and $NE(LG)$ denote the sets of Nash equilibria of the underlying game and leaders' game respectively. Consider a mixed action profile $\hat{\alpha} \in NE(UG)$. For any action $b_k \in \text{supp}(\hat{\alpha}_k)$ choose

$$a_{-k}(b_k) \in \text{argmin}_{c_{-k} \in A_{-k}} u_k(b_k, c_{-k}). \tag{A.2}$$

Define now $\hat{\sigma}^k$ as:

$$\hat{\sigma}^k(a) \equiv \sum\nolimits_{a_k \in A_k} \hat{\alpha}(a_k) \delta_{(a_k, a_{-k}(b_k))}(a). \tag{A.3}$$

If all leaders j different form k follow the strategy defined in (A.3) then leader k is facing the probability on A^{-k} given by $\hat{\alpha}_{-k}$. Consider now a possible strictly profitable deviation $\hat{\tau}^k$ from $\hat{\sigma}^k$. Since by following $\hat{\sigma}^k$ the k leader incurs no punishment cost, the increase in net utility to leader k from $\hat{\tau}^k$ is at least as large as the increase in direct utility, and the direct utility is the utility of the followers. Thus $\hat{\tau}^k$ would have a marginal on A_k that is a profitable deviation for player k from $\hat{\alpha}_k$ against $\hat{\alpha}_{-k}$, a contradiction with $\hat{\alpha} \in NE(UG)$.

The second claim is:

$$\forall \hat{\sigma} \in NE(LG), \text{ if } \hat{\alpha}_k \equiv \hat{\sigma}^k_{A_k}, \text{ then } \hat{\alpha} \in NE(UG). \tag{A.4}$$

Consider in fact a strictly profitable deviation β_k from $\hat{\alpha}_k$ of a player k in the underlying game. Extend β_k to a profitable deviation τ^k in the leaders game of the k^{th} group leader following the construction in equations (A.2) and (A.3). This deviation would insure for group leader k, the same utility as β_k, which would then be higher than $\hat{\sigma}^k$, since the direct utility of τ^k is higher than $\hat{\sigma}^k$, and its punishment cost is zero; a contradiction with the assumption that $\hat{\sigma}^k$ is a best response. □

Lemmas for the Reduction

Lemma 1. *For group-k leader the strategies CC and C^kF^{-k} are weakly dominated by FF.*

Proof. We let $k = 1$. Fix any profile s^{-k} of the other leaders.

Consider CF first. Suppose that $g(CF, s^{-k})_1 = F$; then group 1 must have accepted a proposal FF or FC by the common leader, so that by playing CF or FF group-1 leader gets the same payoff (λ or 0, no punishment). Suppose $g(CF, s^{-k})_1 = C$; then the common leader must have proposed CF as well and group-1 leader gets $\xi < 0$, while in this case by proposing FF she gets 0 and no punishment. The argument in the CC case is analogous.[13] □

In view of this lemma we may assume that group leader k plays only F^kC^{-k} or FF; we let p_k denote the probability of F^kC^{-k}.

Lemma 2. *The probability that the common leader plays either CF or FC is zero.*

Proof. We do it for CF. This proposal is rejected by group 1 who will play F, and accepted for sure by group 2 who will play F and punish the common leader. She is better off by playing FF (strictly if $Q > 0$). □

Proof of Proposition 2

Statement. *In the leaders game with prisoners dilemma underlying game:*

If $P < -\xi$ the equilibria are all $(q, 1, 1)$ for $-\frac{P}{\xi} < q \leq 1$, with outcome FF.

If $P > -\xi$

If $Q < \lambda + \xi$ the equilibria are $(1, 0, 1), (1, 1, 0)$, outcome F^kC^{-k} and $(1, \tilde{p}, \tilde{p})$

If $Q > \lambda + \xi$: equilibrium is $(\hat{q}, \hat{p}, \hat{p})$ if $P + \xi < (\lambda - 1)(Q - (\lambda + \xi))$, otherwise $(1, \tilde{p}, \tilde{p})$.

Proof. There is no equilibrium with $q = 0$ for any $P > 0$, from Lemma 3. Consider $P < -\xi$. We have equilibrium $(1, 1, 1)$ from Lemma 4; from Lemma 7 we have $(q, 1, 1)$ for $\max\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\} < q < 1$; and from Lemma 8 the same equilibrium for $-\frac{P}{\xi} < q < \frac{P}{P+\lambda-1}$ if that interval is nonempty. The last two give $(q, 1, 1)$ for $-\frac{P}{\xi} < q < 1$. Therefore if $P < -\xi$ we have $(q, 1, 1)$ for $-\frac{P}{\xi} < q \leq 1$, as in the statement. Turn to $P > -\xi$. For $Q < \lambda + \xi$ Lemma 5 gives $(1, 0, 1)$, $(1, 1, 0)$ and $(1, \tilde{p}, \tilde{p})$; for $Q > \lambda + \xi$ Lemma 3 gives $(\hat{q}, \hat{p}, \hat{p})$ if $P + \xi < (\lambda - 1)(Q - (\lambda + \xi))$, and Lemma 5 gives $(1, \tilde{p}, \tilde{p})$ for the reverse inequality. □

Lemmas for Proposition 2

We concentrate on the interesting cases in which the relevant inequalities among combinations of parameters hold strictly.

[13] Here it is. Consider CC and suppose first $g(CC, s^{-k})_1 = F$; then group 1 must have accepted a proposal FC by the common leader, and therefore CC and FF yield the leader the same payoff. Suppose $g(CC, s^{-k})_1 = C$ so that her proposal is accepted; the competing offers may have been CC, CF or FF; if all other proposals are CC then her payoff does not change if she plays FF; if there is a CF or an FF by some $\ell \neq 1$ then group-1 leader is strictly better off by playing FF (she gets zero, while with CC she gets $\xi - P$).

Equilibria with $q = 0$. We start with the fact that there are no such equilibria.

Lemma 3. *If* $P > 0$*, there is no equilibrium with* $q = 0$

Proof. If the common leader sets $q = 0$ then the leaders' game is the right panel of table 1 (ignoring the common leader's utility). This game has a unique Nash Equilibrium in dominant strategies in which both group leaders play FF. At this profile of actions of group leaders, CC yields 1, and FF yields 0, to the common leader, hence setting $q = 1$ is the best response. □

Equilibria with $q = 1$. We deal in turn with small P and larger P:

Lemma 4. *If* $P < -\xi$ *then there is a unique equilibrium with* $q = 1$*, with* $(q, p_1, p_2) = (1, 1, 1)$*.*

Proof. Since $\lambda > 1$ and $\xi < -P$, if $q = 1$ we see from table 1 that the action FC is dominant for the first group leader CF for the second). When group leaders play the action profile (FC, CF) then both CC and FF give utility 0 to the common leader, hence $(1, 1, 1)$ is the only equilibrium with $q = 1$. □

Lemma 5. *If* $P > -\xi$*:*

1. *There are two equilibria where group leaders play pure strategies:* $(q, p_1, p_2) \in \{(1, 0, 1), (1, 1, 0)\}$ *if and only if* $Q < \lambda + \xi$*. In these equilibria the outcome is* FC *or* CF*.*
2. *There is an equilibrium where group leaders play a mixed strategy if and only if:*

$$P + \xi > (\lambda - 1)(Q - (\lambda + \xi)) \tag{A.5}$$

The mixed strategy is $\tilde{p}$ *in equation (A.6).*

Proof. If $P > -\xi$ then at $q = 1$ the game among group leaders has three equilibria, the two pure strategies (FF, CF), (FC, FF) and a mixed one with:

$$p_1 = p_2 = \frac{\lambda - 1}{\lambda - 1 + P + \xi} \equiv \tilde{p} \tag{A.6}$$

Note that $\lambda > 1$ and our assumption that $P > -\xi$ insure that $\tilde{p} \in (0, 1)$.

We first consider the possible equilibria where group leaders play pure strategies:

1. If $\lambda + \xi - Q > 0$ then there are two equilibria, $(q, p_1, p_2) = (1, 0, 1), (1, 1, 0)$. This follows because CC gives $(\lambda + \xi - Q)/2$, while FF gives 0 to the common leader.
2. If $\lambda + \xi - Q < 0$ then there are no equilibria $(1, p_1, p_2)$ with $p_i \in \{0, 1\}$, because in this case the utility to the common leader from CC is lower than the one from FF.

We then consider the the possible equilibria where group leaders play a mixed strategy. At any mixed strategy profile (p, p), with $p \in (0, 1)$ of the group leaders the common leader playing CC gets

$$(1 - p)^2 + 2p(1 - p)\frac{\lambda + \xi - Q}{2}$$

and at $\tilde{p}$ this is larger than 0 (hence CC better than FF) if and only if (A.5) holds. □

Equilibria with $q \in (0,1)$. To set up the analysis we assume that the common leader is playing q and compare a group leader's payoffs from FC and FF for each of the two possible strategies CF and FF of the other group leader. From Table 1 we see that in the first case FC is better than FF if and only if

$$q > -P/\xi \tag{A.7}$$

while in the second case FC is better than FF if and only if

$$q > \frac{P}{P+\lambda-1} \tag{A.8}$$

In lemmas 6 and 7 we consider the two extreme possible cases for q:

Lemma 6. *There is no equilibrium with* $0 < q < \min\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\}$.

Proof. The condition on q implies that the action FF is dominant for both group leaders, but the common leader's best response to (FF, FF) is $q = 1$. □

Lemma 7. *There is an equilibrium with any q such that* $\max\{-\frac{P}{\xi}, \frac{P}{P+\lambda-1}\} < q < 1$, *of the form* $(q, 1, 1)$.

Of course the set of such q's may be empty; this is the case when $P > -\xi$.

Proof. The condition on q implies that (FC, CF) is dominant for the group leaders, and at this profile the common leader gets zero both from CC and FF; the conclusion follows. □

Next we consider the intermediate cases for the values of q. At these values of q the game between the group leaders has three equilibria, two pure strategies and one mixed. We deal with pure strategies of group leaders in lemma 8. Observe that $\frac{P}{P+\lambda-1} < -\frac{P}{\xi}$ iff $P > 1 - (\lambda + \xi)$.

Lemma 8.

1. *If* $\frac{P}{P+\lambda-1} < q < -\frac{P}{\xi}$ *then there is no equilibrium with* $p_i \in \{0, 1\}$ *(that is, with group leaders playing pure strategies)*
2. *For any value* $-\frac{P}{\xi} < q < \frac{P}{P+\lambda-1}$, *there is an equilibrium in pure strategies for group leaders of the form* $(q, 1, 1)$.

Proof. For the first case, the two pure strategy equilibria in the resulting group leaders game are (FF, CF) and (FC, FF); consider the first (the second is analogous). In this case CC gives $\frac{\lambda+\xi-Q}{2}$, and FF gives 0. Considering only the cases in which the inequalities holds strictly, it follows that the best response of the common leader to this strategy profile of the group leaders is either $q = 0$ or $q = 1$, hence not in the open interval $(0, 1)$.

For the second case, with q in that range the two pure strategy equilibria in the group leaders' game are (FC, CF) and (FF, FF). At the first profile the common leader gets zero from either CC or FF whence the equilibria; at the second one the common leader gets 1 from CC and zero from FF, hence the best reply is not interior. □

The case of fully mixed equilibrium is proven in the text, Proposition 3.

Proof of Proposition 6

Statement. *The outcomes FC and CF of the underlying game are equilibrium outcomes of the leaders game for all P and Q.*

Proof. We focus on FC. Denote by $BR^{\ell}(a^0, a^1, a^2)$ the best response of leader ℓ to the profile (a^0, a^1, a^2). The proof has three parts, for $\ell \in \{0, 1, 2\}$:

$$FC \in BR^{\ell}(FC, FC, FC) \tag{A.9}$$

So in each step we proceed from the assumption that the other leaders are playing FC and consider the best response of the leader under consideration. We then examine the expected utility from the different possible choices of the leader under consideration, and claim that conclude that his best response is FC.

Consider first $\ell = 2$. Given $a^1 = a^0 = FC$, group 1 will choose F no matter what the other leader offers, because this is the largest utility it can receive, and group 2 has the proposal FC of the common leader. Considering the possible choices of a^2: CC gives a utility $\xi - P$ (because $\xi < 1$, so group 2 will follow leader 2 , but the outcome is FC with payoff (λ, ξ) rater than the implicit promise $(1, 1)$ of leader 2, and hence leader 2 will be punished). CF gives a utility $-P$ (group 2 will follow leader 2 and play F but the outcome is then $(0, 0)$ and so leader 2 gets the 0 utility and the punishment because the realized 0 is smaller than the promised λ). FC gives a utility ξ (because both common leader and leader 2 promise the same utility profile). Finally, FF gives a utility ξ (because the associated utility vector is $(0, 0)$, and common leader is promising ξ). Our claim follows.

Consider next $\ell = 0$. We proceed noting that $a^1 = a^2 = FC$, and thus group 1 is choosing F.

CC gives a utility of $\frac{\lambda+\xi-Q}{2}$ (because group 1 will choose F, following the group leader, while group 2 will choose C, following the common leader, expecting utility 1 rather than the ξ proposed by the group leader; thus the outcome is FC; thus the common leader gets utility $\frac{\lambda+\xi}{2}$, minus $Q/2$ because he is punished by group 2). CF gives a utility of $-Q/2$ (because group 1 will follow leader 1, and group 2 will follow the common leader and play F expecting λ; but the outcome is FF with average utility of groups equal to 0, so the common leader is punished by group 2). FC gives a utility of $\frac{\lambda+\xi}{2}$ (because all leaders are proposing the same action profile). FF gives a utility of $\frac{\lambda+\xi}{2}$ (because the proposal of the common leader will be ignored).

Consider finally $\ell = 1$. Assuming $a^0 = a^2 = FC$, we note that group 1 is considering the utility λ from the common leader (with choice C), and group 2 is considering the utility ξ from both common leader and group leader 2. Group 1 is choosing F, following the common leader, no matter what group leader 1 is going to propose. The choice $a^1 = FC$ gives leader 1 a utility of λ (group 1 is choosing F, because this is then the only proposal they receive, and group 2 is choosing C); but λ is the largest possible utility, hence FC is a best response of group leader 1. □

Proof of Proposition 7

Statement. *For* $Q \leq \min\{\lambda + \xi, (P + \lambda(\lambda - 1 + \xi))/(\lambda - 1)\}$ *the profile* $(1, \tilde{p}, \tilde{p})$ *is an equilibrium.*

Proof. The utility matrix when the common leader plays CC is the following:

	CC	FC	CF	FF
CC	$1,1,1$	$1,1,1$	$\frac{\lambda+\xi-Q}{2}, \xi-P, \lambda$	$1,1,1$
CF	$1,1,1$	$1,1,1$	$\frac{\lambda+\xi-Q}{2}, \xi, \lambda$	$1,1,1$
FC	$\frac{\lambda+\xi-Q}{2}, \lambda, \xi-P$	$\frac{\lambda+\xi-Q}{2}, \lambda, \xi$	$0,-P,-P$	$\frac{\lambda+\xi-Q}{2}, \lambda, \xi$
FF	$1,1,1$	$1,1,1$	$\frac{\lambda+\xi-Q}{2}, \xi, \lambda$	$1,1,1$

Consider first a group leader, given the others' strategies: if she plays FF he gets

$$p\xi + 1 - p = 1 - p(1-\xi)$$

while if she plays FC she gets

$$-pP + (1-p)\lambda = \lambda - p(\lambda + P)$$

so indifference between FF and FC holds if and only if:

$$p = \frac{\lambda - 1}{\lambda - 1 + \xi + P} = \tilde{p}.$$

This is smaller than 1 because $\xi > 0$. As we see from the utility matrix CF yields the same utility as FF and CC is weakly worse.

Consider now the common leader. The reduced utility matrix when she plays CC is this

	CF	FF
FC	$0,-P,-P$	$\frac{\lambda+\xi-Q}{2}, \lambda, \xi$
FF	$\frac{\lambda+\xi-Q}{2}, \xi, \lambda$	$1,1,1$

so by playing CC she gets

$$(1-p)\left(1 + p(\lambda - 1 + \xi - Q)\right).$$

This value is strictly positive if

$$1 + \frac{\lambda - 1}{\lambda - 1 + \xi + P}(\lambda - 1 + \xi - Q) > 0$$
$$\frac{(\lambda - 1 + \xi + P) + (\lambda - 1)(\lambda - 1 + \xi - Q)}{\lambda - 1 + \xi + P} > 0$$
$$(\lambda - 1 + \xi + P) + \lambda(\lambda - 1 + \xi - Q) - (\lambda - 1 + \xi - Q) > 0$$
$$P + \lambda(\lambda - 1 + \xi - Q) + Q > 0$$
$$P - (\lambda - 1)Q + \lambda(\lambda - 1 + \xi) > 0$$
$$(\lambda - 1)Q < P + \lambda(\lambda - 1 + \xi)$$

From the reduced utility matrix in the case in which the common leader plays FF:

	CF	FF
FC	$0, -P, -P$	$0, -P, 0$
FF	$0, 0, -P$	$0, 0, 0$

we see that FF gives zero, less than CC under the above condition.

Consider lastly the utility from playing FC. The utility matrix is

	CF	FF
FC	$-Q/2, -P, -P$	$\frac{\lambda+\xi}{2}, \lambda, \xi$
FF	$-Q/2, 0, -P$	$\frac{\lambda+\xi}{2}, \lambda, \xi$

so her utility is

$$p^2(-Q/2) - p(1-p)(Q - (\lambda + \xi))/2 + (1-p)^2(\lambda + \xi)/2$$

Thus the common leader prefers CC to FC if the following difference is positive:

$$p^2 P + p(1-p)\left((\lambda + \xi) - P\right) + (1-p)^2(2 - (\lambda + \xi))$$

Now the difference would be, since from CC he gets $(1-p)\left(1 + p(\lambda - 1 + \xi - Q)\right)$,

$$\begin{gathered}
(1-p)\left(1 + p(\lambda - 1 + \xi - Q)\right) + p^2(Q/2) + p(1-p)(Q - (\lambda + \xi))/2 - (1-p)^2(\lambda + \xi)/2 \\
(1-p) + p(1-p)(\lambda - 1 + \xi - Q) + p^2(Q/2) + p(1-p)(Q - (\lambda + \xi))/2 - (1-p)^2(\lambda + \xi)/2 \\
(1-p) + p(1-p)\left[(\lambda - 1 + \xi - Q) + (Q - (\lambda + \xi))/2\right] + p^2(Q/2) - (1-p)^2(\lambda + \xi)/2 \\
(1-p) + p(1-p)\left[-1 + \lambda + \xi - Q + Q/2 - (\lambda + \xi)/2\right] + p^2(Q/2) - (1-p)^2(\lambda + \xi)/2 \\
p(1-p)(\lambda + \xi - Q)/2 + p^2(Q/2) + (1-p)^2(2 - (\lambda + \xi))/2 > 0 \\
p(1-p)(\lambda + \xi - Q) + p^2 Q + (1-p)^2(2 - (\lambda + \xi)) > 0
\end{gathered}$$

and for $Q \leq \lambda + \xi$ this is certainly positive for any (ξ, λ) pair in the chicken region. □

We note the following

Corollary. *For $P = 0$ the outcome distribution of the above equilibrium is the same as in the mixed equilibrium of the underlying game.*

Proof. For $P = 0$ we have $\tilde{p} = p(F)$ where $p(F)$ is the probability of F in the mixed equilibrium of the underlying game. Then the claim follows because in the leaders equilibrium: the probability of FF is $\tilde{p}^2$; outcomes FC and FF have probability $\tilde{p}(1-\tilde{p})$; and CC has probability $(1-\tilde{p})^2$. Given $\tilde{p} = p(F)$ this is as in the mixed equilibrium of the underlying game. □

BALIGA, S., D. O. LUCCA, AND T. SJÖSTRÖM (2011): "Domestic political survival and international conflict: is democracy good for peace?" *The Review of Economic Studies*, 78, 458–486.

BALIGA, S. AND T. SJÖSTRÖM (2020): "The strategy and technology of conflict," *Journal of Political Economy*, 128, 3186–3219.

BARRO, R. J. (1973): "The control of politicians: an economic model," *Public choice*, 19–42.

DEWAN, T. AND F. SQUINTANI (2018): "Leadership with Trustworthy Associates," *American Political Science Review*, 112, 844–859.

DOYLE, M. AND N. SAMBANIS (2006): *Making war and building peace: United Nations peace operations*, Princeton University Press.

DUCLOS, J.-Y., J. ESTEBAN, AND D. RAY (2004): "Polarization: concepts, measurement, estimation," *Econometrica*, 72, 1737–1772.

DUTTA, R., D. K. LEVINE, AND S. MODICA (2018): "Collusion constrained equilibrium," *Theoretical Economics*, 13, 307–340.

ELIAZ, K. AND R. SPIEGLER (2020): "A Model of Competing Narratives," *American Economic Review*, 110, 3786–3816.

ESTEBAN, J., L. MAYORAL, AND D. RAY (2012): "Ethnicity and conflict: An empirical study," *American Economic Review*, 102, 1310–42.

ESTEBAN, J.-M. AND D. RAY (1994): "On the measurement of polarization," *Econometrica: Journal of the Econometric Society*, 819–851.

——— (2008): "On the salience of ethnic conflict," *American Economic Review*, 98, 2185–2202.

——— (2011): "Linking conflict to inequality and polarization," *American Economic Review*, 101, 1345–74.

HERMALIN, B. E. (1998): "Toward an Economic Theory of Leadership: Leading by Example," *The American Economic Review*, 88, 1188–1206.

MATĚJKA, F. AND G. TABELLINI (2021): "Electoral competition with rationally inattentive voters," *Journal of the European Economic Association*, 19, 1899–1935.

MEIROWITZ, A., M. MORELLI, R. K. W., AND F. SQUINTANI (2019): "Dispute Resolution Institutions and Strategic Militarization," *Journal of Political Economy*, 127, 378–418.

PRAT, A. AND A. RUSTICHINI (2003): "Games played through agents," *Econometrica*, 71, 989–1026.

RABUSHKA, A. AND K. A. SHEPSLE (1971): "Political entrepreneurship and patterns of democratic instability in plural societies," *Race*, 12, 461–476.