When Conflict is a Political Strategy: a Model of Diversionary Incentives

Melika Liporace*

Bocconi University

September 2021

Latest Version Here

Abstract:

This model revisits the diversionary argument of war by proposing a new mechanism: a population that rebels during a conflict weakens the country's military position; this threat discourages the population to attempt a coup. Being at war thus allows a leader to impose demanding policies without being overthrown. In this context, I show how rally-around-the-flag reactions to conflict can be both rational and efficient. I further prove that purely diversionary incentives exist: international tensions can be initiated with the *only* goal of raising popular support about the conflict. Finally, long-run effects are addressed by allowing rebellion means to be flexible. I find that the population can voluntarily renounce to the freedom to rebel; alternatively, conflicts occur in equilibrium. The strength of the enemy's threat increases the prevalence of barriers to rebellion, while open conflicts are non-monotonically linked to it.

JEL codes: D74, D72, F51

Keywords: Diversionary War, Conflict, Rally around the Flag, Extractive Policy

*I am grateful to Fernando Vega-Redondo for his continued support and to Massimo Morelli and Dominic Rohner for their insights. My thanks further go to Mathias Thoenig, Andrea Pasqualini, Jad Beyhum, and Hormoz Kéchavarz for their precious comments. I gratefully acknowledge financial support from *Fondazione Invernizzi*. Any mistake is my own.

1 Introduction

Do diversionary conflicts exist? The idea is surely attractive: diverting the public's attention abroad in order to hide domestic issues, how easy! But wars are costly. And is the public really that distractible? To believe that diversionary wars exists, one has to believe that all agents are naive : the public, for being distracted so easily, and the instigators of such wars, for not understanding that distraction could be less expensive.

This paper reinterprets diversionary conflicts under a rational light. When no one is naive, can diversionary conflicts exist? I argue that they can. Rather than a distraction, they create a pressure for domestic obedience that no other type of spending could generate. The public does not forget about domestic issues; it cannot act upon it in times of war. This allows leaders to implement demanding domestic policies without risking insurgency. Actually, it can even be rational for a population to support such a conflict.

I begin by exposing three historical examples whose patterns are particularly relevant to the mechanism underlined in this paper. Specifically, I consider Wawro [2005]'s rendition of the 1870 Franco-Prussian war to show how conflict has been designed as a necessary tool to bring independent states to relinquish their sovereignty and rally Germany. Furthermore, I also exploit Zubok [2009]'s vision of the Cold War, to underline the relationship between the official sentiment towards the West and internal popular opposition. Finally, I use Gagnon Jr [2006]'s reinterpretation of the recent Yugoslavian war that emphasizes how conflict has been used as political strategy for demobilizing domestic opposition. These anecdotal evidence seems to point towards patterns that would, in parts, be inconsistent with the classical diversionary argument: patriotic rhetoric and victory on the front are not all that is at stake. Also important seems to be the fact of being at war; and the actual threat posed by the enemy.

In order to formalize these arguments, I introduce a game between two players: a leader, who decides on domestic and foreign policies, and her population, who choses whether to support both domestic and foreign policies. The threat posed by the enemy is captured by a penalty on the population's payoff in case they decide to rebel when a conflict is raging. The leader uses taxation rates to extract wealth from the population. A conflict would then have two consequences: decrease the tax base, as the conflict burns resources; and increase the tax rate, as the threat posed by the conflict would allow the leader to implement harsher policies without risking insurgency. If the latter effect is stronger than the former, the leader benefits from using war as an extractive policy.

The setup distinguishes between low- and high- level conflict. To enter an open war, the leader first has to initiate tensions. However, between these two decisions, the population can express support if tensions were initiated. Because such support would advertise the population's hostility towards the enemy nation, this action is assumed to increase the penalty on the population's payoff in case they rebel during a conflict. Because rebellion decision occurs last, popular support of a belligerent policy can thus be seen as a commitment mean: by advertising their enmity to the other nation, they implicitly tie their hands, as a rebellion would now be more expensive;

this allows the leader to implement demanding policies without having to resort to open war. In this sense, rallies around the flag are both rational and efficient.

However, the very existence of such an opportunity for the population to show support might be detrimental. Indeed, some low-level conflicts are initiated with the only goal of raising popular support. That is, there are contexts in which, if the population were not able to advertise endorsement of the foreign policy, the leader would stay at peace; but the very fact they can show such support is sufficient for the leader to initiate conflict? Using this very restrictive definition of diversionary incentive, I prove that the set of parameters for which such incentive is the driver of conflict is non zero-measure.

Finally, I wonder about long-run effects. Can a population have the incentive to set up barriers to rebellion in order to preserve external peace? When the leader can effectively use the threat of war with an hostile nation as an extraction tool, the population would either preserve peace, at the price of their freedom of rebellion, or preserve their freedom of rebellion, at the price of war. I find that they choose the latter in a non-monotonically function of the strength of the threat. A more threatening enemy makes preserving peace more important; but it also makes it more expensive in terms of commitment to the leader. These two counteracting forces create the hump shaped relationship between the prevalence of conflict and the strength of external threat. Therefore, for important threat, the population would decide to preserve peace by voluntarily renouncing to their freedom of rebellion. This result can account for the evolution of institutions in the long run, in particular in dependence with the institutions of neighboring countries: countries with weaker institutions pose a greater threat, forcing the population to weaken the country's institutions in order to avoid a diversionary conflict, thus spreading bad institutions internationally.

1.1 Related literature

This paper contributes to several strand of the literature. On a doctrinal level, it is built on parts of the classical theory on diversionary wars, as seen in Bodin and Tooley [1955] or Mayer [1969], by underlying the importance of fear in the success of belligerent strategies; and to Levy [1989], I answer that scapegoating does not create nationalism, but does make it salient. Much of the literature on diversionary conflict pertains to foreign policy analysis [see for instance Hagan, 2017, for a survey]. In this strand, the existence of diversionary use of force is either tested or assumed. The empirical evidence on its existence is mixed [Levy, 1989; Chiozza et al., 2004]. Some [Powell, 2014; Jung, 2014; Murray, 2017] assume diversionary conflict in order to empirically study the drivers, implementations or performances of such policies. Others [Oakes, 2012; Davies, 2016] simply assumes them as part of a leader's set of possible strategies to secure her political position. Mostly, this literature recognizes two motivations for diversionary use of force: obtain rally effects; or carry information.

The later justification emerged from economic theory works [Richards et al., 1993; Downs and Rocke, 1994; Hess and Orphanides, 1995]. These papers rely on principal-agent models, in which the outcome of the war is a signal that can determine reelection of the agent. Chiou et al. [2014] twists the argument by arguing that a leader's strategy affects the information environment and thus changes the outcome of a coordination game. Further theoretical contributions [Tarar, 2006; Gent, 2009] include this incentive in international relation models. The former justification, that of rally effects, has been largely highlighted in many related literatures. To the best of my knowledge, they have never been micro founded. It has been justified, as in Theiler [2018], through the social identity theory from sociology; or assumed in theoretical works such as Arena and Bak [2013]. My contribution on this front is thus double: I formalize a new mechanism to justify diversionary conflicts, and I provide a game theoretical insight to diversionary wars justified by rally effects.

In a larger perspective, about conflict theory, I propose a mechanism that contributes to solve the paradox of war, as underlined by Fearon [1995], and which would fall in the agency problem category of Jackson and Morelli [2011]. In particular, the diversionary benefit of war would relate to a political bias as coined by Jackson and Morelli [2007]. Accordingly, Tarar [2006] proves how assuming diversionary motives can lead to the disappearance of the bargaining range. In the political economy literature, I relate to some works on extraction. My mechanism closely relates to Padró i Miquel [2007], that shows how ethnic discrimination within a country permits extraction. Finally, while Acemoglu and Robinson [2005] and Ticchi and Vindigni [2008] see political regimes as a commitment mean for the elite, here a lack of rebellion opportunities is seen, in the long run, as a commitment device for the population.

The remainder of this work is organized as follows. Some anecdotal historical evidence is briefly exposed in Section 2. Section 3 introduces the setup, while Section 4 presents and discusses the equilibrium. Section 5 proves the existence of a diversionnary incentive. Long run effects are explored in Section 6. Section 7 concludes.

2 Motivating examples

In this section, I briefly review three historical examples of the use of war as a policy tool. This anecdotal evidence serves as motivation for model considered. Each example is exposed through the lens of the setup, and mainly relies on one referential work. These expositions are not meant to be exhaustive; they do not – and could not – expose the complete picture of each of the political and historical context. However, these examples should give interesting insights on some recurrent patterns, that are explicited at the end of the section.

2.1 Franco-Prussian War

In the middle of the 19^{th}, what is currently a federated Germany was still heavily divided. Otto von Bismarck, who had just won a war, was now actively working towards a German unification. However, the South States that Bismarck wanted to encompass in the Empire were reluctant to relinquish their independence. As reported by [Wawro, 2005, p.30], these states actually had strong animosities towards Prussia, to which they would prefer France. Bismarck, however, exploited France's territorial ambitions on Belgium, Luxembourg and Rhineland in order to change their opinion. Quickly, the States promised men; but no annexation was effective yet. Three years of tensions and diplomatic crisis followed. Instead of simply declaring war on France, it was important for Bismarck "*to make the French declare war on Prussia, so as to trigger the south German alliances*" [Wawro, 2005, p.37]. First, he instigated a dispute over the Spanish crown succession; but when this did not resolve in war, he wrote a dispatch to the press, falsifying the content of the discussion between a French ambassador and the Prussian king to make it look like the ambassador had been insulted. French opinion ignited and on the 19 June 1870, France had to declare war. Six month later, after a victory on the field, Bismarck also won the four intractable South States. The Constitution of the German Empire was signed; Germany was unified.

First, notice that Bismarck's behavior is difficult to reconcile to the only other economic argument behind diversionary war. Indeed, Bismarck would not have had to signal his competence, as he had just won the Austro-Prussian wars. Another argument often cited is that of war as a mean to revive a patriotic sentiment. But again, this explanation falls short: the South States had no strong nationalist feeling. Furthermore, a mere war was not enough here. Bismarck indeed multiplied tactic in order to not be the instigator of the war. I thus take two lessons from this example. First, it is widely accepted that Bismarck used these tensions with France as a political tool in order to extract concession from the South States. Second, for this to be an efficient tool, France had to look threatening. Here, Bismarck could benefit from this costly war; but, this worked only because the concessions he was extracting from the South States were less costly to them than the threat posed by a French victory.

2.2 Cold War

According to Zubok [2009], in the first month of peace after WWII, the Soviet people were yearning for peace. The war had shaped a strong sense of identity in them, but also deeply scarred their economy. Despite these preferences, both of the people and the elite, Stalin multiplied the action in order to deteriorate the relationship with the West. [Zubok, 2009, p. 29] After 1946 Churchill's iron curtain speech, Stalin replied by accusing the British politician to seek world domination. This changed the public sentiment, as "*the common public wish from now on would not be cooperation with the Western powers but the prevention of war with them. This fear was exactly what Stalin needed to promote his mobilization campaign.*" [Zubok, 2009, p. 53]. Furthermore, "*the winds of a new war also helped Stalin to stamp out any potential discontent and dissent among the elites. The majority of state officials and military officers in the Soviet Union were convinced that the West was on the offensive and had to be contained.*" [Zubok, 2009, p. 60]. After the death of Stalin in 1953, his successor, Khrushchev, launched a de-Stalinization of the country, which involved more liberal policies both internally and externally: from there on, the West was not a threat anymore [Zubok, 2009, p. 104]. This period interestingly corresponds to domestic uprisings and internal turmoil. While the Hungarian Revolution of 1956 is arguably the archetypical mass uprising, Kozlov and MacKinnon [2002] actually documents numerous other popular insurrections, from 1953 to the late 80s. The Cold War ended in 1991 with the collapse of the Soviet Union. This downfall is generally imputed to Gorbachev, who famously advocated liberal policies, and the integration of capitalist principles into the economical system. However, the inadequacy between the ideological systems lead to deceiving results; and popular discontentment rose. Eventually, various revolutions in the Soviet block – coupled with an attempted coup – caused the disintegration of the Soviet Union.

Of course, given the complexity and the length of this conflict, the above exposition barely introduces the issue. However, these few elements all underline a very interesting relationship between internal compliance and external threat. First notice, the ideological confrontation with the West helped Stalin implement policies that were impoverishing his population and discontenting the *nomenklatura*. Of course, other forceful implementations and preventive repressions also accompanied the Stalinian regime, as for example attested by the Trial of the Generals in 1951 or the Slánský Trial in 1952. Yet, it is still interesting to note how the threat of Western victory complemented the other means of repression. One should also note that, despite the absence of open violence in the core of the two blocks, the Soviet Union still had to bear huge due to the conflict, in particular for arming. The second striking element is the timing of popular uprising after Stalin's death, that seem to follow a shift in the state's assessment of the external threat. In February 1956, Khrushchev advertises his *peaceful intentions*; and in October 1956, the major Hungarian Revolution breaks out. Then, Gorbachev comes and establishes a new strategy, with the introduction of capitalist elements to the Soviet system; the Union ends up disintegrating.

2.3 The Yougoslavian War

In the late 20$^{\text{th}}$, the Western world was looking in horror at the heart of Europe: a war was raging. Interpreted by most commentators as an ethnic conflict belonging to a premodern society, Gagnon Jr [2006] shows how the ethnic rhetoric was used and abused by the conservative ruling party, and its president Milošević, in order to instigate a war. It all began in the early 1990, when the Serbian ruling party's position was threatened from all parts: from Slovenia and Croatia, whose population seemed to impose political change; from rival political parties, whose policies were widely endorsed by the public; and from their population, who was mobilizing to ask for competitive elections [Gagnon Jr, 2006, pp. 90-91]. But this conservative party had a plan: induce external violence to demobilize internal opposition. In May 1990, they began instigating violence with Croatia, all the while presenting Croatia as the perpetrator of unjust violence against innocent Serbs. [Gagnon Jr, 2006, pp. 94-95]. They continued provoking as much counteroffensive as possible. "*In Serbia, these conflicts were portrayed as evidence of the Croatian regime's intentions to rid itself of its Serb population*" [Gagnon Jr, 2006, pp. 100]. For years, this worked. Back and forth, popular opposition was rising, and the conflict was worsening [Gagnon Jr, 2006, ch. 4]. Finally, in 1996, the conflict came to an end, and Milošević had to agree to peace. "*Once the Bosnian conflict ended, a popular mass mobilization move-*

ment once again surfaced in Serbia. (...) In response, the grievances that had existed since at least 1990 but which had been demobilized by the wars and the images of threat now burst into the open." [Gagnon Jr, 2006, p. 121]. In February 1998, the conflict in Kosovo began, once again ignited by Serbian forces. "*Meanwhile, faced with growing dissatisfaction and anti-regime mobilization at home, Milošević pointed to the growing unrest in Kosovo as evidence of a continued threat.*"[Gagnon Jr, 2006, p. 125]. However, this time, the international community entered the stage with an ultimatum for Milošević to sign a peace agreement. Upon his refusal, the United States launched several bombing on Serbia, which united all of the Serbian opposition on Milošević's side [Gagnon Jr, 2006, p. 124]. The opposition was again successfully demobilized. Finally in 2000, once the hostilities came to an end, the opposition united, and, endorsed by large popular rallies, successfully demanded elections. Milošević lost.

Gagnon Jr [2006] multiplies evidence to show that the *ethnic* Yougoslavian War was actually a political strategy to weaken internal opposition. Notice how Milošević did not merely declare war, but orchestrated an escalation of violence in order for Croatia and the rest of Yugoslavia to look threatening. In particular, he used state television and newspaper propaganda to show how Serbs' life would be endangered if they were in contact with Croats. This strategy goes beyond nationalist rhethoric. Gagnon Jr [2006, (ch. 2)] indeed provides poll data to show how mild was the ethnic divide before the war. It seems that Milošević's strategy could not have only rested on a tribal instinct of violence against *the others*. The key to the success of his strategy was for the enemy to look threatening.

2.4 From History to Theory

These three examples put into perspective two interesting patterns. First, *being at war* seems to give a leader particular privileges that are revoked once peace is achieved again. Therefore, it seems that that *state of war* is the peculiarity that leaders are after during internal turmoil. Second, more than an enemy, it seems that leaders need a *threatening* enemy to successfully use war for their end. It follows that leaders do not only rely on patriotic sentiments, as often claimed by political commentators.

The only other micro-founded contribution to the diversionary nature of wars does not account for such behaviors; in this setup, in contrast, they are the focal point. In particular, I argue that the state of war is particular because it adds a cost on rebelling. This cost depends on how threatening is the other country. Internal uprising should indeed weaken a country's position on the front; the more threatening the enemy, the more dangerous it is to weaken one's position in the war.

3 Setup

Under study is the incentive for a leader to instigate a conflict with another nation, rather than the bilateral occurrence of war. Therefore, I limit the analysis to one country, which is populated by two players: a leader (she) and citizens whose interests are aligned (they). The leader decides on both internal policies and external policies, while the population's role is to express its agreement with either type of policies. In particular, the population can decide to support belligerent policies; but also to oppose internal policies. In the latter case, the population rebels and ousts the leader in a costly manner.

The policy decided within the country is captured by a continuous scalar τ, which is interpreted as a tax rate.[1] The foreign policy is however discrete, with three level of conflicts: peace, tensions and open war. Open war can occur only if tensions were previously initiated. [2] The decision to initiate tensions is denoted $\theta \in \{0,1\}$, and that of declaring war $\omega \in \{0,1\}$. The population can express support on foreign policy, but only concerning tensions; this choice is captured by $\eta \in \{0,1\}$. Furthermore, the population decides whether to rebel: $\psi \in \{0,1\}$.

The total production from the country is normalized to 1. All costs are expressed as ratio of the country's production. The cost of conflict is κ_t for instigating tensions and κ_w for fighting war. Both of these include tangible and intangible costs. For instance, κ_t can be interpreted as a diplomatic loss of power but also as trade impairments or arming costs. I abstract from the issue of winning or losing: κ_w represents the net cost of war, net of potential advantages derived from winning the war and including further disadvantages from losing it.

Rebelling also comes at a cost. In case of rebellion, the population pays a price ρ in order to oust the leader. However, the population raises this cost to ρ' when they decide to support the leader's decision to initiate tension. It is indeed assumed that a public support has an intrinsic value: coordination and organization would be harder for inconsistent public opinion. This additionally captures how a population that advertised animosities towards a foreign nation would be worse off forfeiting against them. [3]. Indeed, if a rebellion is attempted in time of conflict, the country has to forfeit on the international front. The additional cost of forfeiting, irrespective of the support decision is defined as ϕ_t in case of tensions and as ϕ_w in case of war. In case of rebellion, the leader's payoff is normalized to 0. The population grabs what is left $(1-\phi(h))(1-\rho(h))$, where h represents the history that led to the rebellion.

Finally, taxes are modeled as a transfer of utility. Furthermore, differences in risk attitudes are ignored. Any cost is perceived similarly by both player. For instance, if a war is declared and no revolution occurs, the available interior product is reduced from 1 to $1-\kappa_w$ for both players, who then share this amount according to the taxation rate proposed by the leader.

The timing is as follows:

1. The leader decides whether to initiate tension, $\theta \in \{0,1\}$.

[1] The interests of the leaders might not be pecuniary. Likewise, the concerns of the population might go beyond economical disagreements. This however capture misaligned interests between the leader and her population in a very conservative way.

[2] There is no reason to assume that the outcome from a war instigated after tensions would be different from that of a war that was not preceded by tensions. Therefore, allowing for open war without tension would not change the conditions for war, but could create multiple equilibria. For clarity concerns, I impose this restriction.

[3] Appendix B considers alternative specifications in which the population's decision to support tension, rather than impacting ρ: (i) impacts the entire payoff by a factor $(1-x)$; (ii) directly impacts ϕ_t and ϕ_w. I show how all results carry out.

2. If $\theta = 1$, the population decides on support, $\eta \in \{0, 1\}$; no decision otherwise.

3. If $\theta = 1$, the leader decides whether to intensify the conflict into an open war, $\omega \in \{0, 1\}$; no decision otherwise.
   The leader sets the taxation, $\tau \in [0, 1]$.

4. The population decides whether to rebel, $\psi \in \{0, 1\}$.

The extensive form of this game is depicted in Figure 1.

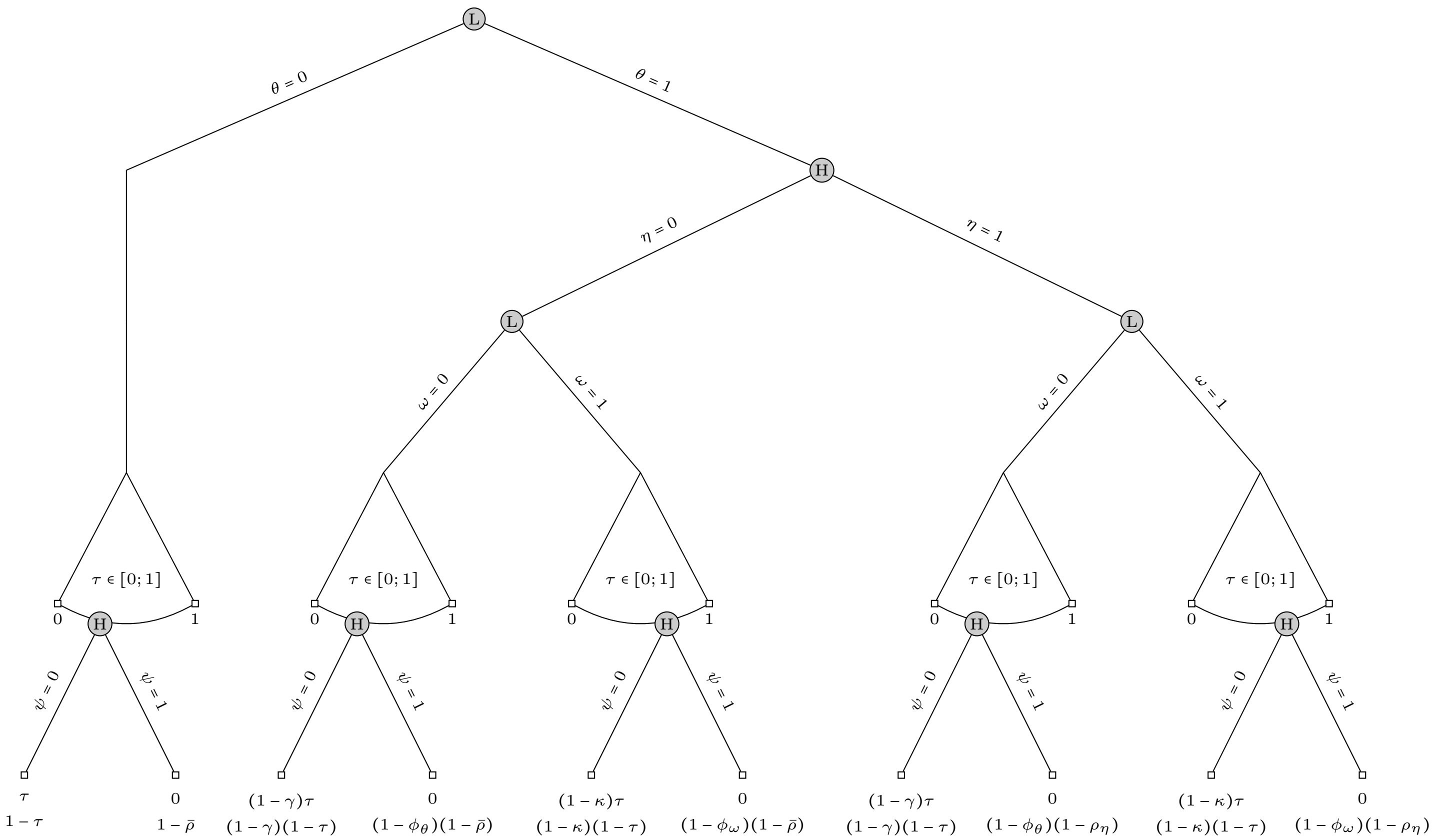

Figure 1: Extensive Form Game

Finally, consider the following tie rule: when a player is indifferent, they choose the status quo – i.e. the more peaceful decision for the leader; no revolution and no support for the population.

4 Equilibrium

I characterize the Subgame Perfect Equlibrium (SPE) of this game by using backward induction. Note that the only tie rule necessary for the existence of an equilibrium is the one pertaining to revolution; however, I do not specify the equilibria under other tie rules, as no results is driven from it.

In the last stage, the population rebels if the tax rate is high enough for a rebellion to lead to a higher payoff than the status quo. Hence, in the penultimate stage, the leader sets the tax rate to make the population indifferent between rebelling or not, i.e.

$$\tau(h) = 1 - \frac{(1-\phi(h))(1-\rho(h))}{1-\kappa(h)}$$

at every partial histories $h \in H_3$ [4] [5] .

When deciding whether to escalate, the leader internalizes all subsequent moves and compares $1 - \kappa_w - (1-\phi_w)(1-\rho(\eta))$ to $1 - \kappa_t - (1-\phi_t)(1-\rho(\eta))$ [6]. Hence, he choses:

- to always escalate to war if $\kappa_w - \kappa_t < (1-\rho')(\phi_w - \phi_t)$
- to never escalate to war if $\kappa_w - \kappa_t \geq (1-\rho)(\phi_w - \phi_t)$
- to escalate only if he is not supported if $(1-\rho')(\phi_w - \phi_t) \leq \kappa_w - \kappa_t < (1-\rho)(\phi_w - \phi_t)$

The population will decide to support the tension only if this decision avoids open war. Indeed, if the parameters are such that the leader takes the same action, irrespectively of the population's decision, then support always worsen the payoff for the population. Hence, for the population to support the tensions, it must be that $(1-\rho)(\phi_w - \phi_t) \leq \kappa_w - \kappa_t < (1-\rho')(\phi_w - \phi_t)$. For such parameters, the population must actually be willing to pay the price of support. In particular, the price of support must be lower than the price of escalating to war. Hence, $\eta = 1$ additionally requires $(1-\phi_t)(1-\rho') > (1-\phi_w)(1-\rho)$.

In the first stage, the leader decision depends on the parameters:

- If $\kappa_w - \kappa_t < (1-\rho)(\phi_w - \phi_t)$, the leader compares peace to war (with no support); he chooses peace if $\rho \geq 1 - \kappa_w - (1-\phi_w)(1-\rho)$.
- If $\kappa_w - \kappa_t \geq (1-\rho')(\phi_w - \phi_t)$, the leader compares peace to tensions (with no support); he chooses peace if $\rho \geq 1 - \kappa_t - (1-\phi_t)(1-\rho)$.
- If $(1-\rho)(\phi_w - \phi_t) \leq \kappa_w - \kappa_t < (1-\rho')(\phi_w - \phi_t)$ and $(1-\phi_t)(1-\rho') > (1-\phi_w)(1-\rho)$, the leader compares peace to supported tensions; he chooses peace if $\rho \geq 1 - \kappa_t - (1-\phi_t)(1-\rho')$.
- If $(1-\rho)(\phi_w - \phi_t) \leq \kappa_w - \kappa_t < (1-\rho')(\phi_w - \phi_t)$ but $(1-\phi_t)(1-\rho') \leq (1-\phi_w)(1-\rho)$, the leader compares peace to war (with no support); hence the condition for peace is again $\rho \geq 1 - \kappa_w - (1-\phi_w)(1-\rho)$.

In order to lighten notation, I define the following values:

[4] Let H_3 be the set of histories until the tax decision node: $H_3 = \{(\theta=0), (\theta=1,\eta=0,\omega=0), (\theta=1,\eta=0,\omega=1), (\theta=1,\eta=1,\omega=0), (\theta=1,\eta=1,\omega=1)\}$

[5] Let $\rho(h) = \begin{cases} \rho' & \text{if } h \in \{(\theta=1,\eta=1,\omega=0),(\theta=1,\eta=1,\omega=1)\} \\ \rho & \text{otherwise} \end{cases}$. Likewise, $\phi(h) = \begin{cases} \phi_w & \text{if } h \in \{(\theta=1,\eta=0,\omega=1),(\theta=1,\eta=1,\omega=1)\} \\ \phi_t & \text{if } h \in \{(\theta=1,\eta=0,\omega=0),(\theta=1,\eta=1,\omega=0)\} \\ 0 & \text{if } h = (\theta=0) \end{cases}$ and $\kappa(h) = \begin{cases} \kappa_w & \text{if } h \in \{(\theta=1,\eta=0,\omega=1),(\theta=1,\eta=1,\omega=1)\} \\ \kappa_t & \text{if } h \in \{(\theta=1,\eta=0,\omega=0),(\theta=1,\eta=1,\omega=0)\} \\ 0 & \text{if } h = (\theta=0) \end{cases}$

[6] By a slight abuse notation, I refer to history $(\theta=1,\eta)$ as (η) for the remainder of this paper.

- $\underline{\kappa}_t = \phi_t(1-\rho')$
- $\bar{\kappa}_t = \phi_t(1-\rho)$
- $r_\rho = \frac{1-\rho'}{1-\rho}$
- $\underline{\kappa}_w = \phi_w(1-\rho')$
- $\bar{\kappa}_w = \phi_w(1-\rho)$
- $r_\phi = \frac{1-\phi_w}{1-\phi_t}$

$\bar{\kappa}_t$ (resp. $\bar{\kappa}_w$) can be interpreted as the cost from which tensions (resp. war) are beneficial *per se* for the leader; and $\underline{\kappa}_t$ (resp. $\underline{\kappa}_w$) as the cost from which tensions (resp. war) are beneficial for the leader only under support. Finally, r_ρ is a ratio representing how affordable support is for the population while r_ϕ indicates how cheap war is to them relatively to tensions.

Given the previous discussion, the SPE is defined in Lemma 1

Lemma 1. *The SPE is as follows:*

- *The leader's equilibrium strategy is:*
  - $\theta = 1$ *if* $\kappa_w < \bar{\kappa}_w$*; or if* $\kappa_t < \bar{\kappa}_t$*; or if* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$*, and* $r_\phi < r_\rho$*, and* $\kappa_t < \underline{\kappa}_t + \rho' - \rho$*;*
    $\theta = 0$ *otherwise.*
  - $\omega(\eta = 0) = 1$ *if* $\kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$*;* $\omega(\eta = 0) = 0$ *otherwise.*
    $\omega(\eta = 1) = 1$ *if* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$*;* $\omega(\eta = 1) = 0$ *otherwise.*
  - $\forall h \in H_3, \tau(h) = 1 - \frac{(1-\phi(h))(1-\rho(h))}{1-\kappa(h)}$
- *The population's equilibrium strategy is:*
  - $\eta = 1$ *if* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$ *and* $r_\phi < r_\rho$*;* $\eta = 0$ *otherwise.*
  - $\forall h \in H_3 \times [0,1]$ *and corresponding* $h' \in H_3$*:*
    $\psi(h) = 1$ *if* $\tau > 1 - \frac{(1-\phi(h'))(1-\rho(h'))}{1-\kappa(h')}$*;* $\psi(h) = 0$ *otherwise.*

Notice that rebellion never occurs in equilibrium,. This directly follows from perfect information. The leader will always set the taxation rate just low enough to avoid popular uprising. Likewise, a war following support never happens in equilibrium. Again, this follows directly from the assumptions that are meant to reflect this paper's interpretation of such behaviors: a way to avoid more severe conflict.

Theorem 1. *There are four possible outcomes: peace, (unsupported) war, unsupported tensions and supported tensions.*

- *War occurs iff* $\kappa_w < \bar{\kappa}_w$*; and* $\kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$*; and either* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$*, or* $r_\phi \geq r_\rho$*.*
- *Supported tensions occur iff* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$*; and* $r_\phi < r_\rho$*; and* $\kappa_t < \underline{\kappa}_t + \rho' - \rho$*.*
- *Unsupported tensions occur iff* $\kappa_t < \bar{\kappa}_t$*; and* $\kappa_w - \kappa_t \geq \bar{\kappa}_w - \bar{\kappa}_t$*.*
- *Peace occurs iff* $\kappa_w \geq \bar{\kappa}_w$*; and* $\kappa_t \geq \bar{\kappa}_t$*; and either* $\kappa_t \geq \underline{\kappa}_t + \rho' - \rho$*, or* $r_\phi \geq r_\rho$*, or* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$*, or* $\kappa_w - \kappa_t \geq \bar{\kappa}_w - \bar{\kappa}_t$ *.*

Remark 1. *In equilibrium, the population can indeed rally around the flag.*

Proof. Lemma 1 is used to find the set of conditions determining the action at each node. They can then be groupped by possible outcomes and redundant conditions are simplified. Details can be found in Appendix A. □

Hence, for peace to exist, it is necessary that neither type of conflict is beneficial *per se*, and that escalation is not either; but it is not sufficient. It must also be that the leader does not have an incentive to initiate conflict in order to initiate a rally around the flag – so either, support is simply not worth the cost of tension, or support is too expensive for the population, or it would be inefficient in preventing escalation, or aggravating conflict is not credible for the leader. This is actually what supported tensions require: for support to be necessary and efficient at preventing escalation, but also for it to be preferred by the population, and by the leader. Unsupported tensions occur when low-level conflict is beneficial to the leader *per se*, but escalation is not. Finally, for a war to happen, it is necessary that war is beneficial *per se* but also that escalation is. Again, these two conditions are not sufficient, as it must also be that the population cannot or is not willing to avoid escalation through support.

4.1 Discussion

Let us now focus one some particular sets of parameters, in order to have a better grasp on the importance of each parameter. The parameters' importance will then be discussed in general.

First, what if raising international tensions was not costly, i.e. $\kappa_t = 0$? If additionally, such tensions can be freely forfeited, that is $\phi_t = 0$, then no unsupported tensions can occur. Indeed, this would mean that tensions are without consequences for either player, so that the leader would never have a strict incentive to initiate tension for the sake of it. However, it is interesting to notice that peace is not ensured either. If $\kappa_w < \bar{\kappa}_t$, war is beneficial *per se*, that is, it decreases the tax base less than it allows to increase the tax rate. There are two possible outcomes then: either the population can and is willing to avoid escalation through support, resulting in supported tensions; or they are not, resulting in war. Now, if $\kappa_t = 0$ but $\phi_t > 0$, peace can never occur. Indeed, this means that the tax base does not decrease for the leader, but the population is threatened by such tensions. Therefore, the leader can freely increase her payoff by initiating tensions. Whether this leads to war, supported or unsupported tensions depends on the other parameters.

On the other end, huge costs of tensions are not enough to insure peace, as it is not even sufficient to avoid tensions. In particular, if $\kappa_t > \phi_t$, that is, if tensions are reducing the countries wealth more when affirmed than when forfeited, it is still possible for tensions to occur in equilibrium, but only if they are supported. Indeed, the population can credibly support such a conflict in order to avoid war, and not rebel in equilibrium. It is surprising because, staying in a low-level conflict creates an additional incentive to rebel, that must be balanced off by a smaller taxation rate. However, while the leader has to pay people in order to avoid rebellion under tensions, he is still paid by the support the population gives her in order to avoid war. Therefore, it can be that, in equilibrium, tensions occur despite their huge cost.

Another interesting case regarding costs occurs when escalation to war has proportional effects for both players, that is $\frac{\kappa_t}{\kappa_w} = \frac{\phi_t}{\phi_w}$. In such a context, unsupported tensions never occur, because if tensions were beneficial *per se*, war *a fortiori* would be. The occurrence of peace depends on the level of fighting costs κ, while the occurrence of war depends on their ratio. Indeed, the highest the above ratio, the more similar the effects of war relative to tensions are on the population, hence the more likely war becomes. Support gets indeed relatively too costly for high value of this ratio.

However, it is worth noting that, in general, higher costs of fighting tend to decrease the occurrence of either type of conflicts. Hence, it tends to have beneficial effects for the society – any type of conflict is indeed Pareto inefficient. Yet, while increasing costs *sufficiently* to ensure peace is unambiguously desirable, some increases in fighting costs might be counter-productive. Indeed, if the increase is not sufficient to insure peace, it ends up creating a higher deadweight loss from fighting a conflict that is now destroying more wealth.

Likewise, if low rebellion costs should, at first sight, be desirable, they are not unambiguously so. First, notice that $\rho = 0$ is rarely desirable, as peace would then rarely be possible. Indeed, in time of peace, the cost of revolution is the only pressure the leader can use in order to extract wealth from his population without being ousted. Decreasing revolution costs below the level that insures the leader the same revenue as in time of conflict would thus worsen the situation for both players, as this would create the inefficiency of actually having to create a conflict in order for the leader to extract wealth from the population without risking to be ousted. Therefore, a very high ρ can actually be optimal from a social perspective. I further analyze this tradeoff in Section 6.

The role of the commitment price ρ' is even more ambiguous. Indeed, if it is too small, it could be inefficient in convincing the leader to renounce to war, and thus would lead to a high-level conflict rather than reaching Pareto preferred supported tensions. But a ρ' that is too high could dissuade the population from preventing war through support. If the leader initiates tensions only for diversionary reasons[7], then a commitment price that is high enough to make popular support impossible would actually insure peace. In such a case, a very high ρ' is preferable. However, if war is beneficial *per se*, a commitment price that is too high could make war avoidable but not avoided. Then, lower ρ' would be socially preferred.

Finally, it can unambiguously be concluded that the higher the costs of forfeiting, the more menacing the threat of conflict. This means a bigger extraction power for the leader, who is hence more prone to enter conflicts, which creates inefficiencies. In Section 6, I further analyze how the mere existence of this threat leads the population to both rationally and efficiently renounce to some of their freedom.

[7] That is, only in anticipation of popular support. I formalize the concept of *diversionary incentives* considered in Section 5.

5 Diversionary Incentive

In this section, I prove the existence of a *diversionary incentive* for conflicts. Loosely speaking, one could call a war *diversionary* when it is used to prevent a leader's position to be questioned. In this setup however, the incentive to use conflict is, by assumption, always that of preventing a rebellion while implementing demanding policies. While I do find that conflict can occur in equilibrium, this does not fully address the question.Tthe population always accepts the internal policy in equilibrium, but popular support of an external conflict is not given. Theorem 1 shows that rally around the flag do occur. The next logical question is thus: does the anticipation of such a support suffice for a leader to initiate war?

The definition of a diversionary incentive thus follows:

Definition. A leader initiates conflict because of a *diversionary incentive* when:

- the leader anticipates the public to support it; and
- the leader would not have initiated it, were it not supported by the public.

Using Lemma 1 and Theorem 1, this definition can be translated in terms of the setup's parameters.

Corollary 1. *A diversionary incentive emerges when:*

1. *The leader initiates conflict:* $\kappa_t < \underline{\kappa}_t - \rho' + \rho$.
2. *Support can credibly be anticipated:* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$ *and* $r_\phi < r_\rho$.
3. *Neither conflict would be initiated otherwise:* $\kappa_t \geq \bar{\kappa}_t$ *and* $\kappa_w \geq \bar{\kappa}_w$.

Proof. Conditions in 1. and 2. correspond to the conditions for supported tensions derived in Theorem 1. Conditions in 3. are derived using the equilibrium strategies specified in Lemma 1; they correspond to the conditions for peace for the alternative game in which the population is not active until the last period (i.e. $\eta \in \{0\}$). □

If war was not beneficial *ex ante*, how could it be a credible threat to escalate the conflict? If the threat is not credible, popular support is impossible. Yet, the threat *can* be credible: once tensions have been instigated, war might become more attractive, making escalation credible. Likewise, while tensions would not be beneficial *per se*, the additional benefit the leader would derive from popular support would make them attractive *ex ante*. Finally, one could argue that, if the support is valuable enough to make the leader willing to instigate tensions that would otherwise be too expensive, then the population cannot be willing to pay the price of support. Again, this does not stand when confronted to the respective formal conditions. Actually, all conditions of Corollary 1 can be met at the same time.

Theorem 2. *There exists a non-zero measure parameter space $\mathcal{D}$ for which the leader has a diversionary incentive to initiate conflict.*

Proof. To prove that $\mathcal{D}$ is non-zero measure, let us find a subset $\mathcal{D}' \subseteq \mathcal{D}$ that is non zero measure. $\mathcal{D}'$ is arbitrarily chosen as follows:

$$\mathcal{D}' = \left\{ (\kappa_t, \kappa_w, \phi_t, \phi_w, \rho, \rho') : \forall (x, \epsilon) \in \mathcal{E}, \begin{array}{lll} \rho = x; & \phi_t = x + \epsilon_2; & \kappa_t = x + \epsilon_4; \\ \rho' = x + \epsilon_1; & \phi_w = x + \epsilon_3; & \kappa_w = x + \epsilon_5; \end{array} \right\}$$

$$\text{where } \mathcal{E} = \left\{ (x, \epsilon) \in (0,1)^6 : \begin{array}{c} 1 - x > \epsilon_4 + \epsilon_1 > \epsilon_5 > \epsilon_4 + (1-\epsilon_1)(\epsilon_3 - \epsilon_2) \\ \epsilon_5 > \epsilon_3 > \frac{3}{2}\epsilon_1 + \epsilon_2 \end{array} \text{ and } \min\{t, \tfrac{1}{3}\} > x \right\}$$

$$\text{with } t = \frac{\sqrt{(\epsilon_1+\epsilon_2)^2 + 4(\epsilon_1+\epsilon_3-\epsilon_5-\epsilon_1\epsilon_3)} - (\epsilon_1+\epsilon_2)}{2}$$

The conditions in $\mathcal{E}$ in particular imply $\epsilon_1 > \max\left\{\frac{\epsilon_5-\epsilon_3}{1-\epsilon_3}; \frac{\epsilon_3-\epsilon_2}{1+\epsilon_3-\epsilon_2}; \frac{1}{3}\right\}$. Indeed, for the interval for ϵ_5 to exist, it must be that $\epsilon_1 > (1-\epsilon_1)(\epsilon_3-\epsilon_2)$, meaning $\epsilon_1 > \frac{\epsilon_3-\epsilon_2}{1+\epsilon_3-\epsilon_2}$. On the other hand for t to be well-defined and positive, it must be that $4(\epsilon_1+\epsilon_3-\epsilon_5-\epsilon_1\epsilon_3) > 0$, hence $\epsilon_1 > \frac{\epsilon_5-\epsilon_3}{1-\epsilon_3}$. Finally, $\epsilon_1 > \frac{\epsilon_3-\epsilon_2}{1+\epsilon_3-\epsilon_2}$ and $\epsilon_3 - \epsilon_2 > \frac{3}{2}\epsilon_1$ implies $\epsilon_1 > \frac{1}{3}$. Because none of these implications contradict any other defined in $\mathcal{E}$, $\mathcal{E}$ is non-zero measure.

It is easy to verify that any vector from $\mathcal{D}'$ belongs to the space of feasible parameters. Indeed $1 - x > \max_{i=1,\dots,6}\{\epsilon_i\}$ so all parameters are lesser than 1. Because $\min_{i=1,\dots,6}\{\epsilon_i\} > 0$ and $x > 0$, all parameters are strictly positive.

All that is left to do is to prove that $\mathcal{D}' \subseteq \mathcal{D}$. I show that for any $v \in \mathcal{D}'$, the conditions from Corollary 1 are all fulfilled:

1. Leader prefers supported tensions to peace:
   Indeed: $\kappa_t - \phi_t(1-\rho') - \rho' + \rho = \epsilon_4 - \epsilon_1 - \epsilon_2 + (x+\epsilon_1)(x+\epsilon_2) < 0$
   It follows from $\epsilon_4 < \epsilon_5 - (1-\epsilon_1)(\epsilon_3-\epsilon_2) < \epsilon_1 + \epsilon_2 - (x+\epsilon_1)(1+\epsilon_2)$ where the last inequality holds because $x < t$.

2. Support can credibly be anticipated:
   (a) Escalation without support is credible:
   Indeed: $\kappa_w - \phi_w(1-\rho) - \kappa_t + \phi_t(1-\rho) = \epsilon_5 - \epsilon_4 - (1-x)(\epsilon_3-\epsilon_2) < 0$.
   It follows from the definition of $\mathcal{E}$, in particular: $\epsilon_3 - \epsilon_2 > \frac{3}{2}\epsilon_1$, $\epsilon_5 - \epsilon_4 < \epsilon_1$, and $x < \frac{1}{3}$.
   (b) Escalation with support is avoided:
   Indeed: $\kappa_w - \phi_w(1-\rho') - \kappa_t + \phi_t(1-\rho') = \epsilon_5 - \epsilon_4 - (1-x-\epsilon_1)(\epsilon_3-\epsilon_2) > 0$.
   It follows from $\epsilon_5 - \epsilon_4 - (1-\epsilon_1)(\epsilon_3-\epsilon_2) + x(\epsilon_3-\epsilon_2) > \epsilon_5 - \epsilon_4 - (1-\epsilon_1)(\epsilon_3-\epsilon_2) > 0$
   (c) Preferred to war by the population:
   $(1-\phi_w)(1-\rho) - (1-\phi_t)(1-\rho') = -(1-x)(\epsilon_3-\epsilon_2-\epsilon_1) - \epsilon_1\epsilon_2 < 0$
   If follows from $\epsilon_3 > \frac{3}{2}\epsilon_1 + \epsilon_2$.

3. Neither conflict would be initiated otherwise:

(a) Tensions are not beneficial *per se*:
Indeed: $\kappa_t - \phi_t(1-\rho) = \epsilon_4 - \epsilon_2(1-x) + x^2 > 0$
It follows from $\epsilon_4 - \epsilon_2(1-x) > \epsilon_4 - \epsilon_2 > 0$, since $\epsilon_4 + \epsilon_1 > \frac{3}{2}\epsilon_1 + \epsilon_2$.

(b) War is not beneficial *per se*:
Indeed: $\kappa_w - \phi_w(1-\rho) = \epsilon_5 - \epsilon_3(1-x) + x^2 > 0$
It follows from $\epsilon_5 - \epsilon_3(1-x) > \epsilon_5 - \epsilon_3 > 0$.

□

Therefore, in a case for which the set of parameters belongs to $\mathcal{D}$, the very possibility for the population to commit through support decreases their payoff. The population would be better off without the freedom of rallying around the flag. The outcome would also be more efficient. However, this is not a general result. The population is better off with this choice when popular support prevents an otherwise unavoidable war. I further wonder about efficiency and rationality of flexible commitment means in section 6.

6 Long Run Effects

In this section, I wonder about the optimality and efficiency of flexible rebellion costs. In the long-run, the cost of popular uprising can indeed be considered endogenous: the institutional, military and legal impairments put in place in order to restrain rebellions are changeable, and the population can contribute to shape them.

Rather than discretely changing ρ to ρ' through the support decision, the population can freely chose the level of ρ. This decision is meant to represent the incentive for a population to put into place the institutional framework that would allow them to easily oust a leader. Obviously, this is a long run consideration. It particularly depends on the diplomatic and military condition of a country; but could not *follow* a short term decision to initiate international tensions. Therefore, from now on, I do not differentiate between tensions and open war anymore: a leader can either chose to initiate conflict, at cost κ, or not. If war is initiated, a population who decides to rebel would still suffer a cost of forfeit, uniquely defined as ϕ. These two parameters are now the only exogenous elements.

The new timing is as follows:

1. The population sets the cost of rebellion, $\rho \in [0,1]$.
2. The leader decides her international position, $\omega \in \{0,1\}$;
   she sets the taxation rate accordingly, $\tau \in [0,1]$.
3. The population decides their domestic position, $\psi \in \{0,1\}$.

The extensive form of the game is depicted in Figure 2. The same tie rules as before are kept. Furthermore, if the population is indifferent between $\rho = 0$ and $\rho > 0$, they set $\rho = 0$.

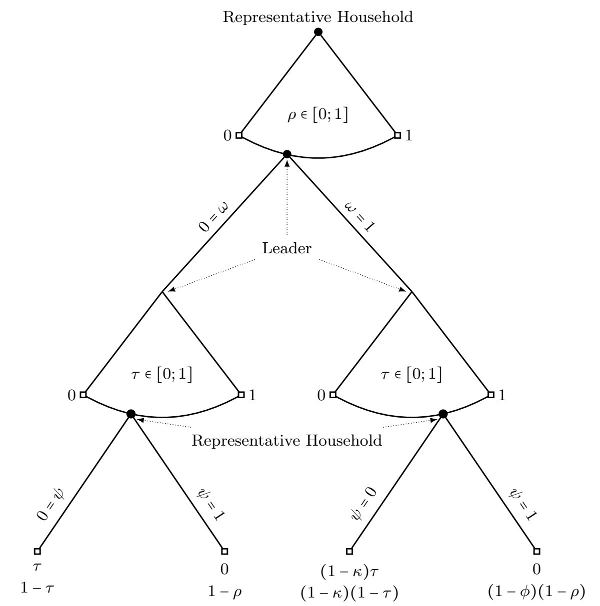

Figure 2: Long-Run Game: Extensive Form

6.1 Equilibrium

I characterize the SPE of this game using backward induction.

Because in the last stage, the population choses to rebel only if their payoff doing so is strictly greater than through the acceptance of the taxation rate, in the penultimate period, the leader sets the taxation rate to make them indifferent between $\phi = 0$ and $\phi = 1$. That is: $\tau : (1-\kappa)(1-\tau) = (1-\phi)(1-\rho)$. It follows that:

$$\tau(h) = 1 - \frac{(1-\phi(h))(1-\rho(h))}{1-\kappa(h)}$$

at every partial histories $h = (\rho, \omega) \in [0,1] \times \{0,1\}$. [8] .

When deciding whether to enter in a conflict, the leader trades off the loss in tax base induced by war with the increased tax rate he could be imposing during conflict. Hence, at history $h = (\rho) \in [0,1]$, he choses peace if $\rho(h) \geq 1 - \frac{\kappa}{\phi}$; and war otherwise.

In the first period, the population can preventing conflict by increasing rebellion cost. They know that any ρ lower than $1 - \frac{\kappa}{\phi}$ would result in a conflict. Therefore, they compare their payoff with $\rho = 1 - \frac{\kappa}{\phi}$ and $\omega = 0$ to $\rho = 0$ and $\omega = \mathbb{1}_{(\kappa<\phi)}$. Obviously, if $\kappa \geq \phi$, a conflict raises more costs than benefits, so that the leader would never start a conflict, and the population has no reason to set anything else than $\rho = 0$. Now, if $\kappa < \phi$, the population compares committing $- \frac{\kappa}{\phi} -$ to being at war $- 1 - \phi$. If $\kappa \geq \phi(1-\phi)$, they decide to increase rebellion costs above 0; otherwise, they do not and let war be used as an extraction tool.

The SPE is characterized in Lemma 2

[8] $\rho(h)$ is trivially defined in h. Let $\phi(h) = \begin{cases} \phi & \text{if } h \in [0,1] \times \{1\} \\ 0 & \text{if } h \in [0,1] \times \{0\} \end{cases}$ and $\kappa(h) = \begin{cases} \kappa & \text{if } h \in [0,1] \times \{1\} \\ 0 & \text{if } h \in [0,1] \times \{0\} \end{cases}$

Lemma 2. *The SPE of the LR game is as follows:*

- *The population's equilibrium strategy is:*
  - $\rho = 1 - \frac{\kappa}{\phi}$ *if* $\phi < \kappa < \phi(1-\phi)$*;* $\rho = 0$ *otherwise.*
  - $\forall h = (\rho, \omega, \tau) \in [0,1] \times \{0,1\} \times [0,1]$ *and corresponding* $h' = (\rho, \omega) \in [0,1] \times \{0,1\}$*:* $\psi(h) = 1$ *if* $\tau > 1 - \frac{(1-\phi(h'))(1-\rho(h'))}{1-\kappa(h')}$ *;* $\psi(h) = 0$ *otherwise.*
- *The leader's equilibrium strategy is:*
  - $\forall \rho \in [0,1]$*,* $\omega(\rho) = 1$ *if* $\rho < 1 - \kappa\phi$*;* $\omega(\rho) = 0$ *otherwise.*
  - $\forall h \in [0,1] \times \{0,1\}$*,* $\tau(h) = 1 - \frac{(1-\phi(h))(1-\rho(h))}{1-\kappa(h)}$

Again, because of perfect information, no rebellion ever occurs.

Theorem 3. *There are three possible equilibrium outcomes: conflict, committed peace and uncommitted peace.*

- *Conflict occurs iff* $\kappa \leq \phi(1-\phi)$*.*
- *Committed peace occurs iff* $\phi < \kappa < \phi(1-\phi)$*.*
- *Uncommitted peace occurs iff* $\kappa \geq \phi$*.*

Remark 2. *(i) In equilibrium, conflicts can occur despite perfectly flexible commitment means; (ii) The strength of the foreign threat has a non-monotonic effect on the prevalence of conflict; (iii) The prevalence of commitment is positively linked with the strength of the foreign threat.*

Proof. This directly follows from Lemma 2. □

Figure 3 depicts the different outcomes as a function of the parameters.

6.2 Discussion

As seen in Theorem 3, for the leader to use conflict as a tool of fiscal extraction, it has to be that conflict is relatively cheap. Interestingly, there are two effects at play. The cheapest conflict, the more the leader can benefit from war, as tax base decreases relatively less than the possible increase in tax rate. But the cheapest the conflict, the less prone is the population to commit in order to avoid it. These two forces create the non-monotonic effect of ϕ on the prevalence of conflict.

Because conflict is Pareto inferior, ϕ has the same non-monotonic effect in social efficiency. However, it has an unambiguous negative effect on the populations' equilibrium payoff. Indeed, an increase in ϕ has two negative effects for the population: on the one hand, in time of conflict, the leader can extract more from them, as the threat is bigger; on the other hand, the population has to pay a higher price in terms of ρ in order to preserve peace. The overall effect is negative

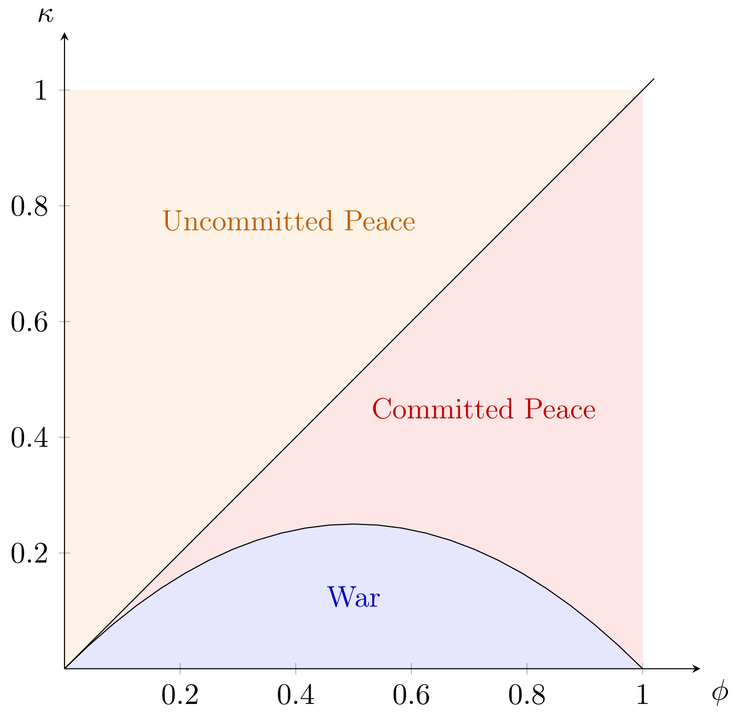

Figure 3: Possible Outcomes in Parameters Space

for the population but mainly positive for the leader. Indeed, her payoffs tend to increase as a function of ϕ; however, because the deadweight loss created by conflict is born by the leader, her payoff is discontinuous at $\phi(1-\phi)=\kappa$; there, the payoff shifts down. Therefore, there are values of ϕ whose decrease could benefit both players. Figures 4 and 5 illustrate this conclusion. They depict the equilibrium payoff of each agents as a function of ϕ, given various κ.

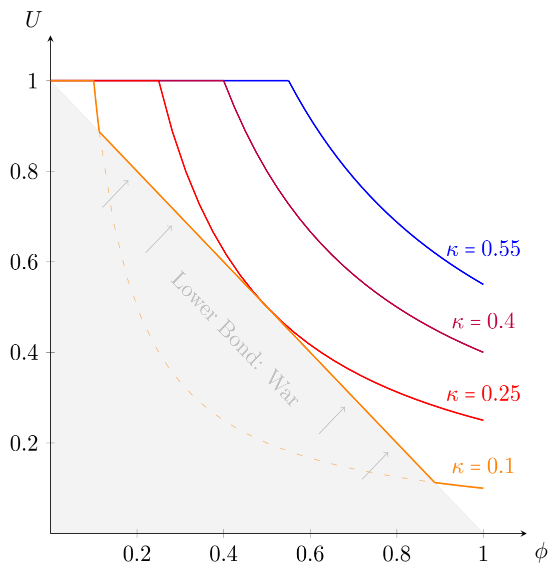

Figure 4: Role of ϕ on Population's Payoff

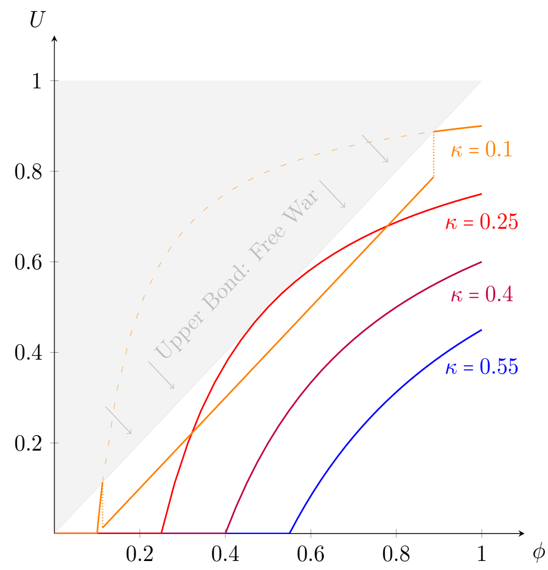

Figure 5: Role of ϕ on Leader's Payoff

Because the deadweight loss created by conflict is born by the leader, the population's payoff is constant in terms of κ as long as $\kappa \leq \phi(1-\phi)$. It is increasing afterwards, as the leader is easier to satisfy through a lower ρ when conflict gets more costly. Again, the leader's payoff is discontinuous in $\kappa = \phi(1-\phi)$. Therefore, although her payoff is decreasing piecewise, there are values of κ whose increase could benefit both players. Furthermore, small κ is also inefficient from a social point of view, as it allows for conflicts to take place. Therefore, for $\kappa < \phi(1-\phi)$, it is unambiguously beneficial to increase κ to at least $\phi(1-\phi)$. Again, Figures 6 and 7 depict the equilibrium payoff of each agent as a function of κ, given various ϕ.

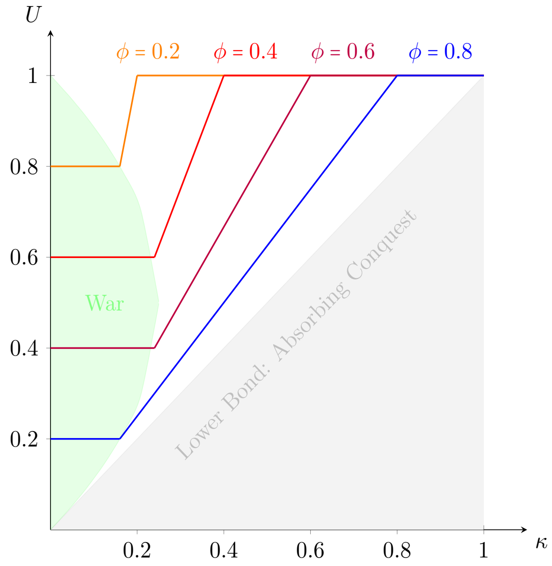

Figure 6: Role of κ on Population's Payoff

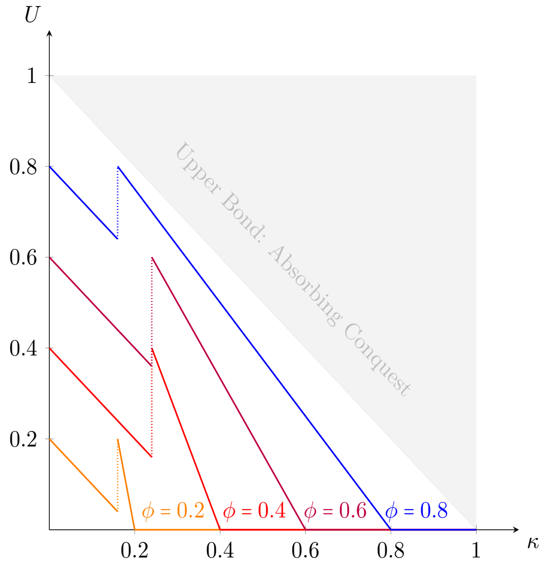

Figure 7: Role of κ on Leader's Payoff

Finally, it is worth emphasizing again how $\rho = 0$ is rarely optimal. Indeed, it is optimal only when the leader cannot use conflict as a fiscal extraction tool, i.e. $\kappa \geq \phi$. Otherwise, the leader can threaten the population to use conflict as a way to implement demanding policies while avoiding rebellion. Then, the population might find optimal to voluntary enchain: prevent the threat posed by the instigation of an external conflict, and set up relatively inefficient institutions. More than rational, this choice is optimal as it prevents an inefficient occurrence of conflict.

In such a context, the population is threatened by their environment as much as by their leader. Indeed, the leader can use conflict as an extraction tool because of the threat of losing the conflict if the population decides to rebel. Therefore, it is the very existence of potential enemy that becomes the threat. The more hostile this enemy, the bigger ϕ, the bigger the threat. Indeed, a very hostile enemy would make any forfeit terrible, thus granted a huge extraction power to the leader. This in terms might mean that the population is better off with very inefficient institutions, i.e. a very high ρ. Notice that neither forfeit, nor conflict actually needs to occur; the mere threat of the environment is sufficient.

These insights might shed new light on the contagion of bad institutions. A country with weak institutions would be a rather menacing enemy, leading neighboring countries to fall prey to their own leaders, and to either weaken their institutions willingfully, or to be impoverished by wars. Because the prevalence of commitment is positively linked to ϕ, one would expect threatening environments to unambiguously weaken institutions. In a similar reasoning, one could justify democratic peace through an opposite virtuous circle: a democratic country would not be much of a threat to another democratic country, were one to forfeit a conflict after its leader initiated it. This would not give a leader any extraction power; thus, neither violence nor depletion of institutions should be feared.

7 Conclusion

This paper wonders whether diversionary conflicts can exist in a rational world. They can. It questions the rationality of rallies around the flag. They are. It further inquires about the possibility for a leader to initiate war with the only goal of generating such rallies around the flag. She can. It finally examines the optimality of voluntary enchainment through restriction to rebellion means in the long run. It is.

I first describe how a conflict can be successfully used to avoid rebellion even when no agent is naive. To do so, I formalize a mechanism that has been alluded to in much of the diversionary literature in political sciences. I argue that beyond patriotism and diversion, insurgency is avoided through fear; in particular, the fear of weakening one's international position, whose intensity is intuitively represented by the *threat* posed by the enemy. In this context, a leader does not need to *win* a war, as argued by the signalling interpretation of Richards et al. [1993], but to *stay* at war. Likewise, she does not need a patriotic or distractible population, as alluded by Levy [1989], but a fierce enemy. A belligerent strategy can then be beneficial to the leader if the decrease in tax base due to the cost of conflict is more than compensated by the increase in tax rate allowed by the state of war.

I then show how rallies around the flag can be interpreted as a commitment to internal peace. Indeed, supporting an aggressive foreign policy tends to make a one's enemies more hostile towards oneself. Thus, rallying around the flag allows a population's leader to implement harsher internal policies without having to resort to open war. Such rally effects in fact occur in equilibrium. Rallying around the flag is thus rational: it is the lesser of two evils; and it is efficient: it prevents further destruction of resources through war.

Because in the setup, a leader's rationale to use force is always that of preventing a rebellion while implementing demanding policies, such conflicts occur in equilibrium. In this sense, diversionary conflicts do occur. However, I further wonder whether conflict can be initiated in the only goal of gathering the popular support about foreign policy, rather than just domestic obedience. This is the considered definition of *diversionary incentive.* I find that such incentive exists. Actually, the set of parameters for which they exist is non-zero measure.

Finally, I explore the long-run effects of a threatful environment on a country's rebellion opportunities. A population might want to purposely set up barriers to rebellion in order to prevent the leader from using inefficient extraction tool. I prove that, indeed, this occurs in equilibrium. By a slight shift in the interpretation of freedom of rebellion, seen as general strong institutions, this results sheds light on the spread of weak institutions: a country whose institutions are weak is a more serious threat to its neighbors, who might, in turn, want to set up weak institutions to preserve peace. I also show how, for other parameters, the population does not commit and conflicts occur. The prevalence of such conflicts is non-monotonically related to the intensity of the enemy's threat is hump shaped. The function is hump-shaped. It reflects two opposite forces: a more hostile country renders peace more valuable, but also more expensive, as the leader is harder to dissuade from war.

This work is relevant in that it reclaims an argument that has been extensively discussed in political sciences, refines it and formalizes the mechanism that underlies it. This allows to rationalize altogether diversionary conflicts and rallies around the flag; it also permits to theoretically prove the existence of pure diversionnary incentives; and to question long terms effect on institutions. Because the mechanism is simple and the setup pared-down, the predictions about the role of each parameters should be clear enough to be brought to the data. Such rigorous empirical analysis could test the empirical credibility of the conclusions presented in this paper while potentially restoring some consensus to the discordant empirical literature on diversionary wars.

References

Acemoglu, D. and Robinson, J. A. (2005). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Arena, P. and Bak, D. (2013). Diversionary incentives, rally effects, and crisis bargaining. *Foreign Policy Analysis*.

Bodin, J. and Tooley, M. J. (1955). *Six books of the commonwealth*, volume 56. B. Blackwell Oxford.

Chiou, J.-Y., Dincecco, M., and Johnston, T. (2014). Autocratic responses to revolutionary threats?

Chiozza, G., Gleditsch, K. S., and Goemans, H. (2004). Civil war, tenure, and interstate insecurity. *Typescript, Harvard University, University of California, San Diego, and the University of Rochester*.

Davies, G. A. (2016). Policy selection in the face of political instability: Do states divert, repress, or make concessions? *Journal of Conflict Resolution*, 60(1):118–142.

Downs, G. W. and Rocke, D. M. (1994). Conflict, agency, and gambling for resurrection: The principal-agent problem goes to war. *American Journal of Political Science*, pages 362–380.

Fearon, J. D. (1995). Rationalist explanations for war. *International organization*, 49(03):379–414.

Gagnon Jr, V. P. (2006). *The myth of ethnic war: Serbia and Croatia in the 1990s*. Cornell University Press.

Gent, S. E. (2009). Scapegoating strategically: Reselection, strategic interaction, and the diversionary theory of war. *International Interactions*, 35(1):1–29.

Hagan, J. D. (2017). Diversionary theory of war in foreign policy analysis. In *Oxford Research Encyclopedia of Politics*.

Hess, G. D. and Orphanides, A. (1995). War politics: An economic, rational-voter framework. *The American Economic Review*, pages 828–846.

Jackson, M. O. and Morelli, M. (2007). Political bias and war. *The American Economic Review*, pages 1353–1373.

Jackson, M. O. and Morelli, M. (2011). The reasons for wars: an updated survey. *The Handbook on the Political Economy of War*, page 34.

Jung, S. C. (2014). Foreign targets and diversionary conflict. *International Studies Quarterly*, 58(3):566–578.

Kozlov, V. A. and MacKinnon, E. M. (2002). *Mass uprisings in the USSR: protest and rebellion in the post-Stalin years*. ME Sharpe.

Levy, J. S. (1989). The diversionary theory of war: A critique. *Handbook of war studies*, pages 259–88.

Mayer, A. J. (1969). Internal causes and purposes of war in europe, 1870-1956: A research assignment. *The Journal of Modern History*, pages 292–303.

Murray, S. (2017). The ?rally-'round-the-flag? phenomenon and the diversionary use of force. In *Oxford Research Encyclopedia of Politics*.

Oakes, A. (2012). *Diversionary war: Domestic unrest and international conflict*. Stanford University Press.

Padró i Miquel, G. (2007). The control of politicians in divided societies: The politics of fear. *The Review of Economic Studies*, 74(4):1259–1274.

Powell, J. M. (2014). Regime vulnerability and the diversionary threat of force. *Journal of Conflict Resolution*, 58(1):169–196.

Richards, D., Morgan, C. T., Wilson, R. K., Schwebach, V. L., and Young, G. D. (1993). Good times, bad times, and the diversionary use of force a tale of some not-so-free agents. *Journal of Conflict Resolution*, 37(3):504–535.

Tarar, A. (2006). Diversionary incentives and the bargaining approach to war. *International Studies Quarterly*, 50(1):169–188.

Theiler, T. (2018). The microfoundations of diversionary conflict. *Security Studies*, 27(2):318–343.

Ticchi, D. and Vindigni, A. (2008). War and endogenous democracy. *Available at SSRN 1136202*.

Wawro, G. (2005). *The Franco-Prussian War: The German Conquest of France in 1870-1871*. Cambridge University Press.

Zubok, V. M. (2009). *A failed empire: the Soviet Union in the Cold War from Stalin to Gorbachev*. Univ of North Carolina Press.

Appendix

A Logical Conditions and Equilibrium Outcomes

A.1 List of Conditions for SPE and their Logical Relationships

First, let us list the conditions on parameters discussed in the backward induction of Section ??.

Condition	Name
$\kappa_t \ \geq \ \phi_t(1-\rho)$	A
$\kappa_w \ \geq \ \phi_w(1-\rho)$	B
$\kappa_t \ \geq \ \phi_t(1-\rho') + (\rho' - \rho)$	C
$\kappa_w - \kappa_t \ \geq \ (1-\rho)(\phi_w - \phi_t)$	D
$\kappa_w - \kappa_t \ \geq \ (1-\rho')(\phi_w - \phi_t)$	E
$\frac{1-\phi_w}{1-\phi_t} \ \geq \ \frac{1-\rho'}{1-\rho}$	F

Table 1: List of the Conditions on Parameters

From them, and using Lemma 1, the four outcomes mentioned in Theorem 1 can be derived, as seen in Table 2.

D		$\bar{D}$ *and* E				$\bar{E}$	
		F		$\bar{F}$			
A	$\bar{A}$	B	$\bar{B}$	C	$\bar{C}$	B	$\bar{B}$
$\theta = 0$ – –	$\theta = 1$ $\eta = 0$ $\omega = 0$	$\theta = 0$ – –	$\theta = 1$ $\eta = 0$ $\omega = 1$	$\theta = 0$ – –	$\theta = 1$ $\eta = 1$ $\omega = 0$	$\theta = 0$ – –	$\theta = 1$ $\eta = 0$ $\omega = 1$
peace -	tensions ~~support~~	peace -	war ~~support~~	peace -	tensions support	peace -	war ~~support~~

Table 2: Equilibrium Path by Parameters' Conditions

Some of the conditions along a given equilibrium path are redundant. First, let us recall that $\kappa_w > \kappa_t, \phi_w > \phi_t$ and $\rho' > \rho$. From there, we find that:

- $C \Rightarrow A$, as $\kappa_t \ \geq \ \phi_t(1-\rho') + (\rho' - \rho) \ > \phi_t(1-\rho') + \phi_t(\rho' - \rho) = \phi_t(1-\rho)$.

- $D \Rightarrow E$, as $\kappa_w - \kappa_t \geq (1-\rho)(\phi_w - \phi_t) > (1-\rho')(\phi_w - \phi_t)$.
- $\bar{A}$ *and* $B \Rightarrow D$, as $-\kappa_t > -\phi_t(1-\rho)$, and $\kappa_w \geq \phi_w(1-\rho)$ implies $\kappa_w - \kappa_t > (1-\rho)(\phi_w - \phi_t)$.
- A *and* $\bar{B} \Rightarrow \bar{D}$, as $-\kappa_t \leq -\phi_t(1-\rho)$ and $\kappa_w < \phi_w(1-\rho)$ implies $\kappa_w - \kappa_t < (1-\rho)(\phi_w - \phi_t)$.
- C *and* $E \Rightarrow B$, as $\kappa_t \geq (1-\rho')\phi_t + \rho' - \rho$ and $\kappa_w - \kappa_t \geq -(1-\rho')\phi_t + (1-\rho')\phi_w$ implies $\kappa_w \geq (1-\rho')\phi_w + \rho' - \rho > (1-\rho')\phi_w + \phi_w(\rho' - \rho) = (1-\rho)\phi_w$.
- C *and* $F \Rightarrow B$, as F can be rewritten as $\rho'(1-\phi_t) + \phi_t - \rho \geq (1-\rho)\phi_w$; and with $\kappa_t \geq \rho'(1-\phi_t) + \phi_t - \rho$, it implies $\kappa_w > \kappa_t \geq \phi_w(1-\rho)$.

A.2 Derivation of the Logical Condition by Outcome

In this subsection, I derive conditions grouped by outcome.

To simplify exposition, I use set notation. Referring to the conditions in Table 1, I denote:

$$X = \left\{(\kappa_w, \kappa_t, \phi_w, \phi_t, \rho, \rho') \in (0;1)^6 \ : \ \textit{condition } X \textit{ fulfilled}\right\} \quad \forall X \in \{A, B, C, D, E, F\}$$

Table 3 summarizes the logical relations between the conditions derived above, along with their set notation.

Logical Condition:	Equivalent for Sets:	Logical Condition:	Equivalent for Sets:
$C \Rightarrow A$	$C \subset A$	$\bar{A}$ *and* $B \Rightarrow D$	$\bar{A} \cap B \subset D$
$D \Rightarrow E$	$D \subset E$	A *and* $\bar{B} \Rightarrow \bar{D}$	$A \cap \bar{B} \subset \bar{D}$
C *and* $F \Rightarrow B$	$C \cap F \subset B$	C *and* $E \Rightarrow B$	$C \cap E \subset B$

Table 3: Summary of the Conditions' Interrelation

It is useful to acknowledge the following implications of Table 3:

- $C \cap A = C$
- $D \cap E = D$
- $\bar{D} \cup \bar{E} = \bar{D}$
- $\bar{D} \cap \bar{E} = \bar{E}$
- $(A \cup \bar{B}) \cap \bar{D} = \bar{D}$
- $(\bar{A} \cup B) \cap D = D$
- $(C \cap E) \cup B = B$
- $C \cap E \cap B = C \cap E$

A.2.1 Condition for War

A high-level conflict can occur in either of the cases described in Table 2. We can write:

$(\bar{D} \cap E \cap F \cap \bar{B}) \cup (\bar{E} \cap \bar{B}) = \bar{B} \cap [\bar{E} \cup (\bar{D} \cap E \cap F)] = \bar{B} \cap (\bar{E} \cup \bar{D}) \cap (\bar{E} \cup F)$

Therefore, war is defined by:

$$\boxed{\bar{B} \cap \bar{D} \cap (\bar{E} \cup F)}$$

A.2.2 Condition for Tensions

A low-level conflict that meet no popular enthusiasm is straightforward to define as:

$$\boxed{\bar{A} \cap D}$$

On the other hand, international tensions that are supported exist when:

$$\boxed{\bar{C} \cap \bar{D} \cap E \cap \bar{F}}$$

A.2.3 Condition for Peace

Let us re-write the conditions for each state of peace. We have:

- $A \cap D = A \cap (\bar{A} \cup B) \cap D = A \cap B \cap D$
- $B \cap \bar{D} \cap E \cap F = B \cap (A \cup \bar{B}) \cap \bar{D} \cap E \cap F = A \cap B \cap \bar{D} \cap E \cap F$
- $B \cap \bar{E} = B \cap \bar{D} \cap \bar{E} = B \cap (A \cup \bar{B}) \cap \bar{D} \cap \bar{E} = B \cap A \cap \bar{D} \cap \bar{E} = A \cap B \cap \bar{E}$
- $C \cap E \cap \bar{D} \cap \bar{F} = C \cap E \cap B \cap \bar{D} \cap \bar{F} = A \cap C \cap E \cap B \cap \bar{D} \cap \bar{F}$

Therefore, we can define peace as:

$$\begin{aligned}
&(A \cap D) \cup (B \cap \bar{D} \cap E \cap F) \cup (C \cap E \cap \bar{D} \cap \bar{F}) \cup (B \cap \bar{E}) \\
&= (A \cap B \cap D) \cup (A \cap B \cap \bar{D} \cap E \cap F) \cup (A \cap B \cap C \cap \bar{D} \cap E \cap \bar{F}) \cup (A \cap B \cap \bar{E}) \\
&= A \cap B \cap [D \cup (\bar{D} \cap E \cap F) \cup (C \cap \bar{D} \cap E \cap \bar{F}) \cup \bar{E}] \\
&= A \cap B \cap \{D \cup \bar{E} \cup [\bar{D} \cap E \cap (F \cup (C \cap \bar{F}))]\} \\
&= A \cap B \cap \{D \cup \bar{E} \cup [\bar{D} \cap E \cap (F \cup C)]\}
\end{aligned}$$

It follows that peace exists when:

$$\boxed{A \cap B \cap (D \cup \bar{E} \cup F \cup C)}$$

While not minimal, I keep this formulation in order to underline its intuitive meaning. Note that the expression also brings:

$$\begin{aligned}
&A \cap B \cap (C \cup F \cup D \cup \bar{E}) \\
&= [A \cap B \cap (C \cup \bar{E})] \cup (A \cap B \cap D) \cup (A \cap B \cap F) \\
&= \{A \cap [(B \cup (C \cap E)) \cap (C \cup \bar{E})]\} \cup (A \cap B \cap D) \cup (A \cap B \cap F) \\
&= \{A \cap [(B \cup C \cup \bar{E}) \cap ((C \cap E) \cup C \cup \bar{E})]\} \cup (A \cap D) \cup (A \cap B \cap F) \\
&= \{A \cap [(B \cup C \cup \bar{E}) \cap (C \cup \bar{E})]\} \cup (A \cap D) \cup (A \cap B \cap F) \\
&= [A \cap (C \cup \bar{E})] \cup (A \cap D) \cup (A \cap B \cap F) \\
&= C \cup [A \cap (\bar{E} \cup D)] \cup (A \cap B \cap F)
\end{aligned}$$

Therefore, condition C is sufficient but not necessary to the preservation of peace.

B Alternative Payoff Specification

B.1 $\eta = 1$ affects the overall payoff

Instead of modifying ρ, let us assume that support $\eta = 1$ implies a change in the total payoff that becomes $(1-p)(1-\rho)(1-\theta(h))$. Similar to the notation in the main text, for every partial history $h \in H_3$, I define $p(h) = \begin{cases} p & \text{if } h \in \{(\theta = 1, \eta = 1, \omega = 0), (\theta = 1, \eta = 1, \omega = 1)\} \\ 0 & \text{otherwise} \end{cases}$.

Furthermore, I consider the following alternative definitions: $\underline{\kappa}_w = (1-p)(1-\rho)\phi_t$ and $\underline{\kappa}_w = (1-p)(1-\rho)\phi_w$. I can thus reformulate Lemma 1, Theorem 1.

. Lemma 1.B.1 *The SPE is as follows:*

- *The leader's equilibrium strategy is:*
  - $\theta = 1$ *if* $\kappa_w < \bar{\kappa}_w$; *or if* $\kappa_t < \bar{\kappa}_t$; *or if* $\underline{\kappa}_w - \underline{\kappa}_t \le \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$, *and* $r_\phi < (1-p)$, *and* $\kappa_t < \underline{\kappa}_t + p(1-\rho)$;
    $\theta = 0$ *otherwise.*
  - $\omega(\eta = 0) = 1$ *if* $\kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$; $\omega(\eta = 0) = 0$ *otherwise.*
    $\omega(\eta = 1) = 1$ *if* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$; $\omega(\eta = 1) = 0$ *otherwise.*
  - $\forall h \in H_3, \tau(h) = 1 - \frac{(1-p(h))(1-\rho)(1-\phi(h))}{1-\kappa(h)}$
- *The population's equilibrium strategy is:*
  - $\eta = 1$ *if* $\underline{\kappa}_w - \underline{\kappa}_t \le \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$ *and* $r_\phi < (1-x)$; $\eta = 0$ *otherwise.*
  - $\forall h \in H_3 \times [0,1]$ *and corresponding* $h' \in H_3$:
    $\psi(h) = 1$ *if* $\tau > 1 - \frac{(1-p(h'))(1-\rho)(1-\phi(h'))}{1-\kappa(h')}$; $\psi(h) = 0$ *otherwise.*

Proof. This follows from backward induction as carried out in the main text, using the new specification of payoff. □

. Theorem 1.B.1 *There are four possible outcomes: peace, (unsupported) war, unsupported tensions and supported tensions.*

- *War occurs iff* $\kappa_w < \bar{\kappa}_w$; *and* $\kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$; *and either* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$, *or* $r_\phi \ge 1 - p$.
- *Supported tensions occur iff* $\underline{\kappa}_w - \underline{\kappa}_t \le \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$; *and* $r_\phi < 1 - p$; *and* $\kappa_t < \underline{\kappa}_t + p(1-\rho)$.
- *Unsupported tensions occur iff* $\kappa_t < \bar{\kappa}_t$; *and* $\kappa_w - \kappa_t \ge \bar{\kappa}_w - \bar{\kappa}_t$.
- *Peace occurs iff* $\kappa_w \ge \bar{\kappa}_w$; *and* $\kappa_t \ge \bar{\kappa}_t$; *and either* $\kappa_t \ge \underline{\kappa}_t + p(1-\rho)$, *or* $r_\phi \ge 1 - p$, *or* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$, *or* $\kappa_w - \kappa_t \ge \bar{\kappa}_w - \bar{\kappa}_t$ *.*

Proof. I show that all implications shown in Appendix A.1 still hold. From there, Appendix A.2 applies directly.

Conditions C, E and F have changed. Condition C now reads $\kappa_t \geq (1-p)(1-\rho)\phi_t + p(1-\rho)$, condition E is $\kappa_w - \kappa_t \geq (1-p)(1-\rho)(\phi_w - \phi_t)$ and F becomes $r_\phi \geq 1-p$. All implications concerned by this change still hold. Indeed:

- $C \Rightarrow A$, as $\kappa_t \geq (1-p)(1-\rho)\phi_t + p(1-\rho) = \phi_t(1-\rho) + p(1-\rho)(1-\phi_t) > \phi_t(1-\rho)$.
- $D \Rightarrow E$, as $\kappa_w - \kappa_t \geq (1-\rho)(\phi_w - \phi_t) > (1-p)(1-\rho)(\phi_w - \phi_t)$.
- C *and* $E \Rightarrow B$, as $\kappa_t \geq (1-p)(1-\rho)\phi_t + p(1-\rho)$ and $\kappa_w - \kappa_t \geq -(1-p)(1-\rho)\phi_t + (1-p)(1-\rho)\phi_w$ implies $\kappa_w \geq (1-p)(1-\rho)\phi_w + p(1-\rho) = (1-\rho)\phi_w + p(1-\rho)(1-\phi_w) > (1-\rho)\phi_w$.
- C *and* $F \Rightarrow B$, as C is rewritten $\kappa_t \geq \phi_t(1-\rho) + p(1-\rho)(1-\phi_t)$; and F means $p \geq 1 - \frac{1-\phi_w}{1-\phi_t}$; so $\kappa_w > \kappa_t \geq \phi_t(1-\rho) + p(1-\rho)(1-\phi_t) \geq \phi_t(1-\rho) + (1-\rho)(1-\phi_t) - (1-\rho)(1-\phi_w) = \phi_w(1-\rho)$.

□

Furthermore, let us adapt the proof of Theorem 2 in this context. In particular, I provide an alternative characterization for $\mathcal{D}' \subseteq \mathcal{D}$. I let the reader verify that all parameters are acceptable and that all conditions of Corollary 1 hold for:

$$\mathcal{D}' = \left\{ (\kappa_t, \kappa_w, \phi_t, \phi_w, \rho, \rho') : \forall (x, \epsilon) \in \mathcal{E}, \begin{array}{lll} \rho = x; & \phi_t = x + \epsilon_2; & \kappa_t = x + \epsilon_4; \\ p = \frac{\epsilon_1}{1-x}; & \phi_w = x + \epsilon_3; & \kappa_w = x + \epsilon_5; \end{array} \right\}$$

where $\mathcal{E}$ is as before

B.2 $\eta = 1$ affects the parameters ϕ

Instead of modifying ρ, let us assume that support $\eta = 1$ implies a change in ϕ. In particular, $\phi(\theta = 1, \eta = 1, \omega = 1) = \phi'_w > \phi_w$ and $\phi(\theta = 1, \eta = 1, \omega = 0) = \phi'_t > \phi_t$. Furthermore, I impose $\phi_w - \phi_t > \phi'_w - \phi'_t$. This ensures that the benefit of escalation is higher without than with support. Without this assumption, there is no room for support, as it would only give the leader an additional incentive to escalate the conflict to war.
I consider the following alternative definitions: $\underline{\kappa}_w = \phi'_t(1-\rho)$ and $\underline{\kappa}_w = \phi'_w(1-\rho)$. I can thus reformulate Lemma 1, Theorem 1.

. Lemma 1.B.2 *The SPE is as follows:*

- *The leader's equilibrium strategy is:*
  - $\theta = 1$ *if* $\kappa_w < \bar{\kappa}_w$; *or if* $\kappa_t < \bar{\kappa}_t$; *or if* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$, *and* $\phi'_t < \phi_w$, *and* $\kappa_t < \underline{\kappa}_t$;
    $\theta = 0$ *otherwise.*
  - $\omega(\eta = 0) = 1$ *if* $\kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$; $\omega(\eta = 0) = 0$ *otherwise.*
    $\omega(\eta = 1) = 1$ *if* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$; $\omega(\eta = 1) = 0$ *otherwise.*
  - $\forall h \in H_3, \tau(h) = 1 - \frac{(1-\phi(h))(1-\rho)}{1-\kappa(h)}$

- *The population's equilibrium strategy is:*
  - $\eta = 1$ *if* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$ *and* $\phi'_t < \phi_w$*;* $\eta = 0$ *otherwise.*
  - $\forall h \in H_3 \times [0,1]$ *and corresponding* $h' \in H_3$*:*
    $\psi(h) = 1$ *if* $\tau > 1 - \frac{(1-\phi(h'))(1-\rho)}{1-\kappa(h')}$*;* $\psi(h) = 0$ *otherwise.*

Proof. This follows from backward induction as carried out in the main text, using the new specification of payoff. □

. **Theorem 1.B.2** *There are four possible outcomes: peace, (unsupported) war, unsupported tensions and supported tensions.*

- *War occurs iff* $\kappa_w < \bar{\kappa}_w$*; and* $\kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$*; and either* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$*, or* $\phi'_t \geq \phi_w$*.*
- *Supported tensions occur iff* $\underline{\kappa}_w - \underline{\kappa}_t \leq \kappa_w - \kappa_t < \bar{\kappa}_w - \bar{\kappa}_t$*; and* $\phi'_t < \phi_w$*; and* $\kappa_t < \underline{\kappa}_t$*.*
- *Unsupported tensions occur iff* $\kappa_t < \bar{\kappa}_t$*; and* $\kappa_w - \kappa_t \geq \bar{\kappa}_w - \bar{\kappa}_t$*.*
- *Peace occurs iff* $\kappa_w \geq \bar{\kappa}_w$*; and* $\kappa_t \geq \bar{\kappa}_t$*; and either* $\kappa_t \geq \underline{\kappa}_t$*, or* $\phi'_t \geq \phi_w$*, or* $\kappa_w - \kappa_t < \underline{\kappa}_w - \underline{\kappa}_t$*, or* $\kappa_w - \kappa_t \geq \bar{\kappa}_w - \bar{\kappa}_t$ *.*

Proof. I show that all implications shown in Appendix A.1 still hold. From there, Appendix A.2 applies directly.
Conditions C, E and F have changed. Condition C now reads $\kappa_t \geq \phi'_t(1-\rho)$, condition E is $\kappa_w - \kappa_t \geq (1-\rho)(\phi'_w - \phi'_t)$ and F becomes $\phi'_t \geq \phi_w$. All implications concerned by this change still hold. Indeed:

- *C* ⇒ *A*, as $\kappa_t \geq \phi'_t(1-\rho) > \phi_t(1-\rho)$.
- *D* ⇒ *E*, as $\kappa_w - \kappa_t \geq (1-\rho)(\phi_w - \phi_t) > (1-\rho)(\phi'_w - \phi'_t)$.
- *C and E* ⇒ *B*, as $\kappa_t \geq (1-\rho)\phi'_t$ and $\kappa_w - \kappa_t \geq -(1-\rho)\phi'_t + (1-\rho)\phi'_w$ implies $\kappa_w \geq (1-\rho)\phi'_w > (1-\rho)\phi_w$.
- *C and F* ⇒ *B*, as $\kappa_t \geq \phi'_t(1-\rho) > \phi_w(1-\rho)$ implies $\kappa_w > \kappa_t \geq \phi_w(1-\rho)$.

□

Furthermore, let us adapt the proof of Theorem 2 in this context. In particular, I provide an alternative characterization for $\mathcal{D}' \subseteq \mathcal{D}$. I let the reader verify that all parameters are acceptable and that all conditions of Corollary 1 hold for:

$$\mathcal{D}' = \left\{ (\kappa_t, \kappa_w, \phi_t, \phi_w, \phi'_t, \phi'_w, \rho) : \forall (x,\epsilon) \in \mathcal{E}, \rho = x \text{ and } \begin{array}{lll} \phi_t = x + \epsilon_2; & \phi'_t = x + \epsilon_2 + \epsilon_{01}; & \kappa_t = x + \epsilon_4; \\ \phi_w = x + \epsilon_3; & \phi'_w = x + \epsilon_3 + \epsilon_{02}; & \kappa_w = x + \epsilon_5; \end{array} \right\}$$

$$\text{where } \mathcal{E} = \left\{ (x,\epsilon) \in (0,1)^6 : \begin{array}{c} (1-x)(\epsilon_3 - \epsilon_2) > \epsilon_5 - \epsilon_4 > \epsilon_{01} - \epsilon_{02} > 0 \\ \epsilon_5 > \epsilon_3 \end{array} \right\}$$