Asymmetric Information and Third-Party Intervention in Civil Wars*

J. Atsu Amegashie

Department of Economics
University of Guelph
Guelph, Ontario
Canada N1G 2W1

E-mail: jamegash@uoguelph.ca
Phone: 519-824-4120 ext. 58945
Fax: 519-763-8497

March 3, 2012

Abstract

Some empirical studies have found that third-party intervention could worsen civil conflicts. However, exactly why this might be the case is not clear. This paper builds a model to explain how a third-party's *expected* intervention in a conflict could worsen the conflict. I study a two-period model of conflict (contest) with two combatants and a third party who is an ally of one of the combatants. The third party is fully informed about the type of her ally but not about the type of her ally's enemy. There is a signaling game with one receiver (i.e., the third party) and two senders (i.e., the two combatants in the conflict). *If the third party will not intervene in a big way*, then there exists a unique perfect Bayesian separating equilibrium in which the third party's *expected* intervention worsens the conflict by energizing her ally's enemy wherein he (i.e., the enemy) exerts more effort than he would in the absence of third-party intervention; this increases aggregate effort in the conflict. Therefore, if the third party will not intervene in a big way, then it might be better not to intervene at all. The third party ignores the signals of his ally, although he (i.e., the ally) is fully informed about the enemy's type. Finally, if *expected* intervention worsens conflicts, then empirical work may overstate any positive effect of *actual* intervention on conflicts.

Keywords: Perfect Bayesian equilibrium, conflict, intuitive criterion, signaling, third-party intervention.
JEL Classification: D72, D74.
*I thank Rene Kirkegaard, Asha Sadanand, Dana Sisak, Ruqu Wang, and Jun Zhang for very helpful comments on an earlier version of this paper. My thanks are also due to Olivier Bos, Bram Cadsby, and Osman Ouattara. This paper is a *substantially revised* version of a paper that was previously titled "Third party intervention in conflicts and the indirect Samaritan's dilemma", circulated as CESifo working paper #2711, and presented at the Royal Military College (Canada), the 2011 CESifo ESP conference in Munich, and the 2011 workshop on contests and tournaments at the Technical University of Berlin.

1. Introduction

The literature on intra-state conflicts is large and growing.[1] Within this literature, research on third-party intervention in intra-state conflicts has received very little attention from economists.

In the post-world war II era and the end of the cold war, there have been numerous third-party interventions in intra-state conflicts. These interventions have been in places such as Bosnia, Somalia, Haiti, the former Soviet republics, Vietnam, and Cambodia and have involved countries like Britain, China, France, the USA and international organizations like the UN. Between 1944 and 1999, Regan (2002) identified 150 intrastate conflicts of which 101 had third-party interventions. He found that third-party interventions tend to worsen conflicts (see also Regan, 2000). Elbadawi and Sambanis (2000) also obtained a similar result in their empirical work. And Diehl et al. (1996) made a similar claim in the case of UN interventions.[2]

Still on the preceding point, Lacina (2006) and Heger and Salehyan (2007) report that civil wars with interventions have significantly higher fatality levels than civil wars without interventions. Of course, there is the identification problem of whether interventions make civil wars worse, or whether really bad civil wars are more likely to

[1] See Blattman and Miguel (2010) and Collier and Hoeffler (2007) for surveys of the literature.

[2] These papers used the duration of conflict to evaluate third –party interventions. Hence, third-party intervention worsens a conflict if it extends the duration of the conflict. In this paper, I use the effect of third-party intervention on the aggregate effort by the warring faction to evaluate the third-party's intervention. This is reasonable because there is likely to be a positive correlation between the duration of conflict and the magnitude of the social costs (i.e., loss of life and property). In the related case of mediation, as opposed to military intervention, Regan and Stam (2000) undertook an empirical analysis of third party intervention in conflicts and found that mediation in the earlier stages of a conflict are more effective and that late mediations worsen conflicts. Dixon (1996) found that mediation efforts and third-party activities to open or maintain lines of communication are the most effective to resolving conflicts.

provoke intervention. As pointed out below, this explains why some scholars have used expected intervention to deal with the endogeneity of actual intervention.

Interventions are usually biased where the third party supports one of the factions in the conflict. For example, using data from the International Crisis Behavior project, Carment and Harvey (2000) found that 140 out of 213 interventions in intrastate conflicts over the period 1918-1994 were clearly biased. In the post-war period, Regan (2000) also found that most interventions were biased. In his empirical work, Regan (2002) found that neutral interventions were less effective in ending conflicts than biased interventions. Betts (1996) argued that the idea of impartial intervention is a delusion and authors such as Watkins and Winters (1997) and Favretto (2009) have argued that biased interventions may be desirable.

While some of the aforementioned empirical studies have found that third-party intervention could worsen conflicts, exactly why this might be the case is not clear. This paper builds a model to explain how a third-party's intervention in a conflict could worsen the conflict. There is a literature that emphasizes the moral hazard effects of third-party intervention in conflicts (e.g., Rowlands and Carment, 1998, 2006; Kuperman, 1996). This literature suggests that domestic groups which would not otherwise resort to political violence or escalate an ongoing conflict may be encouraged to do so by the prospect of outside support.[3] Explicit models of this argument include Rowlands and Carment (1998, 2006) and Favretto (2009). However, they treat the size of the third-

[3]In a May 31, 2006 op-ed in the New York Times, Alan Kuperman claimed that intervention in Darfur was emboldening the rebels to fight on because the rebels who benefited from intervention rejected a proposed agreement. A similar argument has been made by some scholars with respect to the Kosovo crisis (see Grigorian, 2005). However, others do not think that the moral hazard argument is satisfactory (e.g., Grigorian, 2005; Crawford, 2005).

party's military support as exogenous.[4] Other scholars argue that a lack of resolve and credibility within coalitions over the use of force create incentives for the escalation of conflicts (Regan, 1996; Diehl et al., 1996; Harvey, 1998; Walter and Snyder, 1999).

In this paper, I focus on biased interventions. I study a conflict with incomplete information and two combatants who fight over two periods. One of the combatants has an ally (i.e., a third party) who wants to assist him with military support in the conflict. The third party is fully informed about the type of her ally but not about the type of her ally's enemy. The third-party's ally is fully informed about the type of his enemy. There is a signaling game with one receiver (i.e., the third party) and two senders (i.e., the two combatants). *If the third party will not intervene in a big way*, then I find that there is a perfect Bayesian separating equilibrium in which the third-party's *expected* intervention worsens the conflict by inducing the enemy of the third-party's ally to over-invest in arms in order to discourage the third-party from helping her ally or to back off entirely from intervening in the conflict in period 2. Hence the enemy of the third-party's ally displays some bravado (i.e., overinvests in arms). Not only does third-party intervention lead to an increase in the effort of the ally's enemy, it also leads to an increase in the *aggregate* effort (i.e., the sum of the factions' efforts) in the conflict.

The paper also shows that a third-party may rationally mistrust his ally by ignoring the ally's private and valuable information. Such mistrust may exist between

[4] Favretto (2009) uses a model of incomplete information in the framework of Fearon's (1994, 1995) crisis-bargaining model. In contrast to the present model, she assumes exogenous probabilities of success for the warring factions. Rowland and Carment (1998, 2006) assume both complete information and exogenous military support by the third party. Using a crisis bargaining model with no signaling, Werner (2000) considers a model of incomplete information to determine a third-party's incentives to intervene in a conflict. She finds that an attacker can manipulate the stakes of war by making it low enough that, for a third party, the benefits of intervention do not justify the costs (see also Yeun, 2009). In my model, no one can manipulate the stakes of the conflict. And in Werner (2000), the threat of intervention does not worsen the conflict. I discuss the differences between my model and crisis-bargaining models in a subsequent footnote.

allies who could be described as strange bedfellows. An example of this mistrust is the occasional claims by the USA that its allies (i.e., the governments of Pakistan, Iraq, and Afghanistan) are not doing enough to rein in their enemies in their various countries.

The result that the third-party's intervention may worsen the conflict if the size of intervention is not big enough is consistent with the literature which argues that lack of resolve within coalitions over the use of force create incentives for the escalation of conflicts. In section 3.1, I return to this issue and explain why it is only relevant to particular reaction to expected intervention.

To deal with the endogeneity of third-party intervention, Elbadawi and Sambanis (2000) used *expected* intervention instead of actual intervention as the regressor in their empirical work. They found that *expected* third-party interventions worsen conflicts. Akcinaroglu and Radziszewski (2005) also reach a similar conclusion in their empirical work. These empirical papers lend some support to the theoretical result that *expected* intervention may worsen a conflict. Furthermore, as Wagner (2007, p. 229) observed: "expectations about possible interventions may play a role in motivating an internal conflict even if outsiders never intervene in it."

That *expectation* of third-party support could increase the intensity of conflict, at least, for the ally's enemy has implications for empirical work. It suggests that conflicts may worsen prior to a publicly known and biased intervention and may improve after the intervention. Indeed, this effect may exist even if the third party has already intervened in the conflict, so long as she is still not fully informed about the type of her ally's enemy. Hence, part of the subsequent reduction in the intensity of the conflict is not necessarily due to actual intervention *per se*. Therefore, while a biased intervention by a third party may have a positive effect on a conflict, this effect may be overstated in empirical work (i.e., the relevant coefficient in regressions may be biased upwards).[5]

In my two-period model, the intensity of conflict is increased not only because of *expectation* of the third-party's intervention as is the case in period 1, but because the third-party's *actual* intervention reduces the ally's marginal cost of conflict leading the ally to be more aggressive in period 2 than he otherwise would have been. This is consistent with the moral hazard literature of *actual* intervention mentioned above. However, this effect, while present in my model, is not the focus of this paper.

The main point of this paper is *not* to make the *general* point that signaling can lead to undesirable outcomes. A contribution of this paper is to theoretically explain how *expected* third-party intervention could worsen conflicts. Furthermore, as discussed above, other interesting insights such as the implications of the results for econometric work and the conditions under which third-party intervention may worsen conflicts emerge from the analysis. In section 3, I discuss alternative models and compare them with my model.

The paper is organized as follows. Section 2 presents a model and analyses of signaling in a conflict. Section 3 discusses issues of robustness and section 4 concludes the paper.

[5] Notice that this argument is different from the argument in Elbadawi and Sambanis (2000) and Akcinaroglu and Radziszewski (2005) for using expected intervention as an explanatory variable in a regression. Unlike these authors, I am not arguing that expected intervention should be used as an instrumental variable in order to deal with the endogeneity of actual intervention. I am arguing that expected intervention and actual intervention may be distinct explanatory variables in a regression that seeks to explain the intensity of conflict. Still, their finding suggests that expected intervention could have a negative effect on conflicts.

2. A model of third-party intervention and signaling

I use a well-known model of conflict and contests (e.g., see Epstein and Gang 2009; Konrad, 2009) which assumes that the combatants cannot commit to a peaceful resolution through bargaining. Therefore, costly conflict is inevitable. This model is different from crisis-bargaining models pioneered by Fearon (1994, 1995).[6]

Consider two risk-neutral factions, A and B in a conflict over a region (country). Faction A is an incumbent who governs the country and faction B is a challenger to faction A's rule. There is a risk-neutral third party, C, who is an ally of faction A. There are two time periods, 1 and 2. In each period, there could be a battle (conflict) between A and B. Faction A's valuation of controlling a proportion, $P_A^j \in [0,1]$, of the land (country) in period j is $P_A^j V \geq 0$, where $V > 0$, j = 1, 2. Faction B's corresponding valuation is $P_B^j W_H$ with probability $q \in (0,1)$ and $P_B^j W_L$ with probability $1 - q$, where $W_H > W_L > 0$ and $P_A^j = 1 - P_B^j$. If faction A controls a proportion P_A^j of the land

[6] In a seminal two-player crisis-bargaining model in the shadow of conflict, Fearon (1994, 1995) used costly signaling as an explanation of war. In these models, the threat of conflict, not necessarily actual conflict, is used to make credible demands in crisis bargaining. In my model, conflict is inevitable and I do not allow bargaining between the third party and his ally on one hand and the ally's enemy on the other. Crisis bargaining models are typically used to explain the conditions under which conflict will occur and the role of private information is at the heart of this literature (see Walter, 2009, for a review). I am interested in the conditions under which third-party intervention will escalate an *ongoing* conflict. Signaling in crisis bargaining models is different from signaling in the classical sense. As Fearon (1994, note 2) observed "The crisis signals discussed herein are atypical in that they create costs that are paid only if the signaler takes a certain future action ("backing down") rather than regardless of what the signaler does in the future (as in Spence's classical case)." An example of Fearon's signal is a public announcement by a party that it will use force if its demands are not met. If backing down is sufficiently costly (i.e., Fearon's notion of audience costs) to the party that made the threat, then it becomes a credible signal that he will carry out his threat and therefore incur a cost in the future if his demand is not met. Another example is where a party's offer during bargaining gives a signal about his resolve to go to war, which may weaken or strengthen his bargaining position (e.g., Morrow, 1992). The cost of the offer is only incurred if and only if the offer is accepted. Signaling in my model is in the classical Spence case; cost of signals are immediately incurred when they are sent. Also, in these crisis bargaining models, the combatants' probability of victory in a conflict is typically independent of the effort exerted in the conflict (e.g., Fearon, 1995, 1995; Werner, 2000; Yuen, 2009).

(country), the third-party's valuation is $P_A^j S > 0$, where S > 0. Since a player with a higher valuation in contest is the same as a player with a lower unit cost of exerting effort (e.g., see Clark and Riis, 1998), I shall refer to the high-valuation type of faction B as the strong type and the low-valuation type as the weak type. Because of risk-neutrality, the proportions could alternatively be interpreted as probabilities of victory in the conflict with S, V, and W_k as the players' valuations of victory, k = H, L. Note also that the third-party's valuation, S, is a measure of his strength or military capability and, as will be obvious in the subsequent analysis, the third party will intervene in the conflict if his valuation is sufficiently high.

In *each* period, the factions fight over the *same* piece of land or resource. They invest in a composite military good (hereafter referred to as arms, armed investments, or simply investment) that could be thought of being made up of weapons and soldiers. In period 1, I denote the factions' investments in arms by G and in period 2, I denote it by X. So, for example in period 1, factions A and B invest in G_A and G_B units of arms and control the proportions $P_A^1 = G_A/(G_A + G_B)$ and $P_B^1 = 1 - P_A^1$ of the land (country). A similar function describes the proportions in period 2. Throughout the analysis, G_A, G_B, X_A, and X_B are assumed to be non-negative.

I assume that the third party can help faction A in period 2 but not in period 1. For example, in the case of the USA, this may be due to delays in congressional approval of funding for military support of her allies. This assumption is not crucial. What I need is that in period 1 the third party is not fully informed about the type of her ally's enemy (i.e., faction B) and her assistance decision in period 1 is taken before faction B moves.[7] Therefore, the third party could already have intervened in the conflict in period 1 at some level of exogenous military assistance that I have set to zero.[8]

In the absence of third-party intervention, each faction has a unit cost of arms equal to 1. Intervention can take various forms and can be modeled in different ways. I follow the formulation in Chang et al. (2007). In particular, when the third party spends M dollars on military assistance to faction A, it affects faction A's unit cost of arms through some reduced-form relationship such that faction A's unit cost of arms is decreases from 1 to $1/(1 + M)^{\theta}$, where θ measures the degree of effectiveness with which a dollar of assistance reduces faction A's unit cost of arming and $0 < \theta < 1$. I discuss the implications of this formulation of third-party intervention in section 3.1.

The timing of actions is as follows:

Period 1 (or the first battle):

Stage 1: Nature chooses faction B's type (valuation): W_H or W_L. This becomes common knowledge to factions A and B but not to the third party. The third party only knows that $Pr(W_H) = q \in (0,1)$. The valuation, V, of faction A is common knowledge.

Stage 2: Faction B chooses his investment in arms.

Stage 3: Faction A observes B's choice and chooses his investment in arms.

[7] This means that I could demonstrate my result in a one-period model with more stages. However, the two-period model below is more convenient for exposition.

[8] In the model, signaling occurs in period 1. Imagine that there is an *exogenous* third-party military assistance in period 1 but no intervention in period 2. That is, the third party has decided to *exogenously withdraw* from the conflict in period 2. Call this the no-intervention case. In this case, there will be no signaling in period 1. The equilibrium of the game in period 1 when the third party has incomplete information is the same as the equilibrium when the third party has complete information. Now suppose in period 2, the third party decides to withdraw or stay in the conflict based on the outcome of the conflict in period 1. Then the analysis in the paper goes through.

Period 2 (or the second battle):

Faction B's valuation in period 1 is also his valuation in period 2 but this is still only known by factions A and B and may remain unknown to the third party.[9] Faction A's valuation is still common knowledge.

Stage 1: The third party chooses how much help to offer faction A.

Stage 2: Faction A observes the third-party's choice and chooses his investment in arms.

Stage 3: Faction B observes the choices of the third-party and faction A and chooses his investment in arms.

The timing of moves between factions A and B in period 1 is not the same as in period 2. This may come across as odd but it should not. In a situation of war, it is not unreasonable to believe that given that B moved first in the current battle (i.e., conflict in period 1) faction A – perhaps to avoid another surprise attack or to exact revenge and also fight for the land – is likely to move first in the next battle (i.e., conflict in period 2).[10] In fact, in a situation of war, there is no presumption that a faction which launched an offensive attack in one battle will necessarily be the first to launch an offensive attack in the next battle. Furthermore, the desire to exact revenge while at the same time trying to gain control over an asset (e.g., land) is consistent with the aforementioned sequence of moves. The faction exacting revenge (call it faction A) in the current battle will be the first mover and the faction (i.e., B) that responds to this attack will be the second mover. This is not inconsistent with faction B having been the first mover in the previous battle. This may be precisely want faction A may try to avoid by moving first in the current

[9]The assumption that faction B's valuation in period 2 is the same as his valuation in period 1 is crucial. If nature were to move again in period 2, there will be no need for signaling in period 1.
[10] I return to the timing of moves in section 3.

battle.[11] Alternatively, we may say that the faction that wins the battle in period 1 will be the first mover in the battle in period 2. Since there is a *positive* probability that faction A will be victorious in period 1 and the third party observes the outcome of the battle in period 1 before making his decision in period 2, it is reasonable to view the above timing of moves as being consistent with the game in period 2 *conditional* on faction A having been victorious in period 1. The current exercise is to prove a possibility result: *expected* third-party intervention could worsen a conflict. But as discussd in section 3.1, different reactions to expected intervention may have different implications.

2.1 Equilibrium in period 2

The game described above is a signaling game with two senders, A and B, and one receiver (i.e., the third party). I look for perfect Bayesian Nash equilibria of this game and restrict attention to pure strategies.

Consider period 2. Given that in period 2, the third party gives her assistance before the factions engage in conflict and that this assistance cannot be withdrawn, it follows that the complete-information version of the conflict will be played in period 2. This is the game analyzed in Chang, Potter, and Sanders (2007). [12]

In stage 3, faction B of type k chooses $X_{k|B}$ to maximize

[11] For a paper that examines revenge in conflicts, see Amegashie and Runkel (2011).

[12] Chang, Potter, and Sanders (2007) used a very standard model of sequential contests. Their contribution was the introduction of third-party intervention into the standard models of Grossman and Kim (1995), Gershenson and Grossman (2000), Leininger (1993) and Morgan (2003). See Konrad (2009) for a review of sequential contests. Assuming complete information and simultaneous moves, this model of conflict has also been use to analyze third-party intervention by Amegashie and Kutsoati (2007) who endogenized a third-party's choice of her ally while Carment and Rowlands (1998), Rowlands and Carment (2006), and Siqueria (2003) took the third-party's ally as given and examined the effect of third-party's intervention on conflicts. Furthermore, there is a small literature on signaling in conflicts and contests. However, this literature considers only two players. It does not consider a third party or third-party intervention and so its focus is entirely different from this paper.

$$\Pi^2_{k|B} = \frac{X_{k|B}}{X_{k|B} + X_{k|A}} W_k - X_{k|B}, \tag{1}$$

where $X_{k|A}$ is the armed investment of faction A when his opponent is faction B of type k, k = H, L.

Then the following Kuhn-Tucker condition must hold:

$$\frac{\partial \Pi^2_{k|B}}{\partial X_{k|B}} = \frac{X_{k|A}}{(X_{k|A} + X_{k|B})^2} W_k - 1 \le 0; X_{k|B} = 0 \text{ if } \frac{\partial \Pi^2_{k|B}}{\partial X_{k|B}} < 0. \tag{2}$$

The condition in (2) implies that the best-response function for B is

$$X_{k|B} = \max[0, \sqrt{W_k X_{k|A}} - X_{k|A}] \tag{3}$$

Therefore, $X_{k|B} = 0$ if $X_{k|A} \ge W_k$. In this case, faction B will not challenge faction A.

In stage 2, faction A facing faction B of type k chooses $X_{k|A}$ to maximize

$$\Pi^2_{k|A} = \frac{X_{k|A}}{X_{k|B} + X_{k|A}} V - \frac{1}{(1+M)^\theta} X_{k|A}, \tag{4}$$

where M is the third-party's assistance to faction A. As will be shown shortly, this will depend on the third-party's belief that faction B is strong (weak).

From (3), put $X_{k|B} = \sqrt{W_k X_{k|A}} - X_{k|A}$ into (4) and simplify to get

$$\Pi^2_{k|A} = V\sqrt{\frac{X_{k|A}}{W_k}} - \frac{1}{(1+M)^\theta} X_{k|A} \tag{4a}$$

Then the unique optimal investment in arms by faction A noting that $X_{k|B} = 0$ if $X_{k|A} \ge W_k$ is

$$\hat{X}_{k|A} = \min\left[W_k, \frac{V^2(1+M)^{2\theta}}{4W_k}\right] \tag{5}$$

Then, $\hat{X}_{k|B} = \sqrt{W_k \hat{X}_{k|A}} - \hat{X}_{k|A}$ which gives faction B of type k's unique armed investment as

$$\hat{X}_{k|B} = \max\left[0, \frac{V(1+M)^{\theta}}{2} - \frac{V^2(1+M)^{2\theta}}{4W_k}\right] \tag{6}$$

Suppose that M = 0. Then using equation (6), $\hat{X}_{k|B} > 0$ if

$$V\left[\frac{1}{2} - \frac{V}{4W_k}\right] > 0 \Rightarrow V < 2W_k, \tag{7}$$

k = H,L. I assume that (7) holds. Hence faction A cannot deter faction B without the assistance of the third party.[13] This also means, from (5), that

$$\hat{X}_{k|A} = V^2(1+M)^{2\theta}/4W_k. \tag{5a}$$

The equilibrium proportion of the land controlled by faction A, whose opponent is faction B of type k, is

$$\hat{P}^2_{k|A} = \frac{\hat{X}_{k|A}}{\hat{X}_{k|A} + \hat{X}_{k|B}} = \frac{V(1+M)^{\theta}}{2W_k}, \tag{8}$$

where $V(1 + M)^{\theta} \leq 2W_L$.[14]

The proportion of the land to faction B of type k in period 2 is

[13] The parameters of my numerical example are such that this feature of the model is preserved. If $V \geq 2W_k$ for all k then there is no conflict even if the third party does not intervene, k = H, L. In this case, third-party intervention is not necessary, which is not a desirable feature of a model of third-party intervention. In this equilibrium, faction A is sufficiently armed (including the number of soldiers) leading faction B to acquiesce resulting in no conflict. See Grossman and Kim (1995) and Gershenson and Grossman (2000) for a discussion of this equilibrium. This equilibrium is not possible if factions A and B move simultaneously.
[14] This means that the third-party's military assistance does not exceed what is required to deter the weak type of faction B. It implies that the strong type of faction B is not deterred in spite of the third-party's assistance. I construct an example that satisfies this condition.

$$\hat{P}^2_{k|B} = 1 - \frac{(1+M)^{\theta} V}{2W_k}, \tag{9}$$

$k = H, L$.

Faction B of type k's equilibrium payoff in period 2 is

$$\hat{\Pi}^2_{k|B} = \frac{\left(W_k - 0.5(1+M)^{\theta} V\right)^2}{W_k}, \tag{10a}$$

and faction A's payoff is

$$\hat{\Pi}^2_{k|A} = \frac{V^2 (1+\hat{M}(q))^{\theta}}{4W_k}, \tag{10b}$$

$k = H, L$.

Clearly, faction B's payoff in (10a) is decreasing in the third-party's military assistance to faction A while faction A's payoff in (10b) is increasing in the third-party's assistance.

Let $\mu \in [0,1]$ be the third-party's belief in period 2 that faction B (i.e., the opponent of her ally) is a strong type. Then, the third party will choose her assistance M to faction A to maximize

$$\Omega^2_C(\mu) = (1-\mu)\hat{P}^2_{L|A} S + \mu \hat{P}^2_{H|A} S - M. \tag{11}$$

Like most formal models of third-party intervention, the third party does not intervene because he cares about the welfare of faction A (his ally). He only cares about his own material welfare. However, he and his ally benefit more from a higher probability of victory or when his ally controls a bigger proportion of the land.

Putting (8) into (11) and maximizing gives the third-party's optimal military assistance as:

$$\hat{M}(\mu) = \left[\frac{\theta SV}{2} \left(\frac{\mu}{W_H} + \frac{1-\mu}{W_L} \right) \right]^{1/(1-\theta)} - 1. \tag{12}$$

I assume that $\hat{M}(\mu) > 0 \ \forall \ \ \mu \in [0,1)$ but $\hat{M}(\mu) \geq 0$ for $\mu = 1$.

Given (12) and $W_H > W_L$, it is obvious that the third-party's military assistance is decreasing in her belief that her ally's enemy is a strong type. The intuition is that the higher is the third-party's belief that faction B is strong, the smaller is the marginal return to her military assistance.[15] Therefore, this result and equation (10) imply that faction B's payoff in period 2 is increasing in the third-party's belief that he is strong. This explains why faction B may overinvest in arms in period 1. This is costly to him in period 1 but beneficial in period 2 because it will cause the third party to reduce her assistance to faction A.

2.2 Separating Equilibrium in period 1

It is important to note that if faction A is choosing G_A to maximize only his payoff in period 1, then he will choose $G_A = \max[0, \sqrt{VG_B} - G_B]$.

Remark: For want of a better expression, I say that faction A invests in signaling if his investment conveys some information to the third party about faction B's type. Otherwise, faction A does not invest in signaling. Given that the third-party's choice of military support is influenced by his belief that faction B is strong and given that his military support occurs in period 2, faction A invests in signaling if and only if when choosing his investment (effort) in period 1, he includes his payoff in period 2 in his optimization problem. Hence, faction A does not invest in signaling if in choosing his

[15] This can easily be seen by putting (8) into (11).

investment in period 1, he maximizes *only* his payoff in period 1. In particular, faction A does not signal if, for a given G_B, he chooses $G_A = \max[0, \sqrt{VG_B} - G_B]$, where $G_A = 0$ if $G_B \geq V$. It is important to note that this argument is also driven by fact that $G_A = \max[0, \sqrt{VG_B} - G_B]$ is independent of faction B's type. However, by "no signaling", I do not mean that faction A is not informed about B's type, but rather that A's strategy in period 1 is not distorted by a concern to influence the third-party's beliefs.

While it is conceivable that, given some beliefs by the third party, $G_A = \max[0, \sqrt{VG_B} - G_B]$ might be an investment level that maximizes faction A's payoff in period 2, it is reasonable to assume that the third party interprets this level of investment as no signaling by faction A because, given G_B, this is what maximizes faction A's period-1 payoff regardless of the third-party's beliefs.

Let $G^*_{k|B}$ be the investment of faction B of type k in period 1 in the benchmark case of full information, k = H, L. This is the full-information level of investment.

Consider period 1. Recall that in this period I assume that the third party cannot give military assistance.

Lemma 1: *In period 1, there is no separating equilibrium in which faction A invests in signaling.*

Proof: Consider a separating equilibrium in which faction B chooses $G'_{k|B}$ and faction A chooses $G_{k|A} \neq \max[0, \sqrt{VG'_{k|B}} - G'_{k|B}]$, where k = H, L. Suppose faction A's opponent is the strong type of B. Then, given a separating equilibrium, the third party believes that faction B is strong (i.e., $\mu = 1$) and chooses the smallest military assistance to faction A. Hence in this equilibrium faction A's payoff in period 2 is at its minimum. Then regardless of the third-party's out-of-equilibrium beliefs, faction A can profitably deviate from this equilibrium by choosing his investment to *only* maximize his period-1 payoff because, given $\mu = 1$, this will not decrease his payoff in period 2. Hence faction A will choose $G'_{H|A} = \max[0, \sqrt{VG'_{H|B}} - G'_{H|B}]$. Therefore, in a separating equilibrium, faction A will not invest in signaling if faction B is strong.

Now suppose that faction A's opponent is the weak type of faction B. Note that if the third party had complete information and faction B was weak, the equilibrium in period 1, would have been ($G^*_{L|B}, G^*_{L|A}$), where $G^*_{L|A} = \max[0, \sqrt{VG^*_{L|B}} - G^*_{L|B}] > 0$.[16] Therefore, in this signaling game if the third party observes the pair ($G^*_{L|B}, G^*_{L|A}$), it reasonable for him to have the belief that $\mu(G^*_{L|B}, G^*_{L|A}) = 0$. Because types are fully revealed in a separating equilibrium, the weak type of B will choose his full-information investment, $G^*_{L|B}$, in period 1 in a separating equilibrium. Faction A, as the second mover, will also choose his full-information investment, in response to $G^*_{L|B}$ because, given $\mu(G^*_{L|B}, G^*_{L|A}) = 0$, a deviation will not yield a more favorable belief by the third party and deviating from his full-information investment will decrease his payoff in period 1. This proves Lemma 1. **QED**.

[16] That $G^*_{L|A} > 0$ is a consequence of the assumption that $V < 2W_L$. I shall construct a numerical example that is consistent with this assumption.

Candidate Equilibrium: There is a separating equilibrium in period 1 in which the strong type of faction B overinvests by choosing $\hat{G}_{H|B} > G^*_{H|B}$ and the weak type chooses $G^*_{L|B}$. Faction A chooses $G^*_{L|A} = \sqrt{VG^*_{L|B}} - G^*_{L|B}$ if faction B is weak and $\hat{G}_{H|A} = \sqrt{V\hat{G}_{H|B}} - \hat{G}_{H|B}$ if faction B is strong.[17] The third-party's equilibrium beliefs in period 2 are $\mu(G^*_{L|B}, G^*_{L|A}) = 0$ and $\mu(\hat{G}_{H|B}, G^*_{H|A}) = 1$. In period 2, her military assistance to faction A when his enemy is the strong type of faction B is $\hat{M}(1)$ and her assistance when faction A's enemy is the weak type of faction B is $\hat{M}(0)$.

Proof: Based on lemma 1, we know that faction A does not signal in a separating equilibrium. So given that faction B chooses $\hat{G}_{H|B} > G^*_{H|B}$ when he is strong and $G^*_{L|B}$ when he is weak, faction A will choose $\hat{G}_{H|A} = \sqrt{V\hat{G}_{H|B}} - \hat{G}_{H|B} > 0$ and $G^*_{L|A} = \sqrt{VG^*_{L|B}} - G^*_{L|B} > 0$. What we need to show is that faction B will indeed choose the investment levels in the candidate equilibrium. By defining a reasonable set of out-of-equilibrium beliefs, we shall also show that faction A will still not signal even if faction B chose investment levels which were different from those in the separating equilibrium.

Standard refinement criteria (e.g., Cho-Kreps intuitive criterion) for signaling games were developed for games with a single sender. In our game, both factions have the same information which the third-party does not have. This is similar to the signaling

[17] Of course, if the strong type of faction B's investment, $\hat{G}_{H|B} \geq V$, then faction A will choose a zero effort in period 1. The reader should be able to verify that this does not affect any of the analyses. However, in the subsequent numerical example, I choose parameters to ensure that faction A's equilibrium effort in period 1 is positive.

game in Schultz (1996) where two political parties (the senders) have private information about the cost of providing a public good that voters (the receivers) do not have or Bagwell and Ramey (1991) in which two incumbents (the senders) in a market have private information about cost which a potential entrant (the receiver) does not have.[18] In dealing with his two-sender signaling game, Schultz (1996, p. 335) observed that

> "The idea behind the intuitive criterion is that if a player makes an out of equilibrium move and there is only one type of that player which could gain by such a deviation, beliefs should put full weight on that type. However, in our game with two parties with the same information, voters will be confronted with one deviator and a party which does not deviate, if a party makes a deviation.[19] What should they think? We will assume that the voters' belief after a deviation is compatible with the non-deviating party's strategy. That is, if only party *a* deviates and party *b* plays a strategy which prescribes different policies in the two states, the voters should infer the true state from *b*'s policy. On the other hand, if party *b*'s policy is the same in the two states votes cannot learn the true state from *b*. In that case, we use the logic of the intuitive criterion, if the deviation is only beneficial for a deviator in a particular state, voters should infer that this is the true state."

In other words, Schultz (1996) found a way of turning a two-sender game into a single-sender game enabling him to adapt the Cho-Kreps intuitive criterion (see also Daughety and Reinganum, 2007). I follow a similar approach.

In particular, when the third party observes any pair, (G_B, $G_A = \max[0, \sqrt{VG_B} - G_B]$), he infers that faction A is not sending a signal. And so, in this case, the third party can credibly say that he will infer faction B's type based on *only* faction B's action. Accordingly, I specify a *benchmark* set of out-of-equilibrium beliefs for the third party. This *benchmark* specifies the third-party's out-of-equilibrium

[18] See also Hertzendorf and Overgaard (2001) and Daughety and Reinganum (2007) where two incumbent firms, with possibly different costs, send signals to consumers. A difference between all these papers and mine is that the two informed parties move simultaneously while in my case they move sequentially.

[19] This is because parties move simultaneously in his model. As I argue below, in my model deviations by the first mover lead to *correlated* deviations because the factions move sequentially.

beliefs conditional on *no* signaling from faction A. *If faction A does not signal*, let the third-party's out-of-equilibrium beliefs be:

$$\mu(G_B, G_A) = \begin{cases} 1 & \text{if} \quad G_B \in [\hat{G}_{H|B}, \infty) \\ \\ 0 & \text{if} \quad G_B \in [0, \hat{G}_{H|B}) \end{cases}, \tag{13}$$

where $G_A = \max[0, \sqrt{VG_B} - G_B]$.[20] The set of beliefs in (13) is common knowledge.

Now if the third party observes G_B and $G_A \neq \max[0, \sqrt{VG_B} - G_B]$, what should he infer? Like Schultz (1996), I follow a logic that is similar to the logic of the intuitive criterion. Faction A knows that if he does not signal, the third party will follow the beliefs in (13). If $G_B \in [0, \hat{G}_{H|B})$, faction A will not signal because, given (13), the third party will believe that faction B is weak even if faction A does not signal. Hence, given (10b) and (12), faction A's equilibrium payoff in period 2 is at its *maximum*. Therefore, given (13), not signaling strictly dominates signaling for faction A if $G_B \in [0, \hat{G}_{H|B})$. Accordingly, faction A has the incentive to signal *if and only if* $G_B \in [\hat{G}_{H|B}, \infty)$ because, given (13), if he does not signal, the third party will believe that faction B is strong. But then, by the logic of the intuitive criterion, the third party should ignore faction A's signal

[20]Notice that in considering deviations by faction B, I allow faction A to respond to the deviations. If the third party were not in the game, it would not matter whether I allow faction A to respond or if faction A sticks to his subgame perfect equilibrium investment. In either case, faction B will still choose his subgame perfect equilibrium investment. In this model, because deviations by faction B affect the third-party's beliefs, which may be different from his equilibrium beliefs, allowing faction A to respond to deviations by faction B may result in a different choice by faction B than what he will choose if faction A stuck to his equilibrium choice. And since deviations are not observed in equilibrium, we are, of course, asking the following (hypothetical) question: "if faction B could deviate from equilibrium, will he, mindful of the fact that faction A chooses his investment after observing his (i.e., B) investment, choose an investment that is different from the equilibrium investment?" For consistency, we have to assume that if faction B deviates, he deviates as a *first mover* whose actions are observed by faction A. In contrast, being a *second mover*, faction A does not have to worry about faction B responding to his (i.e., A) deviations.

and believe that faction B is strong whenever he observes $G_B \in [\hat{G}_{H|B}, \infty)$ and $G_A \neq \max[0, \sqrt{VG_B} - G_B]$. Hence, regardless of faction A's response to G_B, the third party will follow the beliefs in (13) and so faction A will not signal. That is, the third party will infer faction B's type from only faction B's actions.

Faction B's payoff over the two periods is

$$\Theta_k(G_B, \mu(G_B)) = \hat{\Pi}_k^1(G_B) + \hat{\Pi}_k^2(\hat{M}(\mu)), \tag{14}$$

k = H,L.

In what follows, let $\hat{G}_{H|B}$ be implicitly defined by

$$\Theta_L(G_{L|B}^*, 0) = \Theta_L(\hat{G}_{H|B}, 1). \tag{15}$$

Noting that faction A does not signal (i.e., chooses G_A by only maximizing his period-1 payoff), we can write faction B's payoff in period 1 as

$$\hat{\Pi}_k^1(G_B) = \frac{G_B}{G_B + \sqrt{VG_B} - G_B} W_k - G_B = W_k \sqrt{\frac{G_B}{V}} - G_B, \tag{16}$$

for k = L,H.

Given that $\hat{\Pi}_L^1$ is decreasing in G_B for $G_B \in [G_{L|B}^*, \infty)$ and $\hat{\Pi}_L^2(\hat{M}(1))$ is independent of G_B, it follows that $\Theta_L(G_B, 1)$ is decreasing in G_B for $G_B \in [G_{L|B}^*, \infty)$. Then given $\Theta_L(G_{L|B}^*, 0) < \Theta_L(G_{L|B}^*, 1)$, there exists a unique value of $\hat{G}_{H|B}$ that solves equation (15) *and* satisfies $\hat{G}_{H|B} > G_{L|B}^*$.

I assume that $\hat{G}_{H|B} > G_{H|B}^*$. This holds if $\hat{M}(0)$ is sufficiently bigger than $\hat{M}(1)$. This is because, in this case, the gain to the weak type of pretending to be strong --

which is directly proportional to the difference $\hat{M}(0)-\hat{M}(1)$ -- is so high that it requires a high value of $\hat{G}_{H|B}$ to make him indifferent between his equilibrium payoff and deviating to $\hat{G}_{H|B}$ to obtain $\Theta_L(\hat{G}_{H|B},1)$. It is easy to show that

$$\hat{M}(0)-\hat{M}(1)=\left(\frac{\theta SV}{2}\right)^{1/(1-\theta)}\left(\left(\frac{1}{W_L}\right)^{1/(1-\theta)}-\left(\frac{1}{W_H}\right)^{1/(1-\theta)}\right). \tag{17}$$

The magnitude of (17) depends on the size of the difference between W_H and W_L and the sizes of S and V.

Given (13), a separating equilibrium exists if

$$\Theta_H(\hat{G}_{H|B},1)\geq\Theta_H(G_B,\mu(G_B)), \tag{18}$$

and

$$\Theta_L(G^*_{L|B},0)\geq\Theta_L(G_B,\mu(G_B)) \tag{19}$$

for all G_B.

The weak type of faction B has no incentive to reduce his investment below $G^*_{L|B}$ because, given the beliefs in (13), he reduces his payoff in period 1 without changing his payoff in period 2. Given (15) and (13), the weak type of faction B has no incentive to increase his investment beyond $G^*_{L|B}$. Therefore, the inequality in (19) holds. The strong type of faction B has no incentive to increase his investment beyond $\hat{G}_{H|B}$ because, given the beliefs in (13), this reduces his payoff in period 1 without increasing his payoff in period 2. And I assume that reducing his investment below $\hat{G}_{H|B}$ is not profitable given that this will cause the third-party's belief that he is strong to discontinuously fall from 1 to 0. That is, $\Theta_H(\hat{G}_{H|B},1) \geq \Theta_H(G_B,0)$ for $G_B < \hat{G}_{H|B}$.[21]

Under these conditions, the inequality in (18) holds.

I need to argue that this equilibrium is supported by reasonable out-of-equilibrium beliefs. Because of previous arguments that the third party will infer faction B's type by using only faction B's actions, I can easily adapt the Cho-Kreps "intuitive criterion" to place restrictions on out-of-equilibrium beliefs in this signaling game. The "intuitive criterion" requires that out-of-equilibrium beliefs put no weight on types that have no incentive to deviate from a given equilibrium no matter what the third party would conclude from observing the deviation. Given that the payoff of each type of faction B is strictly increasing in μ, it follows that faction B's equilibrium action dominates any out-of-equilibrium action if his equilibrium payoff is higher than any out-of-equilibrium payoff even if such an out-of-equilibrium action causes the third party to believe that faction B is a strong type (i.e., $\mu = 1$). Accordingly, the separating equilibrium above satisfies the Cho-Kreps "intuitive criterion" if there is no G_B such that

$$\Theta_H(G_B,1) > \Theta_H(\hat{G}_{H|B},1) \text{ and } \Theta_L(G_B,1) < \Theta_L(G^*_{L|B},0), \tag{20a}$$

and

$$\Theta_H(G_B,1) < \Theta_H(\hat{G}_{H|B},1) \text{ and } \Theta_L(G_B,1) > \Theta_L(G^*_{L|B},0). \tag{20b}$$

In a separating equilibrium, the strong type enjoys the most advantageous belief (i.e., $\mu = 1$) by the third party so, given that $\hat{G}_{H|B} > G^*_{H|B}$, he has no incentive to deviate to an investment level that is greater than $\hat{G}_{H|B}$. Hence, any deviation must be to a lower

[21] The numerical example below satisfies this condition with strict inequality. This was verified by plotting $f(G_B) \equiv \Theta_H(\hat{G}_{H|B},1) - \Theta_H(G_B,0)$ on the domain [0, $\hat{G}_{H|B}$].

level of investment. Note that the strong type of faction B will find it profitable to deviate to some $G_{H|B} \in [G^*_{H|B}, \hat{G}_{H|B})$ if the third party will still believe that he is strong. Now since $\Theta_L(G^*_{L|B}, 0) = \Theta_L(\hat{G}_{H|B}, 1)$ as given in (15) and given that $\Theta_L(G_B, 1)$ is decreasing in G_B for $G_B \in [G^*_{H|B}, \infty)$, it follows that $\Theta_L(G^*_{L|B}, 0) < \Theta_L(G_{L|B}, 1)$ for $G_{L|B} \in [G^*_{H|B}, \hat{G}_{H|B})$. Hence, if the strong type of faction B finds a deviation to $G_{H|B} \in [G^*_{H|B}, \hat{G}_{H|B})$ profitable, then the weak type also will find it profitable. Therefore, the beliefs satisfy the Cho-Kreps intuitive criterion for this set of deviations.

Now consider the deviations $G_{H|B} \in [0, G^*_{H|B})$ by the strong type of faction B. Note, from (16), that $\partial\hat{\Pi}^1_H / \partial G_B > \partial\hat{\Pi}^1_L / \partial G_B$ for any given G_B. Therefore, even if the third party still holds the most favorable belief (i.e., $\mu = 1$) for faction B, a reduction of investment in period 1 below $G^*_{H|B}$ hurts the strong type of faction B than it hurts the weak type and, in some cases, it is even beneficial to the weak type because $G^*_{L|B} < G^*_{H|B}$. Therefore, it is reasonable for the third party to believe that such reductions are by the weak type as in equation (13).[22] Hence, the beliefs in (13) satisfy the intuitive criterion.

[22]In this case, the intuitive criterion actually does not tell us what to do. However, the reasoning used here is in the spirit of the D1 condition in Cho and Kreps (1987). The D1 condition requires that we put the *entire* weight on the type that is willing to deviate for a wider range of inferences by the receiver (i.e., uninformed party). In my case, while a decrease (deviation) that the strong type of faction B finds profitable is also profitable to the weak type, the converse is not true. That is, there are some decreases in investment that the weak type finds profitable but are not profitable to the strong type. Hence, it is reasonable for the third party to set $\mu = 0$ in (13) for investments smaller than $\hat{G}_{H|B}$. This is why this kind of reasoning is in the spirit of the D1 condition of Cho and Kreps (1987).

Like any signaling game, this one has multiple equilibria. In particular, there are separating equilibria in which the strong type of faction B chooses $G_{H|B} > \hat{G}_{H|B}$ and the weak type chooses $G^*_{L|B}$. However, for any $G_{H|B} > \hat{G}_{H|B} > G^*_{H|B}$, the strong type of faction B is strictly better off by deviating to $\hat{G}_{H|B}$ if the third party will believe that he is strong. And given (15), this deviation is not profitable for the weak type even if the third party believes that he is strong. Therefore, it is reasonable for the third party to believe that such deviations are by the strong type.[23] Therefore, the strong type of faction B will deviate from any separating equilibrium where $G_{H|B} > \hat{G}_{H|B}$.

Now given (15) and the fact that the weak type does not signal in a separating equilibrium, there cannot be a separating equilibrium with $G_{H|B} < \hat{G}_{H|B}$. Hence, the candidate equilibrium above is the only separating equilibrium.

Suppose that the third-party's valuation, S > 0, is so high that $\hat{M}(1)$ is such that $W_H - 0.5(1+\hat{M}(1))^{\theta} V \leq 0$. Then $W_L - 0.5(1+\hat{M}(0))^{\theta} V < 0$ because $\hat{M}(0) > \hat{M}(1)$ and $W_H > W_L$. Then the third party will give enough military support to faction A to deter faction B in period 2 regardless of faction B's type. Then faction B gains nothing by signaling in period 2, so he will not display any bravado. Therefore, faction B will only display bravado if the third party is not too strong or will not intervene in a big way.[24] Note also that $W_H - 0.5(1+\hat{M}(1))^{\theta} V \leq 0$ is likely to hold if V is sufficiently high.

This gives the following proposition:

---

[23] In this case, the inequalities in (20a) are applicable with the inequality for the weak type being a weak inequality.

[24] In other cases, the requirement to keep civilian causalities at a minimum means that a third party and her ally cannot deploy their full and combined military might.

Proposition 1: *If (i) the third party and his ally are not too strong or if the third party is not willing to intervene in a big way (i.e.,* $W_H - 0.5(1+\hat{M}(1))^{\theta} V > 0$ *),*[25] *(ii) the solution to equation (15) is such that* $\hat{G}_{H|B} > G^*_{H|B}$ *, and (iii) the inequalities in (18) and (19) hold, then there exists a unique*[26] *perfect Bayesian separating equilibrium with bravado that satisfies an adapted intuitive criterion. In this equilibrium, the strong type of faction B is too aggressive by choosing a unique armed investment that is higher than his full-information level of investment in order to induce the third party to reduce his military assistance to faction A (i.e., the third-party's ally) or back off from intervening in the conflict. The weak type of faction B chooses his unique full-information investment.*

Example: Suppose that S = 8, $W_H = 4$, $W_L = 3.2$, V = 3 and $\theta = 0.5$. Then $G^*_{H|B} =$ $0.25(W_H)^2/V =$ 1.333 and equation (15) gives $\hat{G}_{H|B} = 1.9842 > G^*_{H|B}$.

$\Theta_H(\hat{G}_{H|B},1) = 2.034$, $\Theta_L(G^*_{L|B},0) = 0.900, \hat{M}(0) = 2.516, \hat{M}(1) = 1.250, G^*_{L|B} = 0.853$, $G^*_{L|A} = 0.746$ and $\hat{G}_{H|A} = 0.455$. Also, $\Theta_H(\hat{G}_{H|B},1) \geq \Theta_H(G_B,0)$ for $G_B < \hat{G}_{H|B}$.

Given the third-party's military assistance, all equilibrium investment levels and payoffs in period 2 are positive.

[25] For example, if $\theta = 0.5$, this condition is $8(W_H)^2 - SV^2 > 0$.

[26] By this, I mean that focusing on only separating equilibria, the equilibrium in proposition 1 is unique. Under certain conditions, I am able to show rule out pooling equilibria. The proof is available on request. However, focusing on separating equilibria is sufficient for my purposes. This is not a limitation of this paper because its primary goal is to construct an equilibrium in which third-party intervention may worsen a conflict. Therefore, what matters is to show that such an equilibrium exists.

2.4 Aggregate destruction or fatalities in the conflict

As a proxy for aggregate fatalities or destruction in the conflict, I use the aggregate effort (investment) in the conflict. Using aggregate destruction in the conflict is consistent with the criterion used in empirical works, discussed in section 1, which claim that third-party intervention worsens conflicts.

Suppose the conditions in proposition 1 hold. When faction B is strong, the aggregate effort (investment) in period 1 in the separating equilibrium is $\hat{G}_{H|A} + \hat{G}_{H|B} = \sqrt{V\hat{G}_{H|B}} > \sqrt{VG^*_{H|B}}$. Therefore, intervention by the third party increases aggregate effort in period 1. This is the moral hazard effect of third-party intervention which stems from *expected* intervention. If when faction B is weak, aggregate effort in period 1 is $G^*_{L|A} + G^*_{L|B} = \sqrt{VG^*_{L|B}}$.

Note that there is no separating equilibrium in which the third-party's intervention reduces aggregate effort in the conflict. If this were the case, then $\hat{G}_{H|B} \in (G^*_{L|B}, G^*_{H|B})$ would have to be supported as part of a separating equilibrium. However, in this case, the strong type of faction B will be better off by deviating to $G_{H|B} > \hat{G}_{H|B}$ if the third party will still believe that he is strong. Given (15), the weak type of faction A will not deviate even if he will be perceived as strong. Hence, any separating equilibrium with $\hat{G}_{H|B} \in (G^*_{L|B}, G^*_{H|B})$ does not survive the intuition criterion.

In period 2, the aggregate effort (investment) is $\hat{X}_{k|A} + \hat{X}_{k|B} = \sqrt{W_k\hat{X}_{k|A}}$, where $\hat{X}_{k|A}$ is given by equation (5a). Then given that $\hat{X}_{k|A}$ is increasing in the third-party's military assistance regardless of faction B's type, it follows that the third-party's intervention increases the aggregate effort in period 2. This is the moral hazard effect of third-party intervention which stems from *actual* intervention (discussed in section 1).

These results give the following proposition:

Proposition 2*: Suppose the conditions in proposition 1 hold, then third-party intervention may worsen the conflict. In particular, in period 1, the third-party's* ***expected*** *intervention worsens the conflict when faction B is strong but has a neutral effect when faction B is weak. The third-party's* **actual** *intervention worsens the conflict in period 2 regardless of faction B's type.*

In the model, the third party is aware that his intervention increases the cost of the conflict (i.e., proposition 2). However, his objective is to maximize his payoff as given by equation (11). The third-party's intention was not to worsen the conflict neither did he intend to improve it. However, his intervention based on his own economic interest led to a perverse incentive effect in period 1, which the third party was aware of. This is consistent with the definition of the moral hazard of third-party intervention (e.g., Grigorian, 2005).

The third-party's goal is to maximize his payoff in (11) based on the best information about faction B's type. His military assistance in the separating equilibrium is based on full information about faction B's type and his payoff occurs in only period 2 where the factions always play the full-information equilibrium. Therefore, no other strategy by the third party dominates the one studied in this paper.[27,28]

In the separating equilibrium, the strong type of faction B displays bravado because of the third-party's limited information about faction B. Furthermore, the third party can nullify this behavior if she can commit to a given level of assistance based on her prior beliefs. This can be summarized in the following corollary:

Corollary 1: *The expectation of a third-party's assistance to an ally coupled with the third-party's limited information about the strength of the enemy of her ally can be strategically exploited by the enemy. However, the ability of the enemy to strategically gain from his superior information no longer exists if the third party can commit to a given level of assistance based on her prior beliefs.*

It is noteworthy that although faction A knows faction B's type, he cannot communicate this information to his ally (i.e., the third party) to avoid the over-aggressive behavior of faction B. A reason is that faction A's information is "soft information" not "hard information". But more importantly, given this "soft information, the third party cannot believe whatever faction A says because faction A has the incentive

[27] The evaluation of welfare in a game of incomplete information is not straightforward. For example, in evaluating the third-party's welfare, should we use his *ex ante* welfare or *ex post* welfare (see, Holmstrom and Myerson, 1983). In my case, it is *ex post* welfare because the third party decides on his military assistance in the second period after he has updated information, if any, based on the factions' armed investments observed in the first period. One can use the envelope theorem to show that the payoff function in (11) is decreasing in μ (i.e., the third-party's belief that faction B is strong) which means that the third-party's payoff is minimized if he believes with certainty (i.e., $\mu = 1$) that faction B is strong. However, using this logic to evaluate welfare fails to recognize that any equilibrium beliefs other than those in a separating equilibrium may be false although they may be rational. For example, using this logic the third party might choose to always hold the possibly false belief that faction B is weak in order to maximize his payoff. In fact, if beliefs are subjective, rational ignorance may be bliss. My model is based on objective beliefs. I assume that third party prefers the truth and so if a separating equilibrium is achievable, he prefers this equilibrium because his beliefs are not only rational but also true.

[28] For example, one may argue that alternatively the benchmark out-of-equilibrium beliefs could be defined with respect to what the third party will do if faction B does not signal. However, if faction B does not signal in which case the different types of faction B choose their unique full-information investment, then the third party does not need faction A's signal to infer faction B's type. And if instead, at least, one type of faction B signals, then following the approach above makes sense because it simplifies the third-party's signal extraction problem and ultimately leads to full revelation of types.

to lie about faction B's type when faction B is strong. Furthermore, faction B's first-mover advantage in period 1 makes it difficult for faction A to credibly convey his information to the third party. This leads to the following proposition:

Proposition 3: *Because of the ally's incentive to lie to the third party, it is optimal for the third party to rationally ignore the private and valuable information of his ally and instead base his decisions solely on the costly signals of the ally's enemy.*

3. Further discussion and robustness of results

3.1 On the effects of and nature of third-party intervention

As shown above, the third-party's actual intervention in period 2 makes his ally too aggressive and thus escalates the conflict. This effect of actual intervention is not the focus of this paper. The focus of this paper is on expected intervention.

In this paper, a third-party's *expected* intervention causes his *ally's enemy* to be more aggressive. However, there are instances where it can be plausibly argued that a third-party's expected intervention can cause his *ally* to be more aggressive. For example, rebel groups in Libya may have fought harder in order to convince Western powers that providing them air support may create a significant chance of Gaddafi's removal. Also, after the 1991 Kuwait War, the Kurds and Shias rose up against Saddam Hussein expecting American military support.

However, the effects of expected intervention in the preceding paragraph cannot explain the argument by some scholars that a lack of resolve and credibility within coalitions over the use of force create incentives for the escalation of conflicts (Regan, 1996; Diehl et al., 1996; Harvey, 1998; Walter and Snyder, 1999). In fact, the effect of expected intervention in the preceding paragraph is likely to arise for precisely the opposite reason: the expectation by the *ally* that the third party has a strong resolve and so will intervene in a big way. As indicated above in the case of rebel groups who fought Gaddafi in Libya, if the ally expects the third party to intervene in a big way, he may fight harder in order to convince the third party to intervene.

The present model yields an effect of expected intervention that crucially depends on whether the third party intervenes in a big way and it is driven by the reaction of *ally's enemy*. Unlike the effect in the preceding paragraph, the ally's enemy (not the ally) fights harder if the third party is not expected to intervene in a big way and does not change his behavior if the third party intervenes in a big way. The analyses also show that a third party may ignore the signals by an informed ally, a result that cannot be obtained in a model which predicts that a third-party's expected intervention will make his ally more aggressive. It is pointless for the ally to use extra aggression to signal to the third party if these signals will be ignored.

One may argue that the third party (e.g., the USA) may be directly involved in the conflict by committing troops and so military assistance need not take the form of reducing the ally's cost of conflict. In a previous version of this paper, I showed that this would not affect the results.[29] What is important for my results is not whether the third party is directly involved in the conflict. What I need is that the third-party's assistance or effort is decreasing in his belief that the enemy is strong. Moreover, third parties do intervene in conflicts by giving financial support for military expenditure without directly getting involved in the conflict (e.g., the USA's support of Angola's UNITA rebels during the Cold War). Elbadawi and Sambanis (2000) define external intervention "... as a unilateral intervention by one (or more) third party government(s) in a civil war in the

[29] The proof is available on request.

form of military, *economic* or mixed assistance in favor of either the government or the rebel movement involved in the civil war."

There is nothing in the model which suggests that the third-party's assistance is not an in-kind transfer. The third-party's assistance could involve military training, the provision of intelligence information, military equipment, etc. These could have the effect of reducing his ally's cost of effort.

3.2 An alternative analysis of the third-party's behavior

In the present model, the third-party's assistance is decreasing in his belief that his ally's enemy is strong. As explained earlier, this is because the return to his investment (assistance) is smaller when his ally's enemy is strong. However, another plausible scenario is that the third party is more likely to assist his ally when the ally's enemy is strong. For example, the benefits of defeating a strong enemy might be higher because there may otherwise be a higher probability of such an enemy challenging the third-party's ally in the future if he is not defeated soundly in the current period. Or losing to a strong enemy makes it more difficult to fight back in the future. That is, if you lose to a strong enemy you are less likely to get a second chance relative to losing to a weak enemy. In either case, the enemy has the incentive to feign weakness (i.e., under-invest in arms) rather than display bravado (overinvest in arms).

It is possible to model how the aforementioned higher future benefit of defeating a stronger enemy may arise. However, I simply consider a reduced-form version of this argument by assuming that the third-party's benefit when a strong enemy is defeated is

S_H and when a weak enemy is defeated is S_L, where $S_H \geq S_L > 0$. Then, analogous to equation (11), the third party will choose her assistance M to faction A to maximize

$$\tilde{\Omega}_C^2(\mu) = (1-\mu)\frac{V}{2W_L}(1+M)^{\theta}S_L + \mu\frac{V}{2W_H}(1+M)^{\theta}S_H - M \quad (21)$$

The third-party's optimal military assistance is:

$$\tilde{M}(\mu) = \left[\frac{\theta V}{2}\left(\mu\frac{S_H}{W_H} + (1-\mu)\frac{S_L}{W_L}\right)\right]^{1/(1-\theta)} - 1 \quad (22)$$

Then $\partial\tilde{M}/\partial\mu > 0$ if $S_H/W_H > S_L/W_L$. Given that $W_H > W_L$, a necessary condition for this result is $S_H > S_L$. Clearly, if the benefit of defeating a strong enemy is higher than the benefit of defeating a weak enemy, then it is possible that the third party's assistance to his ally is increasing in his belief that his ally's enemy is strong. Then we can construct an equilibrium in which the ally's enemy underinvests in arms in period 1. Also, when $S_H/W_H = S_L/W_L$, then $\partial\tilde{M}/\partial\mu = 0$. Then there will be no signaling by either faction and third-party intervention will not worsen the conflict.

Of course, in this paper, I have assumed that $S_H = S_L = S$. Or more generally, I have assumed that $S_H/W_H < S_L/W_L$. In this case, equilibria with bravado exist and third-party intervention may worsen the conflict.

3.3 The timing of moves in periods 1 and 2

As mentioned previously, faction B moves before faction A in period 1 (i.e., the first battle) while faction A moves before faction B in period 2 (i.e., the second battle). In section 2, I provided a number of reasons to justify this timing of moves. Still, I wish to point out an implication of this sequence of moves.

It turns out that the third-party's military assistance in period 2 is increasing in his belief that faction B is strong if faction B moves before faction A in period 2. With this timing, the third-party's optimal assistance is

$$\widehat{M}(\mu) = \left(\frac{\theta S}{2V} \left(\mu W_H + (1-\mu) W_L \right) \right)^{1/(1+\theta)} - 1, \tag{23}$$

where it is immediately obvious that, given $W_H > W_L$, $\partial \widehat{M}(\mu)/\partial \mu > 0$. In this case, the strong type of faction B would rather feign weakness. We cannot construct equilibria with bravado in this case.

It is important to note that the assumption that faction A moves before faction B in period 2 does not bias the analysis towards one particular result because as shown in subsection 3.2, the relationship between the third-party's belief that faction B is strong and the size of his military assistance is different depending on the parameters of the model (i.e., depends on the sign of $S_H/W_H - S_L/W_L$). Therefore, the result in proposition 1 simply gives conditions under which third-party intervention might worsen a conflict.

3.4 Faction A is the first mover in period 1

The game is harder to solve if faction A is the first mover in period 1. In this case, defining the benchmark out-of-equilibrium beliefs is no longer straightforward because it is not clear what it means to claim that the second mover is not signaling. It is important to reiterate that the previous argument that faction A, as the second mover, does not signal if he chooses $G_A = \max[0, \sqrt{VG_B} - G_B]$ and that it is rational for the third party to base his beliefs on only faction B's actions is correct because $G_A = \max[0, \sqrt{VG_B} - G_B]$ is independent of faction B's type. This is why faction A's choice of

$G_A = \max[0, \sqrt{VG_B} - G_B]$ cannot convey any information to the third party. In contrast, even if faction B, as a second mover, chooses G_B to maximize only his payoff in period 1, this *might* be a signal to the third party because he will choose

$G_{k|B} = \max[0, \sqrt{W_k G_A} - G_A]$ which is clearly dependent on his type. However, faction A, as the first mover, could now take actions to nullify faction B's signals. For example, suppose the third party were to say that, for a given G_A, if he observes $G_B =$ $\sqrt{W_H G_A} - G_A > 0$, then he will believe that faction B is strong. But then faction A could choose G_A such that $\max[0, \sqrt{W_H G_A} - G_A] = 0$ which implies that $\max[0, \sqrt{W_L G_A} - G_A] = 0$, so the third-party cannot infer faction B's type. All this shows that if faction A moves first in period 1, the game is more complicated. Therefore, the timing of moves was also chosen to make the model tractable. Among others, it led to useful insights such as the conditions under which the third party will rationally ignore the signals of his ally.

4. Conclusion

Expected third-party intervention may have a perverse effect on conflicts. This paper has shown that this may occur through a display of bravado by the enemy of the third-party's ally. However, as discussed in section 3, not all equilibria would be characterized by a display of bravado by faction B (enemy of the third-party's ally). Depending on the parameters of the model, one can construct a separating equilibrium in which faction B underinvests in arms. And even in the separating equilibrium stated in proposition 2, bravado does not occur if the enemy of the third-party's ally is weak.

When faction B is weak, the *expected* third-party's intervention does not worsen the conflict in period 1. Hence, using signaling as a possible mechanism through which a third-party's intervention might lead to a perverse outcome, the paper sheds some light on the conditions under which *expected* third-party intervention may or may not worsen conflicts. Indeed, if the third party will not actually intervene in a big way, then it might be better if the third party does not intervene at all.

It is also noteworthy that the third party may rationally ignore the private information of his ally which, as explained in section 1, appears to be consistent with claims by the USA that its allies (i.e., Pakistan, Iraq, and Afghanistan) are not doing enough to rein in its enemies. And, as explained in section 1, the results of this paper have implications for econometric work that attempt to estimate the effect of third-party interventions on conflicts.

It should be obvious from the analysis that the sequential moves in period 1 and the fact that faction A could perfectly observe faction B's choice simplified the third-party inference problem. The problem would have been a lot harder if factions A and B moved simultaneously in period 1. Solving the game in this case is left as an exercise for future research.

References

Akcinaroglu, S., and Radziszewski, E. (2005) Expectations, Rivalries, and Civil War Duration. *International Interactions* 31: 349-374.

Amegashie, J.A., and Kutsoati, E. (2007). (Non)intervention in Intra-State Conflicts. *European Journal of Political Economy* 23: 754-767.

Amegashie, J.A., and Runkel, M. (2011). The Paradox of Revenge in Conflicts. *Journal of Conflict Resolution*, forthcoming.

Bagwell, K., and Ramey, G. (1991). Oligopoly Limit Pricing. *Rand Journal of Economics* 22: 155–172.

Betts, R.K. (1996). The Delusion of Impartial Intervention. In: Cocker, C.A., Hampson, F.O. (Eds). *Managing Gobal Chaos: Sources of and Responses to International Conflict*. United States Institute of Peace Press, Washington DC, pp. 333-343.

Blattman, C., and Miguel, E. (2010). Civil War. *Journal of Economic Literature* 48: 3 - 57.

Carment, D., and Harvey, F. (2000). *Using Force to Prevent Ethnic Violence: An Evaluation of Theory and Evidence*. Westport, CT: Praeger.

Carment, D., and Rowlands, D. (1998). Three's Company: Evaluating Third-Party Intervention in Intrastate Conflict. *Journal of Conflict Resolution* 42: 572-599.

Chang, Y.-M., Potter, J., and Sanders, S. (2007). War and Peace: Third-party Intervention in Conflict. *European Journal of Political Economy* 23: 954-974.

Cho, I-K., and Kreps, D. (1987). Signaling Games and Stable Equilibria. *Quarterly Journal of Economics* 102: 179-221.

Clark, D.J., and Riis, C. (1998). Competition over more than one Prize. *American Economic Review* 88: 276-289.

Collier, P., and Hoeffler, A. (2007). Civil War. In *Handbook of Defense Economics*. K. Hartley and T. Sandler (Eds). Elsevier North Holland.

Crawford, T.W. (2005). Moral hazard, intervention and internal war: A conceptual analysis. *Ethnopolitics* 4: 175 - 193.

Daughety, A.F., and Reinganum, J.F. (2007). Competition and Confidentiality: Signaling Quality in a Duopoly when there is Universal Private Information. *Games and Economic Behavior* 58: 94-120.

Diehl, P., Reifschneider, F. J., and Hensel, P. R. (1996). United Nations Intervention and Recurring Conflict. *International Organization* 50: 683 - 700.

Dixon, W.J. (1996). Third-Party Techniques for Preventing Conflict Escalation and Promoting Peaceful Settlement. *International Organization* 50: 653-81.

Elbadawi, I., and Sambanis, N. (2000). External Intervention and the Duration of Civil Wars. *World Bank Policy Research Paper 2433*.

Epstein, G.S., and Gang, I.N. (2009). Good Governance and Good Aid Allocation. *Journal of Development Economics* 89: 12-18.

Favretto, K. (2009). Should Peacemakers Take Sides? Major Power Mediation, Coercion, and Bias. *American Political Science Review* 103: 248-63.

Fearon, J. D. (1994). Domestic Political Audiences and the Escalation of International Disputes. *American Political Science Review* 88: 577 - 592.

Fearon, J. D. (1995). Rationalist Explanations for War. *International Organization* 49 :379– 414

Gershenson, D., and Grossman, H.I. (2000). Civil Conflict: Ended or Never Ending? *Journal of Conflict Resolution* 44: 807-821.

Grigorian, A. (2005). Third-party intervention and escalation in Kosovo: Does moral hazard explain it?. *Ethnopolitics* 4: 195 - 213.

Grossman, H.I., and Kim, M. (1995). A Theory of the Security of Claims to Property. *Journal of Political Economy* 103: 1275-1288.

Harvey, F. (1998). Deterrence Failure and Ethnic Conflict: The Case of Bosnia. in *Peace in the Midst of Wars:Preventing and Managing International Ethnic Conflicts*, ed. David Carment and Patrick James, Columbia, SC: University of South Carolina Press, 1998.

Heger, L., and Salehyan, I. (2007). Ruthless Rulers: Coalition Size and the Severity of Civil Conflict. *International Studies Quarterly* 51: 385 - 403.

Hertzendorf, M.N., and Overgaard, P. B. (2001). Price Competition and Advertising Signals. *Journal of Economics and Management Strategy* 10: 621-662.

Holmstrom, B., and Myerson, R. B. (1983). Efficient and Durable Decision Rules with Incomplete Information. *Econometrica* 61: 1799-1819.

Konrad, K.A.(2009). *Strategy and Dynamics in Contests*. New York: Oxford University Press.

Kuperman, A. (1996). The Other Lesson of Rwanda: Mediators Sometimes do More Damage Than Good. *SAIS Review* 16: 221-40.

Lacina, B (2006). Explaining the Severity of Civil Wars. *Journal of Conflict Resolution* 50: 276 - 289.

Leininger, W. (1993). More Efficient Rent-Seeking – a Munchhausen Solution. *Public Choice* 75: 43-62.

Morgan, J. (2003). Sequential Contests. *Public Choice* 116: 1-18.

Morrow, J.D. (1992). Signaling Difficulties with Linkage in Crisis Bargaining. *International Studies Quarterly* 36: 153 - 172.

Regan, P. (1996). Conditions for Successful Third Party Intervention in Intrastate Conflicts. *Journal of Conflict Resolution* 40: 336-359.

Regan, P. M. (2000). *Civil Wars and Foreign Powers: Outside Intervention in Intrastate Conflict*. Ann Arbor, Michigan: University of Michigan Press.

Regan, P. M. (2002). Third-Party Intervention and the Duration of Intrastate Conflicts. *Journal of Conflict Resolution* 46: 55-73.

Regan, P.M., and Stam, A. (2000). In the Nick of Time: Conflict Management, Mediation Timing, and the Duration of Interstate Disputes. *International Studies Quarterly* 42: 239-260.

Rowlands, D., and Carment, D. (1998). Moral hazard and conflict intervention. In *The political economy of war and peace*, ed. Murray Wolfson. Norwell, MA: Kluwer Academic Publishers.

Rowlands, D. and Carment, D. (2006) Force and Bias: Towards a Predictive Model of Effective Third-Party Intervention. *Defence and Peace Economics* 17: 435-456.

Schultz, C. (1996). Polarization and Inefficient Policies. *Review of Economic Studies*, 63: 331–344.

Siqueira, K. (2003). Conflict and Third-Party Intervention. *Defence and Peace Economics* 14: 389-400.

Wagner, R. H. (2007). *War and the State: the Theory of International Politics*. Ann Arbor, MI: University of Michigan Press.

Walter, B. F. (2009). Bargaining Failures and Civil War. *Annual Review of Political Science* 12: 243-261.

Walter, B., F., and Snyder, J. (1999). *Civil Wars, Insecurity, and Intervention*. New York: Columbia University Press.

Watkins, M., and Winters, K. (1997). Intervenors with Interests and Power. *Negotiation Journal* 13: 119-142.

Werner, S. (2000). Deterring Intervention: The Stakes of War and Third-Party Involvement. *American Journal of Political Science* 44: 720-732.

Yuen, A. (2009). Target Concessions in the Shadow of Intervention. *Journal of Conflict Resolution* 53: 745-773.